ELSEVIER

Journal of Chromatography B, 782 (2002) 57–66

JOURNAL OF CHROMATOGRAPHY B

www.elsevier.com/locate/chromb

Review

Mass spectrometry and tandem mass spectrometry, alone or after liquid chromatography, for analysis of polymerase chain reaction products in the detection of genomic variation

James J. Walters, Karen F. Fox, A. Fox*

Department of Microbiology and Immunology, University of South Carolina, School of Medicine, Columbia, SC 29208, USA

Abstract

The availability of the sequences of entire bacterial and human genomes has opened up tremendous opportunities in biomedical research. The next stage in genomics will include utilizing this information to obtain a clearer understanding of molecular diversity among pathogens (helping improved identification and detection) and among normal and diseased people (e.g. aiding cancer diagnosis). To delineate such differences it may sometimes be necessary to sequence multiple representative genomes. However, often it may be adequate to delineate structural differences between genes among individuals. This may be readily achieved by high-throughput mass spectrometry analysis of polymerase chain reaction products.

Keywords: Reviews; Genomics; Polymerase chain reaction products

Contents

1. Introduction

There are numerous established molecular biology techniques for detection of genomic variation. Changes in properties that indirectly reflect structural changes are assessed. These approaches include (1) detection of mobility changes most commonly involving electrophoresis (e.g. agarose gel electrophoresis, single strand conformation polymorphism or capillary electrophoresis), (2) hybridization characteristics (e.g. DNA microarrays) or (3) enzymatic recognition of sequence variation (e.g. restriction

*Corresponding author. Tel.: +1-803-733-3288; fax: +1-803-733-3192.

E-mail address: afox@med.sc.edu (A. Fox).

1570-0232/02/$ – see front matter
PII: S1570-0232(02)00563-9

enzymes [sequence changes] or cleavases [secondary structure]). These techniques have been reviewed numerous times by others [1–5] and are beyond the scope of the current review.

This review instead will focus on the use of mass spectrometry as a stand-alone technique or in conjunction with on-line high-performance liquid chromatography (LC). Mass spectrometry directly detects simple structural changes including single nucleotide substitutions or deletions/insertions (e.g. single nucleotide polymorphisms, SNPs, in normal individuals) whilst tandem mass spectrometry identifies more complex sequence changes (e.g. as may arise in cancer). The most widely used approach allowing focus on the genetic region of interest is the polymerase chain reaction (PCR), which exponentially amplifies a specific section of DNA from a genome. Thus the work discussed here primarily involves mass spectrometry analysis of PCR products.

Some genes are present as a single copy, others exist in multiple copies and may display sequence differences. For example, the ribosomal RNA (rRNA) operon in bacteria is present in varying copy number [6], e.g. *B. subtilis* sub-group 168 contains 10 operons, some of which lack a 3 bp, ''TTA'' insert [7]. In higher order genomes (e.g. human) there are generally two copies. These alleles can be identical (homozygotes) or have sequence distinctions (heterozygotes) [8]. Thus, there must be adequate mass spectral resolution to detect the presence of two operons [7] or both alleles [8] displaying minor sequence variation.

In some circumstances, there may be multiple differences among copies, creating a complex sample difficult to resolve by MS or MS–MS alone. Additionally, at an early stage of cancer development, normal cells will greatly outnumber mutated cells. Thus any proposed technology must detect the mutant gene when it is present at low abundance relative to the corresponding normal gene. LC in these instances may be necessary prior to mass spectral analysis.

2. MS analysis

MS analysis of DNA may broadly be broken down into two types of techniques. Matrix assisted laser desorption ionization time-of-flight (MALDI-TOF) MS or electrospray ionization (ESI) in conjunction with mass analyzers of varying degrees of resolution. To discriminate SNPs or other changes in sequence by MS it is essential that double stranded (ds) PCR products be denatured and the mass of the single strands determined. Otherwise mass changes in one strand (e.g. G to C) may be neutralized by mass changes on the complementary strand (e.g. C to G) [9,10]. It has been observed with some ESI sources that external heating of the sample is required to denature the ds PCR products [10,11] whilst in other instances ds PCR products are inherently denatured into the two complementary strands [10]. In ESI-MS analysis, resolution is often adequate to allow detection of multiple strands as would be observed for heterozygotes (four strands, two from each allele). In MALDI-TOF MS analysis, ds PCR products are also often dissociated into single strands, however MS resolution is usually inadequate to discriminate these strands [12–15].

In MALDI-TOF MS, spectra are simple since a single mass peak is often generated with a charge of (± 1) for each molecular species present, which allows complex nucleic acid mixtures to be analyzed [16–19]. In contrast, ESI creates multiple charges for each oligomer or polymer. Each charge state is observed as a single peak in the mass spectrum. The charge state series created potentially makes mixture analysis difficult to interpret because of possible overlapping charge states [8,20]. For simple mixtures, this can be overcome by the use of software such as MaxEnt (Micromass, Danvers, MA). This algorithm produces a transformed spectrum from multiply charged ESI raw spectra. M_w of the components is plotted automatically without any prior knowledge (see Fig. 1).

In MS analysis of PCR products, the detection limit, mass accuracy and resolution are adversely affected with increasing M_w. In some cases, genetic variations lead to M_w changes that are substantially different for two PCR products selected for analysis (e.g. insertions/deletions of several nucleotides); in these instances, the demands for mass accuracy are lowered, and estimating the length is sufficient. Therefore, discrimination of higher M_w PCR products is simplified. In other cases variations are reflected by single nucleotide insertion/deletions, or

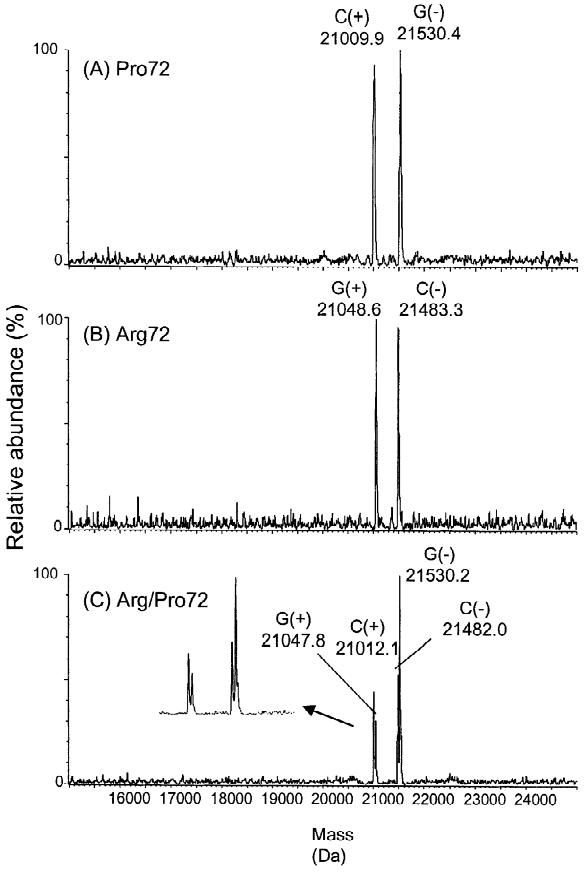

Fig. 1. Deconvoluted ESI quadrupole mass spectra of the three genotypes; 69 bp PCR products were amplified using human genomic DNA. (A) Pro72 and (B) Arg72 are homozygotes whilst (C) Pro/Arg72 is a heterozygote. Thus (A) and (B) show two peaks whilst (C) has four. The insert shows an expanded region of a mass spectrum. The sense and antisense strands are designated as (+) and (−), respectively.

substitutions. Insertion/deletions require discrimination of losses or gains of the mass of a nucleotide phosphate (289, 304, 313, or 329 Da, for C, T, A or G, respectively). However, single nucleotide substitutions result in a mass change between 9 and 40 Da. This limits the size of PCR products that can be selected for analysis since mass accuracy decreases with increasing M_w.

MALDI-TOF MS has demonstrated utility for the characterization of PCR products, which differ in nucleotide length. In one study, differentiation of PCR products generated from normal and mutated cystic fibrosis transmembrane conductance regulator genes (e.g. 60 and 57 mers) allowed determination of the 3 bp deletion in CF patients [14,21]. In a second study, the presence of normal polymorphic alleles of the carbonic anhydrase gene were differentiated. Homozygotes and heterozygotes were differentiated by the presence of either a 133, or 152 mer, or both [12]. In a third study, MALDI-TOF MS was used to identify *Legionella pneumophila*, the causative agent of Legionnaires disease. Identification was based on detection of two PCR products, a 108 mer from the 5S rRNA gene and a 168 mer from the Mip gene [22]. In all of these examples, high resolution was not required, due to study design (i.e. large mass differences between the products that need to be discriminated).

Resolution at the level of a base pair insertion is required for Sanger sequencing and the mode of detection must be able to distinguish mixtures. MALDI-TOF MS spectra of mixtures are readily interpreted since each oligonucleotide in a sequencing ladder primarily produces one mass peak. MALDI-TOF MS spectra can be directly compared to sequencing ladders obtained by gel electrophoresis. However, analysis takes seconds compared to hours needed for electrophoresis [17,23]. Again, as the sequencing products increase in length, the utility of MALDI-TOF MS for this type of analysis is diminished. Due to the complexity of spectra, sequencing ladders are rarely analyzed by ESI-MS. Multiple peaks are generated for each of the numerous DNA species, many of which differ by only a single nucleotide addition.

MALDI-TOF MS has made significant progress in the recent past with analysis of PCR products where base additions and deletions have been detected [24]. However, the limited resolution has precluded precise characterization of PCR products for base substitutions, and is generally restricted to small stretches of DNA. As noted above, the two strands of ds PCR products are dissociated but the two strands are not generally resolved on MALDI-TOF MS analysis [15]. This makes it difficult to resolve the four strands of DNA that would be present in a heterozygote and limits the ability of MALDI-TOF MS to genotype single nucleotide substitutions in increasingly large DNA.

It has been reported that MALDI-TOF MS has the resolution to discriminate a C or T substitution (15.0

Da) in an ss PCR product, 69 bases long [25]. However, the size of the PCR products have been minimized because resolution decreases rapidly as the size increases. Indeed, above 65 mers the resolution is poor [26]. Further, most recent analyses utilizing MALDI-TOF MS have successfully employed primer extension reactions, restriction digested DNA, labeled probes bound to captured ss PCR product, or other strategies to minimize the M_w of the DNA under analysis for the purpose of genotyping [27–35]. In these instances, the sequence must be known to allow design of the primers/probes, limiting the technique to detection of known variations (e.g. SNPs).

MALDI-TOF MS is likely to be important for diagnosis of known genetic alterations associated with genetic disease. However, it will be more difficult to apply to population studies designed to determine genetic diversity where all potential sequence differences may not be known prior to analysis.

The work of Potier et al. [36] was among the first analyses of larger DNA molecules. Using ESI-quadrupole MS analysis of 40-, 72-, and 132-mer synthetic oligonucleotides, resolution and mass accuracy was decreased upon increasing molecular mass. The measured and predicted M_w differed by <1, 3, and 14 mass units, respectively.

Unfortunately, PCR products require more careful sample handling than synthetic oligonucleotides prior to MS analysis because of the complexity of the sample matrix. For adequate mass accuracy and sensitivity, clean-up of the PCR reactions is essential prior to mass spectral analysis. PCR reaction mixtures are complex, containing amplification enzyme, deoxynucleotide triphosphates (dNTPs), oligonucleotide primers and buffers (using metal ions including Na^+), non-ionic detergents and albumin. Low M_w species (e.g. dNTPs and primers) preferentially ionize over the higher molecular PCR products, dramatically dampening the signal and increasing the complexity of spectra.

Purification of PCR products from oligonucleotide primers and dNTPs for ESI-MS analysis has focused on differential solubility during precipitation (usually with ethanol) [37] adsorption to ionic columns, or retention on molecular sieves [38]. Affinity chromatography has also been used to prepare ss PCR products. In this case pairs of primers are used one of which is biotin labeled. The resulting ds PCR product binds to streptavidin-labeled magnetic beads and after washing, ss or ds PCR products are released by denaturation [24,25].

Monovalent and divalent cations bound to DNA have been removed by washing with ammonium acetate [39] and/or chelating agents (e.g. trans-1,2-diaminocyclohexane-*N*,*N*,*N*^{7},*N*^{7}-tetraacetic acid, CDTA) [40]. Purification of higher M_w ds PCR products from oligonucleotide primers and dNTPs is most simply performed with a simple ethanol precipitation in the presence of $[NH_4]^+$ acetate [7–10,37,41]. The poly-anionic phosphate backbone of DNA forms strong complexes with metal cations adversely affecting mass accuracy upon MS analysis. Adduction is reduced significantly by replacing the mono-valent and di-valent cations complexed to PCR products with $[NH_4]^+$ by washing with ammonium acetate [39,41]. The adsorbed ammonium ions dissociate from the PCR product on MS analysis [39]. After precipitation, there are still sufficient metal ions present, which may require the sample be subject to counter current microdialysis [10,11,42] and/or suspended in a basic solution of imidazole/piperidine during analysis. This eliminates residual metal ion complexes [38,41,43].

In regards to instrument configuration, the highest resolution analysis of nucleic acids is achieved using magnetic sector and FTICR analyzers. These instruments allow differentiation of single base substitutions [9,37,38,44]. PCR products as large as 500 mers have been successfully analyzed, using FTICR MS [45]. However, at the current time, even using FTICR, the mass accuracy to analyze single nucleotide substitutions has been limited to a 114 bp product [38].

Despite their high resolution, analysis with magnetic sector instruments has involved low M_w restriction digests. For example, a 91 bp PCR product was subjected to restriction enzyme digestion generating a 49 bp final product for ESI MS analysis [37]. ESI-magnetic sector MS was also able to distinguish single base substitutions (G or C polymorphism) between two 44 bp PCR products from a heterozygous patient for the collagen alpha 2 type I gene [44].

As noted above, M_w changes may be substantially

different for two PCR products selected for analysis (e.g. several nucleotides or greater); when the demands for mass accuracy are lowered discrimination of higher M_w PCR products is simplified. A 114- and 89-bp PCR product is generated from the interspace region of *B. subtilis* and *B. cereus*, respectively. These products were readily discriminated using ESI-FTICR MS. More importantly, the PCR product from *B. subtilis* W23 differed from the predicted molecular mass (based upon published sequence from *B. subtilis* strain 168) by ~40.0 Da suggesting a G to C switch [38]. This was subsequently confirmed by conventional Sanger sequencing [46]. Accurate characterization of the number of tandem repeats in a human allele including a nucleotide substitution [11] has also been demonstrated. It is clear that ESI-FTICR MS has superior resolution and mass accuracy, and is the only configuration to distinguish single nucleotide substitutions in PCR products above 100 bp.

ESI-quadrupole MS has also demonstrated the mass accuracy to allow differentiation of bacterial interspace region PCR products in the size range of 89–130 bp that differ by one or more nucleotides [20]. A mixture of 111/114 PCR products was also detected by ESI MS for some strains of *B. subtilis*. Although, 114 bp products alone are generated from other strains. It was demonstrated that this reflected genomic diversity among strains of *Bacillus subtilis* [7].

As mentioned, instrumental mass resolution is more critical for genotyping small genetic changes in DNA. Additionally, human genomic DNA, as opposed to cloned genes, must be used as the template in genotyping. Human DNA with a bi-allelic SNP would contain three potential genotypes; one homozygote for each allele, and a heterozygote, which would contain one copy of both alleles. Discriminating homozygotes from one another is primarily an issue of mass accuracy. However, in a heterozygote, four strands (two from each allele) of similar mass must be discriminated within one acquired spectrum. This is an issue not only of mass accuracy, but also resolution.

The ''low resolution'' analyzers (i.e quadrupole and quadrupole ion-trap) generally coupled to ESI have demonstrated the resolution to distinguish a G or C polymorphism in a 69 bp PCR product [8,10]. However, this 40.0 Da shift is the largest existing mass difference with regards to SNPs. The most common SNP is a C or T transition (results in a predicted 15.0 Da difference) and accounts for about two-thirds of all SNPs. Homozygous C or T polymorphisms have also been distinguished in intact PCR products [8].

3. HPLC MS analysis

In MALDI-TOF MS, the sample is embedded in a solid matrix and ablated by a laser, and therefore is not readily amenable to on-line separation techniques such as HPLC and CE. Conversely, in the ESI source, sample enters as a liquid by a continuous flow. As discussed, the analysis of nucleic acid mixtures by direct ESI-MS is limited because complex spectra are produced making interpretation difficult. Coupling ESI-MS to LC could offer on-line separation of the mixtures as well as provide sample clean-up, automation, and the additional advantage of performing SIM or MRM. The use of SIM, or MRM, results in an increase in specificity by monitoring only the m/z values of interest.

On-line HPLC–MS of nucleic acids requires a solvent system which allows both optimal chromatographic separation and efficient ionization. As previously mentioned, ESI-MS of nucleic acids is complicated by the high affinity for metal cations for the phosphate backbone resulting in highly adducted spectra complicating interpretation and limiting sensitivity. The addition of organic bases [43] to the injection solvent has been most successful in achieving efficient ionization. Unfortunately, the most widely used organic bases (piperidine and imidazole) are relatively non-volatile, and may build up in the MS lenses causing loss of sensitivity, and therefore are not amenable to a practical HPLC–MS method. On-line cation-exchange chromatography has also been utilized in order to remove metal cations prior to ESI MS analysis. After desalting, mass spectra of synthetic oligonucleotides revealed no significant sodium adducted peaks [47].

Traditionally, HPLC of nucleic acids have focused on anionic exchange chromatography which utilizes a gradient of increasing salt concentration for elution. A rapid method (relative to gel electrophoresis)

for efficient separation, quantification and purification of both DNA restriction fragments and PCR products has been described [48]. This method is successful for separation of nucleic acids when used in conjunction with non-specific detectors such as variable wavelength detectors. However, this does not provide the structural information afforded by MS, which is required for determining genetic variation.

Ion-pair (IP) reverse phase (RP) HPLC has also been utilized for the efficient separation of nucleic acids. This consists of a hydrophobic stationary phase and water–methanol mobile phase containing ion-pairing agents which form non-covalent cross-links between the nucleic acid and the stationary phase [47]. The nucleic acid is eluted by increasing the organic concentration in the mobile phase.

Conventionally, triethylammonium (TEA) acetate has been used as an ion-pairing agent. However, it has been shown that as concentration of TEA acetate increases, chromatographic resolution improves, but ESI signal is dampened. This ion suppression is presumably caused by the presence of acetate [49,50]. An ion pair reverse phase (IP-RP) HPLC–MS solvent system for analysis of synthetic oligonucleotides replacing the acetate component with 1,1,1,3,3,3-hexafluoro-isopropanol (HFIP) was thus developed [51]. TEA (~10 m*M*) in MeOH was shown to be optimal for chromatographic separation and ESI MS analysis of oligonucleotides.

Others have utilized this system to genotype known SNPs [52,53]. Small restriction enzyme digested 15–20 mer DNA fragments were analyzed in SIM mode, which was required for optimal sensitivity/specificity. However mutations may be found in different locations within genes or large regions of genes (e.g. cancer). In this case, total ion monitoring would be necessary, since the mass (or m/z) would not be known prior to analysis.

Another group developed a capillary IP-RP HPLC–MS method, which allows separation and ionization of ss and ds DNA utilizing a custom designed polymer based column [54]. Using styrene and divinylbenzene as copolymers for packing a capillary column, chromatographic separation is efficient and single base resolution is achieved up to synthetic 18 mers. Restriction digested plasmids were also analyzed, successfully resolving fragments ranging in size from 51 to 587 bp. ESI of the ds DNA is observed, without strand dissociation. Thus, this would not discriminate single base substitutions in PCR products. Interestingly, 100 m*M* TEA acetate was utilized. In contrast to Apffel et al. who observed ion suppression at concentrations above 10 m*M*. The difference may reside in the flow-rates, low μl/min flow-rates in the latter study [54].

To screen for previously undetected polymorphisms or mutations, total ion monitoring mass spectrometry (MS) is required. As noted above, an IP-RP HPLC ESI-MS procedure for genotyping known polymorphisms has been described. However, SIM proved essential for optimal sensitivity. However, (in unpublished work) using this solvent system, intact PCR products generated highly adducted mass spectra with poor sensitivity. The solvent system was optimized to improve the sensitivity of MS analysis while retaining the chromatographic characteristics of the separation system. The primary change was increasing the TEA concentration to 368 m*M* and neutralizing the solution to a pH of ~10 with HFIP. Silica-based C_{18} columns traditionally used for IP-RP-LC analysis of nucleic acids are unstable at pH above 7.0. Therefore, XTerra-MS, a recently introduced C_{18} column with a usable pH range from 2 to 12, was utilized instead.

Two polymorphisms of the p53 gene were genotyped. (1) A cytosine (C) or guanine (G) transversion, designated C↔G (G↔C on the opposite strand), were each detected by a 40.0 Da change upon ESI quadrupole MS analysis and (2) a cytosine/thymine (T) transition, designated C↔T (G↔A on the opposite strand). This SNP is discriminated by a 15.0 Da change on one strand (C↔T) and a 16.0 Da change on the other (G↔A). Mass spectra, obtained by LC–MS were comparable to those observed with direct ESI-MS (compare Figs. 1 and 2).

4. MS–MS analysis

MALDI-TOF MS and ESI-MS analysis can detect known single nucleotide substitutions. ESI-MS might more readily detect previously unobserved point mutations in larger stretches of DNA [8,10, 24,32,55]. ESI-MS–MS has the potential to evaluate

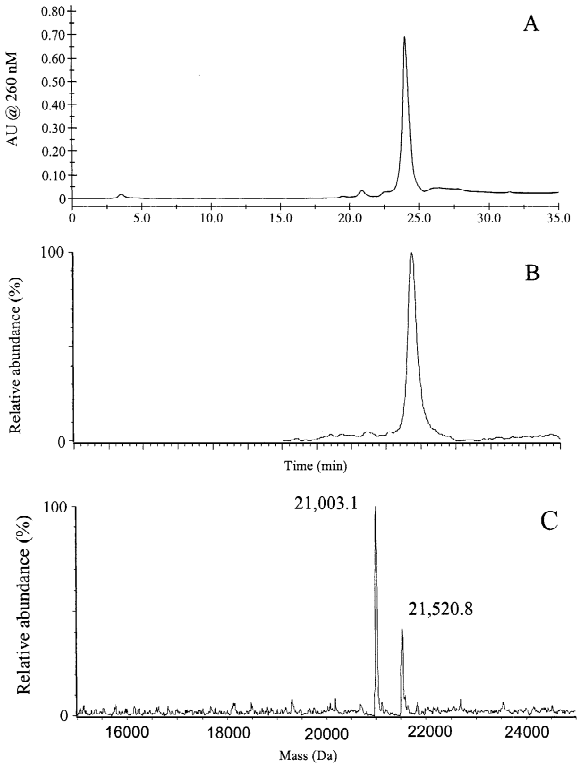

Fig. 2. IP-RP-LC analysis of a 69 bp PCR product amplified from the cloned human p53 proline variant of codon 72. (A) UV absorbance at 260 n*M*, (B) total ion chromatogram using ESI-MS and, (C) the deconvoluted mass spectrum.

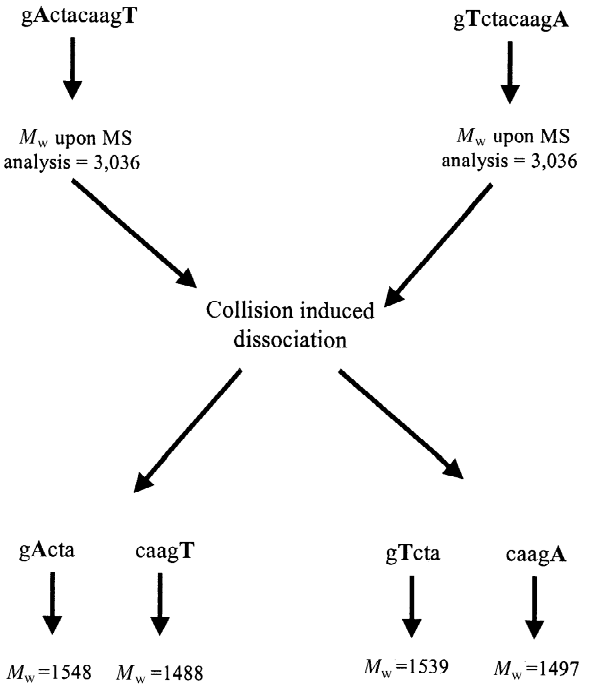

Fig. 3. Demonstration of MS–MS analysis of a 10 mer. M_w upon MS is identical making them indistinguishable. After CID, fragments are created which separate the sequence differences and are detected in MS–MS. The 9.0 Da mass unit shift corresponds to an A or T substitution.

more complex sequences (e.g. nucleotide substitutions at two or more sites). For example, two PCR products with the same base composition may still differ in sequence. These PCR products would have the same mass and could not be discriminated with MS alone, but would require MS–MS analysis. Fig. 3 illustrates how two pieces of DNA (10 mers) with the same nucleotide composition (and the same M_w) differing in sequence can produce fragments of differing M_w (after MS–MS).

In MS–MS, intact DNA (with a known charge) is fragmented by collision-induced dissociation (CID) by reaction with a neutral collision gas (e.g. Ar or He) creating a spectra of detectable product ions. In the early 1990s McLuckey and co-workers were able to obtain the sequence of small oligonucleotides using ESI-MS–MS [56,57]. This work defined two series of ions representing breakages in the phosphodiester bridges from the 3′ and 5′ ends, respectively. Ions a–d are numbered from the 5′ end while the corresponding ions w–z are numbered from the 3′ end. Fig. 4 illustrates possible fragmentation patterns for higher M_w DNA that can be observed on MS–MS analysis. Secondary ions can be generated by an additional break in the primary fragment, also at a phosphodiester bridge, creating an internal fragment. They also observed that the major ions in the spectra represent initial loss of a base (usually adenine) followed by breakage at the immediate 3′ C–O bond from which the base was lost.

Since this work, it has been shown that sequence can be inferred a priori from product ion spectra of short oligonucleotides (e.g. 15 mers) [58]. Using a triple quadrupole mass spectrometer with nozzle-skimmer fragmentation produced di-nucleotides. This was followed by CID and MS–MS analysis, which allowed determination of the two adjacent nucleotides. The position of modified nucleotides

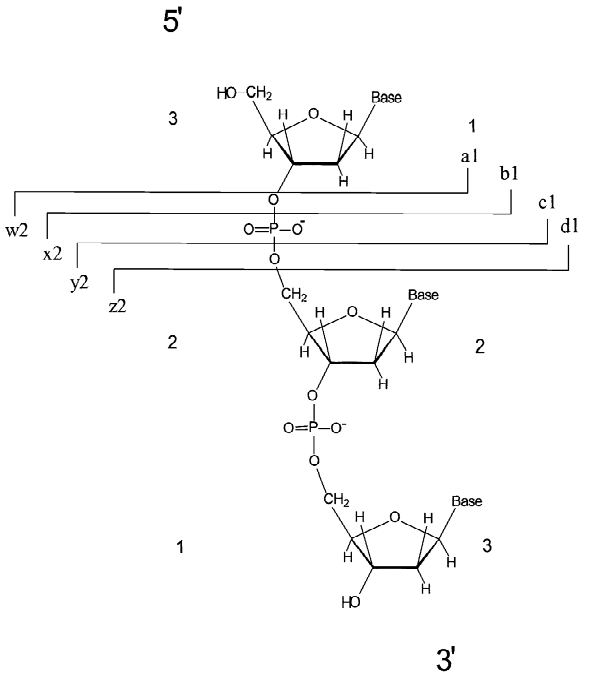

Fig. 4. Interpretation of fragmentation patterns of nucleic acids using MS–MS analysis.

present in tRNA (76 mer) have also been determined by this approach [59]. With the quadrupole ion trap, MS–MS fingerprints of restriction digested PCR products are also readily obtained [10,52]. Oligonucleotides up to 100 bases have been sequenced using MS–MS without resorting to sequencing ladders [60,61].

However, complete sequencing from fragmentation patterns of large pieces of DNA is not something readily achieved at present. This is primarily due to the numerous fragments created, each of which have multiple potential charge states. Only, high resolution MS–MS with isotopic resolution has the ability to assign charge states in product ion spectra [62]. Alternatively MS–MS can be used as a fingerprinting tool to recognize mass shifts in product ion spectra of different PCR products [10]. These mass shifts should correspond to identifiable sequence variations. In some instances, it may be adequate to use MS–MS to recognize a sequence change without defining the exact nature of the sequence.

It might be useful to consider possible respective uses of MS and MS–MS in genomics. In studying a group of unknowns it would be required that the sequence of the region of interest first be known for one ''standard human genomic DNA''. Primers would then be designed against this known sequence. The masses of each PCR product (which would not have been previously sequenced) would be determined using MS. If the masses fall into two clear groups a polymorphism would have been identified. The use of MS–MS would then be used as a complementary technique, to determine the presence of genetic differences other than an SNP. Some of these changes might result in regions with the same nucleotide composition and mass but different sequence. Such changes would be missed on MS analysis. Most product ion spectra would fall into two patterns representing the SNP. The purpose would be to identify aberrant PCR products that did not fall into these two fragmentation patterns. In most instances, these PCR products would represent only a small portion of samples. Only PCR products with aberrant MS–MS fingerprints would then be subject to conventional Sanger sequencing. Prior MS and MS–MS analysis, used as a screening technique, could drastically reduce the amount of sequencing in genomic population studies. Assignment of peaks in product ion spectra of PCR products demonstrate that fragmentation patterns are interpretable based on known sequence [10]. However, in a population study it might not be necessary to identify all, or indeed any peaks, in product ion spectra.

Ion traps and triple quadrupoles are both well suited for high-throughput analysis. These instruments work exquisitely in MS analysis for M_w determination of PCR products [41]. It is well-recognized due to the differences on configuration of the two instruments (''separation in time'' vs. ''separation in space'') that ion traps display much greater sensitivity for obtaining product ion spectra [63,64]. The difference in sensitivity for the two instruments becomes exaggerated as oligonucleotides or PCR products of increasing mass are analyzed. An additional advantage of the ion trap is the capability to perform MS^n. This allows the sequential fragmenting of ions of interest providing additional information on the sequence differences between PCR products.

ESI-MS and MS–MS have been used for the analysis of SNPs and more complex genetic varia-

tions. Double stranded (ds) PCR products were studied. A 69 bp arginine mutant PCR product [5′-x(C_{17})x(G_{38}) x-3′] with a negating switch has the same mass as the proline variant [5′-x(G_{17})x(C_{38})x-3′] but are readily distinguishable on ion trap MS–MS analysis. MS and MS–MS are powerful and complementary techniques for analysis of DNA [10]. In summary, MS can readily distinguish SNPs but MS–MS is required to identify nucleotide substitutions at two or more sites.

5. Concluding statements

ESI-MS and MS–MS are complementary techniques. MS analysis allows discrimination of single nucleotide polymorphisms. MS–MS allows differentiation of PCR products of different sequence, even if they have the same nucleotide composition. The combination of MS and MS–MS have great potential for use in the study of diversity of bacterial and human genomes. LC–MS (and LC–MS–MS) will certainly be necessary when complex mixtures of PCR products require analysis. MS, MS–MS, LC–MS and LC–MS–MS are all amenable to high-throughput and automated analysis.

At the current time quadrupole MS analysis has been shown to have adequate mass accuracy to detect a C to T switch in a 59 bp PCR product from homozygotes. The more difficult tasks of detecting a C to T switch in a heterozygote or an A to T switch in homozygote or heterozygote also remains to be studied. Analysis of higher M_w PCR products, while retaining the discrimination to detect single nucleotide substitutions, is also feasible. Two approaches can be employed to enhance the ability of MS to detect smaller differences in increasingly larger polymers. Firstly, higher resolution mass spectrometers can be used. Further research using FTICR instruments clearly needs to be performed. Quadrupole time-of-flight MS instruments (e.g. the Q-TOF) may also be of utility since they have the potential for higher resolution MS and MS–MS analysis with high throughput. Alternatively, samples can be manipulated to increase the mass differences of PCR products. For example, stable isotope (^{13}C, ^{15}N) labeled nucleotide tri-phosphates can be used as reagents in PCR [65]. Selectively labeling thymine triphosphates with $^{13}C/^{15}N$ isotopes in a PCR reaction would increase the mass difference between a C to T transition from 15.0 to 27.0 Da. Such labeling does not change the length or sequence or other properties of the DNA, and the M_w of PCR products only increases slightly.

The groundwork has been laid for increasingly cost effective, time efficient, accurate, sensitive DNA analysis using mass spectrometry. Modifications in sample preparation may eventually allow integration of PCR, clean-up and mass spectral analysis into a single user-friendly instrument.

Acknowledgements

This research was supported by a grant (to AF and KF) from NIH (Human Genome Project, R21 HG01810-01).

References

[1] D. Ravine, J. Inherit. Metab. Dis. 22 (1999) 503.
[2] P.Y. Kwok, Annu. Rev. Genomics Hum. Genet. 2 (2001) 235.
[3] E. Carrilho, Electrophoresis 21 (2000) 55.
[4] A.J. Nataraj, I. Olivos-Glander, N. Kusukawa, W.E. Highsmith Jr., Electrophoresis 20 (1999) 1177.
[5] J.M. Jaklevic, H.R. Garner, G.A. Miller, Annu. Rev. Biomed. Eng. 1 (1999) 649.
[6] V. Gurtler, V.A. Stanisich, Microbiology 142 (1996) 3.
[7] Y. Shaver, M. Nagpal, K.F. Fox, R. Rudner, A. Fox, Mol. Microbiol. 42 (2001) 101.
[8] J.J. Walters, W. Muhammad, K.F. Fox, A. Fox, D. Xie, K.E. Creek, L. Pirisi, Rapid Commun. Mass Spectrom. 15 (2001) 1752.
[9] D.S. Wunschel, D.C. Muddiman, K.F. Fox, A. Fox, R.D. Smith, Anal. Chem. 70 (1998) 1203.
[10] M.T. Krahmer, J.J. Walters, K.F. Fox, A. Fox, K. Creek, L. Pirisi-Creek, D.S. Wunschel, R.D. Smith, D. Tabb, J. Yates, Anal. Chem. 72 (2000) 4033.
[11] J.C. Hannis, D.C. Muddiman, Rapid Commun. Mass Spectrom. 13 (1999) 323.
[12] Y.H. Liu, J. Bai, Y. Zhu, X. Liang, D. Siemieniak, P.J. Venta, D.M. Lubman, Rapid Commun. Mass Spectrom. 9 (1995) 735.
[13] Y.H. Liu, J. Bai, X. Liang, D.M. Lubman, P.J. Venta, Anal. Chem. 67 (1995) 3482.
[14] L.Y. Ch'ang, K. Tang, M. Schell, C. Ringelberg, K.J. Matteson, S.L. Allman, C.H. Chen, Rapid Commun. Mass Spectrom. 9 (1995) 772.

[15] C. Jurinke, D. van den Boom, V. Collazo, A. Luchow, A. Jacob, H. Koster, Anal. Chem. 69 (1997) 904.
[16] U. Pieles, W. Zurcher, M. Schar, H.E. Moser, Nucl. Acids Res. 21 (1993) 3191.
[17] D.J. Fu, K. Tang, A. Braun, D. Reuter, B. Darnhofer-Demar, D.P. Little, M.J. O'Donnell, C.R. Cantor, H. Koster, Nat. Biotechnol. 16 (1998) 381.
[18] S. Mouradian, D.R. Rank, L.M. Smith, Rapid Commun. Mass Spectrom. 10 (1996) 1475.
[19] M.T. Roskey, P. Juhasz, I.P. Smirnov, E.T. Takash, S.A. Martin, L.A. Haff, Proc. Natl. Acad. Sci. 93 (1996) 4724.
[20] Y.A. Johnson, M. Nagpal, M.T. Krahmer, K.F. Fox, A. Fox, J. Microbiol. Methods 40 (2000) 241.
[21] M.J. Doktycz, G.B. Hurst, S. Habibi-Goudarzi, S.A. McLuckey, K. Tang, C.H. Chen, M. Uziel, K.B. Jacobson, R.P. Woychik, M.V. Buchanan, Anal. Biochem. 230 (1995) 205.
[22] G.B. Hurst, M.J. Doktycz, A.A. Vass, M.V. Buchanan, Rapid Commun. Mass Spectrom. 10 (1996) 377.
[23] D.J. Fu, N.E. Broude, H. Koster, C.L. Smith, C.R. Cantor, Genome Res. 12 (1996) 137.
[24] P.L. Ross, P. Belgrader, Anal. Chem. 69 (1997) 3966.
[25] P.L. Ross, P.A. Davis, P. Belgrader, Anal. Chem. 70 (1998) 2067.
[26] Y. Zhu, L. He, J.R. Srinivasan, D.M. Lubman, Rapid Commun. Mass Spectrom. 11 (1997) 987.
[27] L.A. Haff, I.P. Smirnov, Nucl. Acids Res. 25 (1997) 3749.
[28] A. Harksen, P.M. Ueland, H. Refsum, K. Meyer, Clin. Chem. 45 (1999) 1157.
[29] Y. Wada, K. Mitsumori, T. Terachi, O. Ogawa, J. Mass Spectrom. 34 (1999) 885.
[30] J. Li, J.M. Butler, Y. Tan, H. Lin, S. Royer, L. Ohler, T.A. Shaler, J.M. Hunter, D.J. Pollart, J.A. Monforte, C.H. Becker, Electrophoresis 20 (1999) 1258.
[31] T.J. Griffin, W. Tang, L.M. Smith, Nat. Biotechnol. 15 (1997) 1368.
[32] T.J. Griffin, L.M. Smith, Anal. Chem. 72 (2000) 3298.
[33] T.J. Griffin, L.M. Smith, Trends Biotechnol. 18 (2000) 77.
[34] Z. Fei, L.M. Smith, Rapid Commun. Mass Spectrom. 14 (2000) 950.
[35] T.A. Worrall, B.J. Schmeckpeper, J.S. Corvera, R.J. Cotter, Anal. Chem. 72 (2000) 5233.
[36] N. Potier, A. Van Dorsselaer, Y. Cordier, O. Roch, R. Bischoff, Nucl. Acids Res. 22 (1994) 3895.
[37] Y. Naito, K. Ishikawa, Y. Koga, T. Tsuneyoshi, H. Terunuma, R. Arakawa, Rapid Commun. Mass Spectrom. 9 (1995) 1484.
[38] D.C. Muddiman, D.S. Wunschel, C. Liu, L. Pasa-Tolic, K.F. Fox, A. Fox, G.A. Anderson, R.D. Smith, Anal. Chem. 68 (1996) 3705.
[39] J.T. Stults, J.C. Marsters, Rapid Commun. Mass Spectrom. 5 (1991) 359.
[40] P.A. Limbach, P.F. Crain, J.A. McCloskey, J. Am. Soc. Mass Spectrom. 6 (1995) 27.
[41] M.T. Krahmer, Y.A. Johnson, J.J. Walters, K.F. Fox, A. Fox, M. Nagpal, Anal. Chem. 71 (1999) 2893.
[42] C. Liu, Q. Wu, A.C. Harms, R.D. Smith, Anal. Chem. 68 (1996) 3295.
[43] M. Greig, R.H. Griffey, Rapid Commun. Mass Spectrom. 9 (1995) 97.
[44] T. Tsuneyoshi, K. Ishikawa, Y. Koga, Y. Naito, S. Baba, H. Terunuma, R. Arakawa, D.J. Prockop, Rapid Commun. Mass Spectrom. 11 (1997) 719.
[45] D.C. Muddiman, A.P. Null, J.C. Hannis, Rapid Commun. Mass Spectrom. 13 (1999) 1201.
[46] M. Nagpal, K.F. Fox, A. Fox, J. Microbiol. Methods 33 (1998) 211.
[47] C.G. Huber, J. Chromatogr. 806 (1998) 3.
[48] Y.F. Maa, S.C. Lin, C. Horvath, U.C. Yang, D.M. Crothers, J. Chromatogr. 508 (1990) 61.
[49] A. Apffel, J.A. Chakel, S. Fischer, K. Lichtenwalter, W.S. Hancock, Anal. Chem. 69 (1997) 1320.
[50] D. Wunschel, K.F. Fox, A. Fox, M.L. Nagpal, K. Kim, G.C. Stewart, M. Shahgholi, J. Chromatogr. 776 (1997) 205.
[51] A. Apffel, J.A. Chakel, S. Fischer, K. Lichtenwalter, W.S. Hancock, J. Chromatogr. 777 (1997) 3.
[52] S.J. Laken, P.E. Jackson, K.W. Kinzler, B. Vogelstein, P.T. Strickland, J.D. Groopman, M.D. Friesen, Nat. Biotechnol. 16 (1998) 1352.
[53] P.E. Jackson, G.S. Qian, M.D. Friesen, Y.R. Zhu, P. Lu, J.B. Wang, Y. Wu, T.W. Kensler, B. Vogelstein, J.D. Groopman, Cancer Res. 61 (2001) 33.
[54] A. Premstaller, H. Oberacher, C.G. Huber, Anal. Chem. 72 (2000) 4386.
[55] S. Zhang, C.K. Van Pelt, G.A. Schultz, Anal. Chem. 73 (2000) 2117.
[56] S.A. McLuckey, G.J. Van Berkel, G.L. Glish, J. Am. Soc. Mass Spectrom. 3 (1992) 60.
[57] S.A. McLuckey, S. Habibi-Goudarzi, J. Am. Chem. Soc. 115 (1993) 12085.
[58] J. Ni, C. Pomerantz, J. Rozenski, Y. Zhang, J.A. McCloskey, Anal. Chem. 68 (1996) 1989.
[59] J. Rozenski, J.A. McCloskey, Anal. Chem. 71 (1999) 1454.
[60] D.P. Little, F.W. McLafferty, J. Am. Chem. Soc. 117 (1995) 6783.
[61] D.P. Little, R.A. Chorush, J.P. Speir, M.W. Senko, N.L. Kelleher, F.W. McLafferty, J. Am. Chem. Soc. 116 (1994) 4893.
[62] R.D. Smith, J.A. Loo, C.G. Edmonds, C.J. Barinaga, H.R. Udseth, Anal. Chem. 62 (1990) 882.
[63] J.V. Johnson, R.A. Yost, P.E. Kelley, D.C. Bradford, Anal. Chem. 62 (1990) 2162.
[64] M.T. Krahmer, K.F. Fox, A. Fox, Int. J. Mass Spectrom. 191 (1999) 321.
[65] X. Chen, Z. Fei, L.M. Smith, E.M. Bradbury, V. Majidi, Anal. Chem. 71 (1999) 3118.