Optimized Suppression of Adducts in Polymerase Chain Reaction Products for Semi-Quantitative SNP Genotyping by Liquid Chromatography-Mass Spectrometry

Herbert Oberacher and Walther Parson
Institute of Legal Medicine, Innsbruck Medical University, Innsbruck, Austria

Georg Hölzl and Peter J. Oefner
Stanford Genome Technology Center, Palo Alto, California, USA

Christian G. Huber
Instrumental Analysis and Bioanalysis, Department of Chemistry, Saarland University, Saarbrücken, Germany

While electrospray ionization mass spectrometry has shown great potential for the identification of genotypes in DNA sequences amplified by polymerase chain reaction (PCR), the quantitative determination of allele frequencies remains challenging because of the presence of cationic adducts in the mass spectra which severely impairs accuracy of quantitation. We report here on the elaboration of an optimized desalting protocol for ion-pair reversed-phase high-performance liquid chromatography-electrospray ionization mass spectrometry (ICEMS) of PCR amplicons which facilitates the genotyping of single nucleotide polymorphisms (SNPs). Chromatographic purification at temperatures between 50 and 70 °C using monolithic reversed-phase columns and acetonitrile gradients in aqueous, 20–30 mmol/l butyldimethylammonium bicarbonate enabled the mass spectrometric analysis of nucleic acid solutions containing up to 1.7 mol/l sodium chloride. A further improvement in removal of metal cations was achieved upon the addition of 5–10 mmol/l ethylenediaminetetraacetic acid (EDTA) to the sample solution prior to liquid chromatography. ICEMS was used for the semi-quantitative genotyping of SNPs amplified from the tetraploid genome of potato cultivars. Using a quadrupole ion trap mass spectrometer, allele frequencies were determined with an accuracy of 2–9% by measuring the relative intensities of the signals corresponding to the molecular mass of each of the alleles in the deconvoluted mass spectra. ICEMS results correlated well with those obtained by pyrosequencing, single nucleotide primer extension, and conventional dideoxy sequencing. (J Am Soc Mass Spectrom 2004, 15, 1897–1906) © 2004 American Society for Mass Spectrometry

Single nucleotide polymorphisms (SNPs) are important genetic markers for identity testing, genotype-phenotype association studies, and reconstruction of evolution [1]. A large number of methods have been devised for determining the allelic state of SNPs and have been reviewed comprehensively [2, 3]. The exact determination of allele frequency in polyploid organisms and, eventually, in populations is essential in association studies and, in due consequence, there is growing interest in methodology suitable for quantitative determination of allele abundances [4–6]. Enzymatic amplification of DNA sequences by the polymerase chain reaction (PCR) is usually the first step in the genotyping of SNPs [7]. During the past decade, mass measurements by matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) and electrospray ionization mass spectrometry (ESI-MS) have become a valuable extension to the pool of methodologies suitable for allele determination after amplification by PCR [8–13]. While MALDI-MS is predominantly utilized for the high-throughput analysis of short products of primer extension minisequencing reactions [11], ESI-MS is applicable to the direct mass analysis of single- and double-stranded PCR amplicons ranging in size up to more than 500 base pairs [14, 15].

Sample purity represents one of the most critical factors for the success of mass spectrometric analyses of nucleic acids. PCR mixtures may contain varying amounts of amplification enzyme, deoxynucleotide

Published online November 11, 2004

Address reprint requests to Dr. C. G. Huber, Instrumental Analysis and Bioanalysis, Department of Chemistry, Saarland University, P.O. Box 151150, 66041 Saarbrücken, Germany. E-mail: christian.huber@mx.uni-saarland.de

1044-0305/04/$30.00
doi:10.1016/j.jasms.2004.09.004

Received June 23, 2004
Revised September 3, 2004
Accepted September 4, 2004

triphosphates, detergents, albumin, and other ionic buffer components including metal ions like sodium, potassium, and magnesium. Especially detergents, deoxynucleotide triphosphates, and primers deteriorate ionization efficiency due to competitive ionization [16, 17]. Metal ions, on the other hand, interact with the nucleic acids via substitution of variable numbers of protons in the phosphodiester groups, resulting in dispersion of pseudomolecular ions among several different species and decreased sensitivity [18]. Accurate mass measurements are severely hampered by such cation adduction, because the numerous signals for the higher charge states of adducted species merge into one unresolved, broadened peak shifted to higher masses relative to the fully protonated species.

Commonly applied off-line and on-line sample purification protocols for nucleic acids prior to mass spectrometric analysis include multiple ethanol precipitation [19], membrane filtration [20], solid-phase extraction [21], microdialysis [14], affinity purification [22, 23], cation-exchange [24], ligand-exchange [13], magnetic particles [25], size exclusion chromatography [26], and high-performance liquid chromatography (HPLC) [27, 28]. Ion-pair reversed-phase high-performance liquid chromatography hyphenated to electrospray ionization mass spectrometry (ICEMS) has been demonstrated to be a rapid and highly effective analytical tool for the characterization of PCR amplified sequences [15, 29–31]. ICEMS enabled the detection of SNPs both in homozygous and heterozygous individuals by intact molecular mass measurements using low femtomol amounts of DNA fragments up to lengths of more than 100 base pairs within a few min [32].

As outlined above, the efficient removal of adducted metal cations is essential both for correct allele identification by intact molecular mass determination and allele quantification based on mass spectrometric signal intensities. In this report, we therefore elaborated a protocol for improved purification of PCR amplicons by liquid chromatography. The applicability of the method is demonstrated by the semi-quantitative genotyping of SNPs in the polyploid genome of potato plants.

Experimental

Chemicals and Samples

Acetonitrile (HPLC gradient-grade) and water (HPLC grade) were obtained from Merck (Darmstadt, Germany). Ethylenediaminetetraacetic acid (EDTA, 99%), triethylamine (analytical reagent grade), and butyldimethylamine (analytical reagent grade) were purchased from Fluka (Buchs, Switzerland). Stock solutions (0.50 M) of triethylammonium bicarbonate or butyldimethylammonium bicarbonate were prepared by passing carbon dioxide gas (AGA, Vienna, Austria) through a 0.50 mol/l aqueous solution of the amines at 5 °C until pH 8.4–8.9 was reached. The synthetic oligodeoxynucleotide $(dT)_{16}$ (M_r 4805.2) was ordered from Microsynth (Balgach, Switzerland) and used without further purification.

PCR amplification of a 61-bp fragment from the Y-chromosomal locus M9 (allele G) was performed as described in reference [31]. The total reaction volume was 20 μl including 1× PCR buffer (Qiagen, Hilden Germany), 0.2 mmol/l each dNTP, 1 μmol/l forward primer (AACGGCCTAA GATGGTTGAA T), 1 μmol/l reverse primer (GAACGTTTGA ACATGTCTAA ATTAAAGA), 1× Q-Solution (Qiagen), 1 unit HotStar Taq Polymerase (Qiagen) and 10 ng DNA. PCR was performed on a Gene Amp PCR System 9700 (Applied Biosystems) comprising 45 cycles of 95 °C for 15 s, 54 °C for 30 s, and 72 °C for 1 min following initial denaturation at 95 °C for 15 min and final extension at 72 °C for 10 min.

PCR amplifications of a 68-bp segment containing the potato SNPs 44/114 + 118 (CTCCTGTGAC CTAAACCCCT GAAATTCT(T/C)C AA(A/C)TGGATTC CTTAAGCTAC CATTTTGATA TTTCCCTG) and of a 71-bp segment containing the SNP 44/174 (GCTACCATTT TGATATTTCC CTGAAAGAGC ACTTGGATGC AAC(A/G)CTCCCT TGGAAACAAG GTACTCAGCC G) were performed in 50 μl volumes containing 45 ng of genomic DNA extracted as described in [4], 25 μl of PCR Master Mix (Qiagen), and 0.4 μmol/l of each primer ($44/114-118_f$ CTCCTGTGAC CTAAACCCCT G $44/114\text{-}118_r$ CAGGGAAATA TCAAAATGGT AGC; $44/174_f$ GCTACCATTT TGATATTTCC CTG; $44/174_r$ CGGCTGAGTA CCTTGTTTCC AA; all primers were obtained from Microsynth; the indexes r and f are used to distinguish between the reverse and forward primers, respectively). Amplifications were carried out in a thermocycler (Mastercycler Personal, Eppendorf, Hamburg, Germany) comprising 35 cycles of 94 °C denaturation for 60 s, 54–56 °C annealing for 60 s, and 72 °C extension for 60 s. Prior to amplification the enzyme was activated by a 95 °C incubation for 15 min. Following a final extension step at 72 °C for 10 min, samples were chilled to 4 °C.

Capillary High-Performance Liquid Chromatography Coupled to Electrospray Ionization Mass Spectrometry

A low-pressure gradient micropump (Model Knauer 1100, Knauer, Berlin, Germany) was used to generate a primary flow of 100–150 μl which was then reduced to a constant secondary flow of 2.0 μl/min by means of a tee-piece (Valco Instruments, Houston, TX) and a 375 μm o.d. fused silica capillary with an internal diameter of 50 μm and a total length of 1 m. Chromatographic separation was performed in a monolithic, poly-(styrene/divinylbenzene) capillary column (60 × 0.2 mm i.d.) prepared according to a published protocol [33]. The column temperature was thermostated with a heat-jacket made from 3.3 mm o.d. copper tubing, which was heated by a circulating water bath (Model K20 KP, Lauda, Lauda-Königshofen, Germany). 500-nl volumes

of sample were injected onto the column through a nano-injection valve (Model C4-1004, Valco Instruments).

Nucleic acids were detected and mass analyzed by ESI-MS using a quadrupole ion trap mass spectrometer (LCQ, Thermo Electron Corporation, San Jose, CA). The capillary column was connected directly to the spray capillary (fused silica, 90 μm o.d., 20 μm i.d., Polymicro Technologies, Phoenix, AZ) by means of a microtight union (Upchurch Scientific, Oak Harbor, WA). In order to enhance detection sensitivity, acetonitrile was added post-column to the column effluent at a flow rate of 2.0 μl/min using a syringe pump and a tee-piece [34]. Electrospray voltage was set at 3.1 kV (Figures 1, 2, and 3) or 5.0 kV (Figures 4, 5, and 6). A nitrogen sheath gas flow of 40 (Figures 1, 2, and 3) or 100 arbitrary units (Figures 4, 5, and 6) was employed. The temperature of the heated capillary was set to 250 °C. Total ion chromatograms and mass spectra were recorded on a personal computer with the Xcalibur software version 1.2 (Thermo Electron Corporation). Deconvolution of raw mass spectra was performed with the BioWorks Software package from Thermo Electron. Mass calibration and tuning were performed according to a published optimization procedure [35].

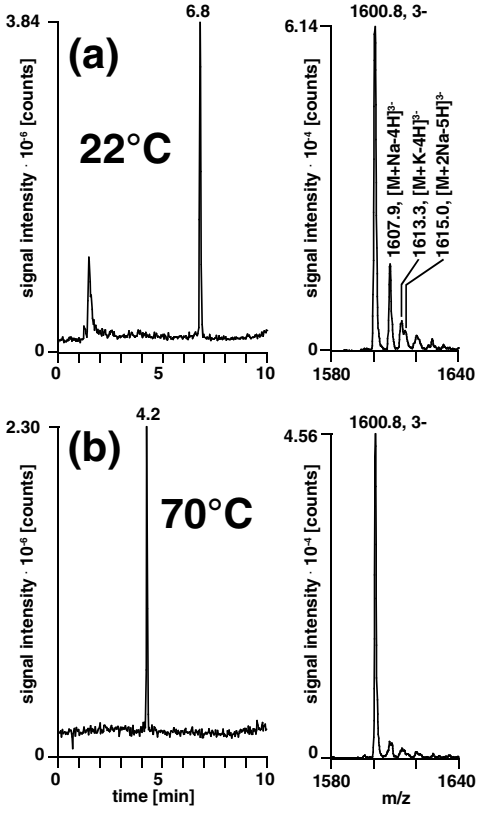

Figure 1. Desalting efficiency of ICEMS at different column temperatures. Column, PS-DVB monolith, 60 × 0.20 mm i.d.; mobile phase, (A) 25 mmol/l triethylammonium bicarbonate, pH 8.40, (B) 25 mmol/l triethylammonium bicarbonate, pH 8.40, 20% acetonitrile; linear gradient, 10–60% B in 10.0 min; flow-rate, 2.0 μl/min; temperature, (**a**) 22 °C, (**b**) 70 °C; sheath liquid, 2.0 μl/min acetonitrile; scan, 1000–3000; sample, 104 fmol of $(dT)_{16}$.

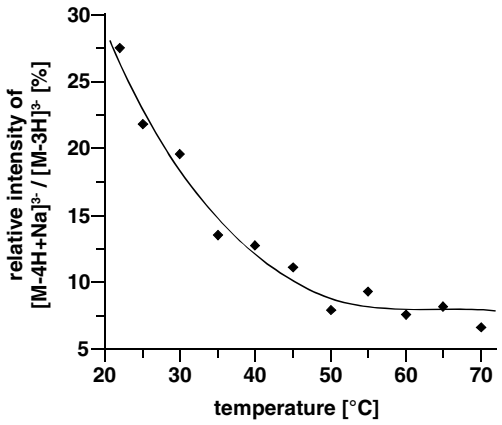

Figure 2. Effect of column temperature on the relative signal intensity of $[M - 4H + Na]^{3-}$ of $(dT)_{16}$. Temperature, 22–70 °C. Other experimental conditions as in Figure 1. Data points represent the averages of three measurements, the standard deviations varied between 0.3 and 4.7% (absolute values).

Results and Discussion

Effect of Column Temperature on Desalting Efficiency of ICEMS

The influence of column temperature on the removal of adducted metal cations was investigated through the triplicate analysis of a synthetic 16-mer oligonucleotide, dissolved in HPLC grade water, at column temperatures between 22 and 70 °C. Figure 1 illustrates the total ion current chromatograms and raw mass spectra obtained at temperatures of 22 and 70 °C. At least five mass spectra of the oligonucleotide were extracted from the reconstructed ion current chromatograms (RICC) and averaged. Apart from a decrease in chromato-

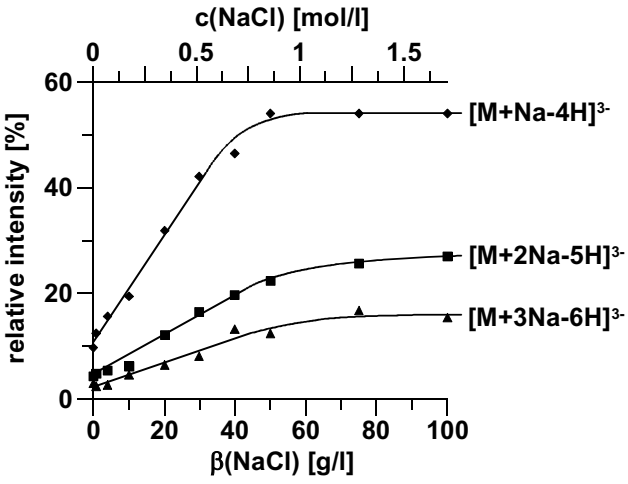

Figure 3. Effect of sodium chloride content on the relative signal intensities of $[M - 4H + Na]^{3-}$, $[M - 5H + 2Na]^{3-}$, and $[M - 6H + 3Na]^{3-}$ of $(dT)_{16}$. Temperature, 70 °C; samples, 104 fmol of $(dT)_{16}$ dissolved in 0 to 1.71 mol/l sodium chloride solutions. Other experimental conditions are as in Figure 1. Data points represent the averages of three measurements, the standard deviations varied between 0.0 and 6.3% (absolute values).

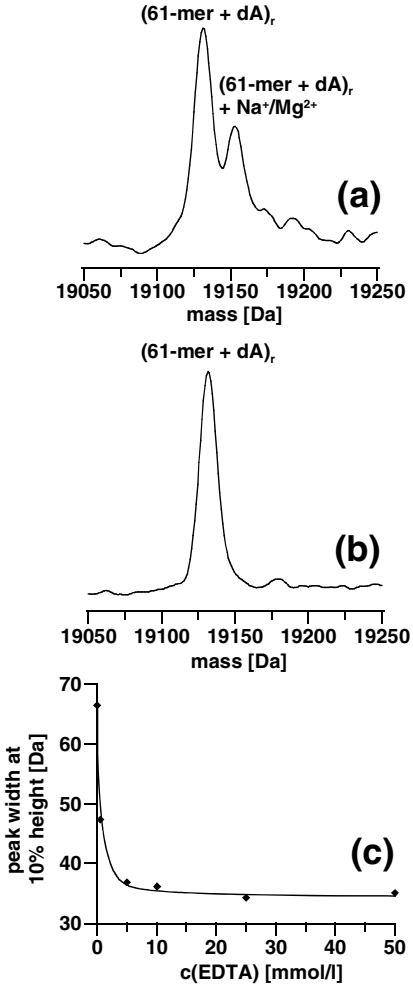

Figure 4. Influence of the EDTA concentration on the desalting efficiency of ICEMS. (**a**) Deconvoluted mass spectrum without EDTA, (**b**) deconvoluted mass spectrum after addition of EDTA (25 mmol/l) to the crude PCR mixture, (**c**) peak width at 10 % height of the reverse single strand of the M9 amplicon in the deconvoluted mass spectrum as a function of EDTA concentration; mmol/l mobile phase, (a) 25 mmol/l butyldimethylammonium bicarbonate, pH 8.40, (B) 25 mmol/l butyldimethylammonium bicarbonate, pH 8.40, 40% acetonitrile; linear gradient, 5–70% B in 10.0 min; flow-rate, 2.0 μl/min; temperature, 70 °C; scan, 500–2000; samples, 500 nl of crude PCR mixtures containing 100 fmol of amplicon and 0 to 25 μmol EDTA. Other experimental conditions are as in Figure 1. Data points in (**c**) represent the averages of three measurements, the standard deviations varied between 2.6 and 5.7 Da.

graphic retention at elevated temperature, there was no significant difference between the chromatograms. Both mass spectra showed predominant signals for the 3− charge state of the oligonucleotide. Although the oligonucleotide was dissolved in pure water, adduct signals were observed in all mass spectra because of the ubiquitous presence of small amounts of metal cations. The signal intensity of the $[(dT)_{16} - 4H + Na]^{3-}$ species relative to $[(dT)_{16} - 3H]^{3-}$ was used to characterize the desalting performance of ICEMS at the various column temperatures (Figure 2). The relative signal intensity of $[(dT)_{16} - 4H+Na]^{3-}$ decreased from 28% at 22 °C to 8% at 50 °C and remained practically constant at 6-10% above 50 °C. The quality of the deconvoluted mass spectra both with respect to mass accuracy and signal-to-noise ratio was clearly improved because of the reduction in cationic adducts.

We believe that the observed threefold reduction of adduct formation at column temperatures higher than 50 °C has to be attributed to a more efficient exchange of adducted Na^+, K^+, and Mg^{2+} ions with triethylammonium from the eluent during the chromatographic process. Upon transfer to the gas phase, dissociation of the adducted amine from the sugar-phosphate backbone of the nucleic acids occurs, leaving a proton with the oligonucleotide, which results in signal enhancement, charge state reduction, and cation adduct suppression [36]. Moreover, improved desalting can also be a consequence of higher temperature of the electrosprayed column effluent. This hypothesis is supported by experiments of Liu et al. [37] and Walters et al. [38], who reported an increased desalting performance by heating nucleic acids solutions prior to ESI-MS.

Elevated column temperature carries the additional advantage of facilitating detection of sequence variation. Base substitutions are difficult to identify in double-stranded PCR products by mass measurements, because A-T and G-C base pairs have almost identical masses. Nevertheless, they can be readily distinguished by measuring the molecular masses of the corresponding single strands, in which base substitutions lead to mass differences between 9 and 40 Da. In order to achieve maximum desalting efficiency and complete denaturation of the DNA double strands, we thus recommend the chromatographic purification of PCR amplicons at column temperatures of 60–70 °C.

Characterization of Salt Tolerance of ICEMS

Although a vast number of protocols were introduced in the last years for the purification of nucleic acids prior to ESI-MS, only few reports focused on the efficiency of these methods. In a report by Liu et al. [37], the salt tolerances of multiple buffer exchange, organic base addition, and microdialysis were studied by adding high concentrations of sodium chloride to the analyzed nucleic acid solutions. This group reported that microdialysis offered the highest desalting performance among the three techniques, which enabled the mass spectrometric detection of nucleic acids from solutions containing 0.5 mol/l sodium chloride.

To investigate the desalting performance of ICEMS, $(dT)_{16}$ solutions containing 0–1.7 mol/l sodium chloride were analyzed in triplicate at a column temperature of 70 °C. As above, the signal intensities of $[(dT)_{16}$-$4H + Na]^{3-}$, $[(dT)_{16} - 5H + 2Na]^{3-}$, and $[(dT)_{16} - 6H + 3Na]^{3-}$ relative to the abundance of $[(dT)_{16} - 3H]^{3-}$ served as indicators for the desalting capability of

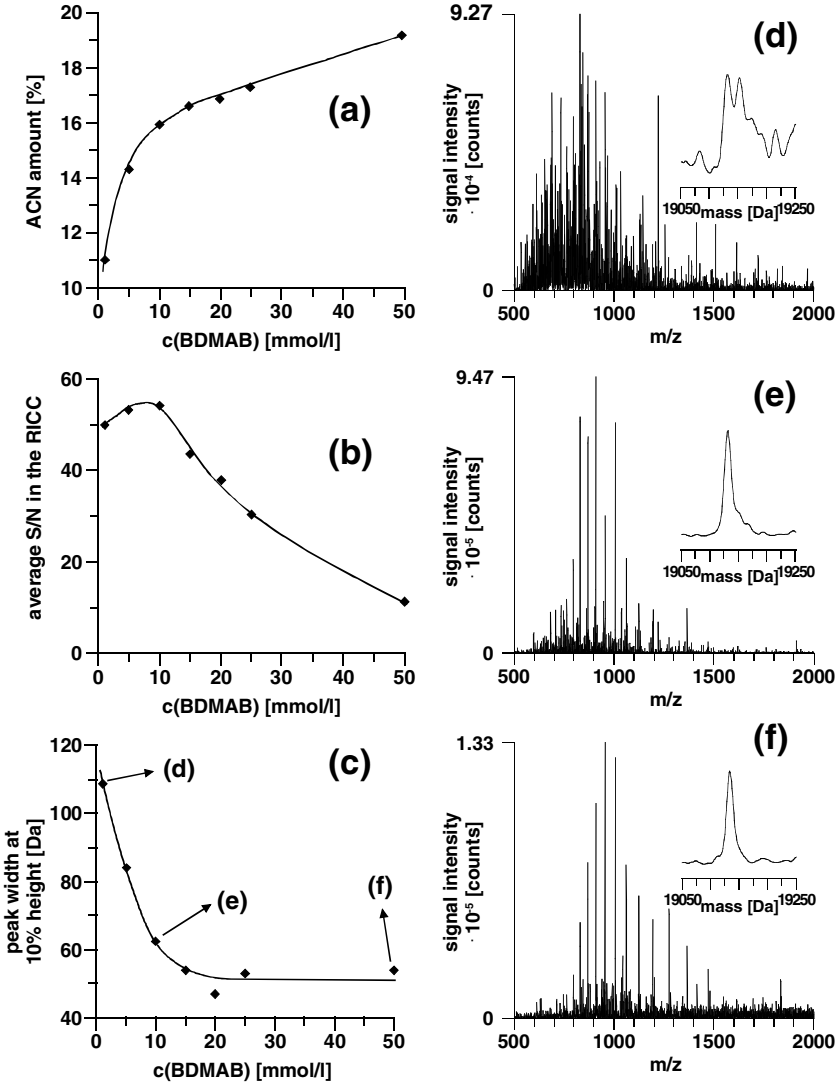

Figure 5. Influence of the BDMAB concentration (mol/l) in the mobile phase on (**a**) the acetonitrile concentration necessary for elution (corrected for gradient delay), (**b**) the signal-to-noise ratio in the RICC, and (**c**) the peak width at 10% height in the deconvoluted mass spectra of the reverse strand of the 61-bp M9 amplicon from the BDMAB concentration. (**d**), (**e**), (**f**) Illustrate raw and deconvoluted mass spectra of the M9 amplicon. Mobile phase (A), at different BDMAB concentrations. Mobile phase, (A) mmol/l mobile phase, (A) 1–50 mmol/l butyldimethylammonium bicarbonate, pH 8.40, (B) 1–50 mmol/l butyldimethylammonium bicarbonate, pH 8.40, 40% acetonitrile; samples, 500 nl of crude PCR mixtures containing 100 fmol of PCR amplicon and 10 μmol EDTA. Other experimental conditions are as in Figure 1. Data points in (**a**), (**b**), **c**) represent the averages of three measurements; the standard deviations varied between 0.0 and 0.3% (absolute values) in (**a**), 1–5 in (**b**), and 0.04–1.7 Da in (**c**).

ICEMS. Figure 3 shows that the relative signal intensities of mono-, di-, and tri-sodium adducts increased linearly between 0 and 0.7 mol/l sodium chloride and reached a plateau at higher concentrations with saturation values of 54, 26, and 15%, respectively. It is important here to note that the sodium-free species remained the most abundant at all sodium chloride concentrations, which enabled the correct determination of the molecular mass of the oligonucleotide. Moreover, the results indicate that even higher concentrations of sodium can be removed by ICEMS. No other sample purification method has shown such a high salt tolerance.

Improvement of Desalting Efficiency via Addition of Chelating Compounds

The experiments described above have shown that ICEMS is very efficient in the removal of monovalent metal cations. Nevertheless, relatively high concentrations of magnesium ions are present in PCR mixtures as

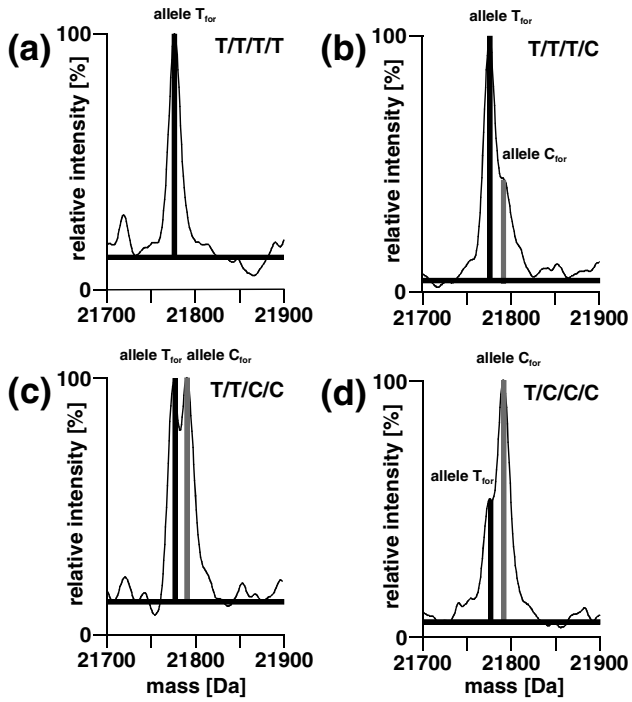

Figure 6. Comparison of the genotyping results obtained for four different allelic combinations of the potato SNP 44-174 by measuring the relative peak intensities of the forward strands of the two possible alleles in the deconvoluted mass spectra. Samples, 500 nl of crude PCR mixtures containing 100 fmol of amplicon and 10 μmol EDTA. Other experimental conditions are as in Figure 1.

a cofactor for DNA polymerase and other polyvalent ions may be introduced through contamination of reagents or genomic DNA. Polyvalent ions are quite difficult to remove from the DNA backbone because of their strong electrostatic interaction and a chelating effect in the binding to adjacent phosphodiester groups. As a consequence, magnesium adducts are observed especially in the mass spectra of products of PCR. The suppression of such adducts is of utmost importance for correct allele identification and determination of allelic ratios, because signals of adducted species may be misinterpreted as being related to alternative alleles.

We utilized a 61-bp amplicon of the human Y-chromosomal locus M9 to characterize the influence of bivalent ions on mass measurements by ICEMS. Figure 4a depicts the deconvoluted mass spectrum obtained from the analysis of 500 nl of a crude PCR mixture containing approximately 100 fmol amplicon. Upon comparison of the measured molecular mass with the theoretical mass, the most abundant signal was identified as the monoadenylated reverse strand of the amplified DNA sequence (M_r 19133). A signal for a second species was observed in the deconvoluted mass spectrum with a mass difference of 23 relative to the signal of highest intensity, which would be compatible with a sodium adduct (+22 mass units), a magnesium adduct (+22 mass units), the exchange of a cytosine by an adenine (+24 mass units), or, eventually, the exchange of a thymine by a guanine (+25 mass units). Since the PCR product was amplified from the hemizygous Y-chromosome, the latter two options can be excluded here. However, for autosomal loci or in the case of polyploid genomes, the mass spectrum could be readily misinterpreted as revealing alternative alleles. The presence of a sodium adduct is unlikely based on the experiments described above, and hence, adduction with magnesium ions is a reasonable interpretation of the second signal.

Adduction with magnesium makes accurate quantitative genotyping impossible and necessitates an improvement in the sample clean-up strategy. Based on previous work performed by Limbach et al. [39], who reported on the use of chelating agents in the spray solvent for the abstraction of bivalent cations from ribonucleic acids, we considered ethylenediaminetetraacetic acid (EDTA) as an additive to enhance the suppression of magnesium adducts. Since the presence of EDTA in the electrosprayed solvent deteriorated signal quality through competitive ionization, EDTA was added to the sample solution prior to analysis by ICEMS and not to the chromatographic eluent. EDTA complexed magnesium ions and other polyvalent ions. The complexes eluted in the void volume and hence, did not disturb the mass spectrometric analysis of PCR products. Figure 4b illustrates the deconvoluted mass spectrum of the reverse strand obtained upon analysis of 500 nl of a PCR mixture after addition of 12.5 mmol/l EDTA. A clean mass signal was observed and adduct signals were completely suppressed, which not only yielded a measured molecular mass very close to the expected value but also significantly increased the signal-to-noise ratio in the mass spectrum.

The optimal EDTA concentration necessary for the complete desalting of the PCR product was evaluated by means of addition of different amounts of the chelators to the PCR mixture. The peak widths at 10% height in the deconvoluted mass spectrum were taken as an indicator for the desalting efficiency. As shown in Figure 4c, a concentration of at least 5 mmol/l EDTA, which was 2 times higher than the employed $MgCl_2$ concentration, was necessary to ensure optimal desalting performance and to eliminate magnesium adduction. As far as we know, the combination of EDTA addition and ICEMS represents the only cleaning procedure that can remove all kinds of metal ions that bond to PCR amplified nucleic acids with such high efficiency. Even the combined use of ethanol precipitation and microdialysis could not achieve the same level of desalting efficiency (compare Figure 1 in reference [17]).

Optimization of Ion-Pair Reagent Concentration

The proper choice of a chromatographic eluent suitable for ICEMS is very important in order to maintain chromatographic separation efficiency while still guaranteeing low limits of mass spectrometric detection. Eluents of low electric conductivity and high content of volatile organic solvent are most suitable for nucleic

acid analysis by ESI-MS [27, 28, 34, 40, 41]. In ion-pair reversed-phase HPLC of nucleic acids, an ion-pair reagent such as triethylammonium acetate or butyldimethylammonium acetate is added to the mobile phase to ensure chromatographic retention of nucleic acids on a nonpolar stationary phase [42]. Since amines have also been shown to be quite effective in the suppression of cationic adducts [43], it was evident to investigate the effect of ion-pair reagent concentration with respect to the efficiency of adduct elimination, chromatographic performance, and mass spectrometric detectability.

PCR mixtures containing 100 fmol of the 61 bp PCR amplicon in 500 nl were analyzed using a gradient of 2–28% acetonitrile in 1 to 50 mmol/l butyldimethylammonium bicarbonate. To ensure highest desalting efficiency, 10 μmol EDTA were added per 20 μl sample solution and runs were performed at 70 °C. The influence of the concentration of ion-pair reagent on chromatographic retention of the reverse single strand on the monolithic column can be deduced from Figure 5a. The portion of acetonitrile in the eluent necessary to elute the PCR product increased steeply form 11.0 to 17.3% in the concentration range of 1–25 mmol/l butyldimethylammonium bicarbonate, whereas higher concentrations of ion-pair reagent concentration ensued only a slight increase in retention.

Figure 5b summarizes the effect of ion-pair reagent concentration on mass spectrometric detectability. For this analysis, the signal-to-noise ratio was measured in the RICC of the 61-bp PCR amplicon. It becomes obvious that the maximum signal-to-noise ratio was reached at 10 mmol/l butyldimethylammonium bicarbonate. At 25 mmol/l butyldimethylammonium bicarbonate, the concentration commonly used for ICMS analyses, the signal-to-noise ratio drops to about 50% of the maximum. Interestingly, the signal-to-noise ratio also deteriorates at concentrations of butyldimethylammonium bicarbonate lower than 10 mmol/l, presumably due to a lower content of acetonitrile during elution of the PCR product.

Finally, the desalting performance of ICEMS was examined as a function of the concentration of ion-pair reagent. Metal cations were efficiently removed in 15–50 mmol/l butyldimethylammonium bicarbonate (Figure 5c). However, as soon as the concentration of ion-pair reagent dropped below 15 mmol/l, metal cations were only inefficiently displaced by butyldimethylammonium ions and considerable amounts of mono-, di-, and tri- sodium and/or magnesium adducts were present in the deconvoluted mass spectrum (see insets in Figure 5c).

Semi-Quantitative Genotyping of SNPs

The optimization experiments outlined above have clearly shown that chromatographic separation at temperatures above 60 °C, concentrations of 20–30 mmol/l butyldimethylammonium bicarbonate in the eluent, and the addition of 10 mmol/l EDTA to the sample strongly support the suppression of cation adduction in ESI-MS of nucleic acids, which is essential for correct genotyping both in diploid and polyploid genomes by ICEMS. The strategy applied here for the semi-quantitative determination of allele frequencies is based on the chromatographic purification of PCR fragments containing the polymorphic sites under denaturing conditions and subsequent mass spectrometric detection of the single-stranded species that are characteristic for the different alleles. Since the allelic ratio in the genome may be considered to be conserved during PCR, liquid chromatography, and mass spectrometry, the relative signal intensities in the deconvoluted mass spectrum were utilized to deduce the relative allelic ratios.

While diploid organisms show only three possible allelic ratios for a biallelic SNP (1:0, 1:1, 0:1), five ratios are distinguishable in tetraploids (4:0, 3:1, 2:2, 1:3, and 0:4). For the determination of allelic ratios of SNPs in the tetraploid potato genome, PCR products of SNP 44-174 containing a T > C polymorphism were amplified and analyzed by ICEMS. The series of multiply charged ions were deconvoluted to yield the molecular masses of the single strands for the identification of the different alleles as well as their signal intensities for their relative quantitation (Figure 6 and Table 1). The molecular masses of the single strands corresponding to the two alleles differed by 16 mass units, which facilitated their differentiation in homo- and heterozygous samples using an ion trap mass spectrometer. Each sample was analyzed in triplicate except the sample showing a 1:1 ratio of the two alleles, which was run ten times. The standard deviations of the allelic ratios were between 1.6 and 4.7%.

In samples with 1:1 allelic ratio, the absolute differences between the measured and the theoretical allele

Table 1. Semi-quantitative genotyping of four different allelic combinations of the potato SNP 44–174 by measuring the relative peak intensities of the forward and reverse single strands of the two possible alleles T and C in the deconvoluted mass spectrum

Genotype	T/T/T/T		T/T/T/C		T/T/C/C		C/C/C/T	
Allele	T	C	T	C	T	C	T	C
Theoretical frequency [%]	100	0	75	25	50	50	25	75
Average measured allele frequency [%]	100	0	72	28	52	48	34	66
Standard deviation [%]	0.0		4.7		3.0		1.6	
Number of measurements	3		3		10		3	
Absolute deviation from expected	—		±3		±2		±9	

Table 2. Measured and theoretical molecular masses M_r used for the determination of the haplotypes of SNPs 44 − 114 + 118 by ICEMS

Measured M_r [Da]	Theoretical M_r [Da]	Assignment	Haplotypes
20652	20653	$(C,C)_{for}$	
20692	20692	$(T,A)_{for}$	
20966	20966	$(C,C)_{for}$ + dA	
21004	21005	$(T,A)_{for}$ + dA	C,C/T,A
21196	21195	$(T,A)_{rev}$	
21236	21236	$(C,C)_{rev}$	
21506	21508	$(T,A)_{rev}$ + dA	
21549	21549	$(C,C)_{rev}$ + dA	

frequencies were less than 2%. However, for samples with an unbalanced allele distribution, deviations of 3 and 9% from the theoretical allele frequencies were measured. The major reason for the observed deviations was limited mass resolution of the quadrupole ion trap mass analyzer, which could not separate the two oligonucleotide species to baseline. Due to this overlap of signals, the intensity of the lower abundant species is usually overestimated. This problem could be readily alleviated by using mass analyzers that provide higher resolution, such as linear ion trap-, time-of-flight-, or ion cyclotron resonance mass spectrometers.

In a second example, we compared the results of ICEMS analysis for the quantitative haplotyping of the potato SNP 44-114 and SNP 44-118 with data obtained by three other techniques, namely pyrosequencing, single nucleotide primer extension, and conventional dideoxy-sequencing. For the haplotyping by ICEMS, SNP 44-114 and SNP 44-118 were coamplified in a single PCR product. A comparison of the measured molecular masses with the theoretical masses calculated from the sequences of the possible haplotypes (T,A; T,C; A,C; C,C) clearly revealed the haplotypes T,A and C,C as the only combinations present (Table 2). Semi-quantitative genotyping was performed by measuring the relative signal intensities of the forward strand of each haplotype in the deconvoluted mass spectra. The peaks were resolved nearly to baseline due to a molecular mass difference of 40 Da, which enabled a more accurate determination of allele frequencies as compared to the previous example (Table 3). The deviation of the measured T,A frequency from the theoretical value was in all cases lower than 6%. Moreover, the ICEMS results correlated well with the results obtained by pyrosequencing, single nucleotide primer extension, and conventional dideoxy-sequencing (Table 3).

Conclusions

Careful purification and control of experimental conditions enables the exhaustive removal of adducted metal cations from nucleic acid molecules. In order to enhance desalting efficiency during chromatographic purification of oligonucleotides, EDTA should be added to the sample mixture to capture bivalent cations and separation should be performed at elevated temperature with concentrations of ion pair reagent between 20 and 30 mmol/l. The efficient desalting of nucleic acids not only effectuates better signal-to-noise ratios in the raw and deconvoluted mass spectra, but also results in sharper mass spectrometric signals, which enables a better resolution of different alleles of PCR-amplified DNA sequences. Extracting the signal intensities from deconvoluted mass spectra of PCR amplicons facilitates the rapid, semi-quantitative determination of relative allele ratios by ICEMS with an accuracy of better than 10%. Nevertheless, there is a large potential for improving quantitative accuracy further by using high-resolution mass spectrometers. In contrast to most other established methods for allele quantitation, PCR amplicons can be rapidly analyzed within less than 15 min by ICEMS without any other post-PCR sample treatment. Moreover, rapid column regeneration and equilibration protocols compatible with the monolithic separation media hold the potential for a further increase in sample throughput [31]. A major advantage of mass spectrometric investigation of polymorphisms rests within the possibility to investigate, simultaneously, both individual single strands of the PCR products, which enhances confidence of identification and quantitative accuracy.

Table 3. Genotyping of the three heterozygous combinations of the potato SNPs 44 − 114 + 118 by measuring the relative peak intensities of the forward strands of the two possible alleles T,A and C,C in the deconvoluted mass spectra

Genotype	T,A/T,A/T,A/C,C		T,A/T,A/C,C/C,C		T,A/C,C/C,C/C,C	
Allele	T,A	C,C	T,A	C,C	T,A	C,C
Frequency [%]	73	27	52	48	31	69
Theoretical frequency [%]	75		50		25	
Deviation [%]	2		−2		−6	
Pyrosequencing [%]	71.1		48.4		26.8	
Single nucleotide primer extension [%]	70.3		49.8		32.3	
Sequencing [%]	75.5		50.1		26.3	

The allele frequencies are compared to the values obtained by pyrosequencing, single nucleotide primer extension, and conventional dideoxy sequencing. (Data were adapted from reference [4]).

Acknowledgments

The authors thank Andreas M. Rickert for providing the potato DNA samples. Part of the experiments described in this work has been performed at the Institute of Analytical Chemistry and Radiochemistry, Leopold-Franzens-University, Innsbruck. This work has been supported by the National Institutes of Health (grant no. HG00205).

References

1. Wang, D. G.; Fan, J. B.; Siao, C. J.; Berno, A.; Young, P.; Sapolsky, R.; Ghandour, G.; Perkins, N.; Winchester, E.; Spencer, J.; Kruglyak, L.; Stein, L.; Hsie, L.; Topaloglu, T.; Hubbel, E.; Robinson, E.; Mittmann, M.; Morris, M. S.; Shen, N.; Kilburn, D.; Rioux, J.; Nusbaum, C.; Rozen, S.; Hudson, T. S.; Lander, E. S. Large-Scale Identification, Mapping, and Genotyping of Single-Nucleotide Polymorphisms in the Human Genome. *Science* **1998**, *280*, 1077–1082.
2. Gut, I. G. Automation in Genotyping of Single Nucleotide Polymorphisms. *Hum. Mutat.* **2001**, *17*, 475–492.
3. Kristensen, V. N.; Kelefiotis, D.; Kristensen, T.; Borresen-Dale, A. L. High-Throughput Methods for Detection of Genetic Variation. *Biotechniques* **2001**, *30*, 318–322.
4. Rickert, A. M.; Premstaller, A.; Gebhardt, C.; Oefner, P. J. Genotyping of SNPs in a Polyploid Genome by Pyrosequencing. *Biotechniques* **2002**, *32*, 592–603.
5. Sham, P.; Bader, J. S.; Craig, I.; O'Donovan, M.; Owen, M. DNA Pooling: A Toll for Large-Scale Association Studies. *Nature Rev. Gent.* **2002**, *3*, 862–871.
6. Null, A. P.; Nepomuceno, A. I.; Muddiman, D. C. Implications of Hydrophobicity and Free Energy of Solvation for Characterization of Nucleic Acids by Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2003**, *75*, 1331–1339.
7. Reynolds, R.; Sensabaugh, G. Analysis of Genetic Markers in Forensic DNA Samples Using the Polymerase Chain Reaction. *Anal. Chem.* **1991**, *63*, 2–15.
8. Haff, L. A.; Smirnov, I. P. Single-Nucleotide Polymorphism Identification Assays Using a Thermostable DNA Polymerase and Delayed Extraction MALDI-TOF Mass Spectrometry. *Genome Res.* **1997**, *7*, 378–388.
9. Li, J.; Butler, J. M.; Tan, Y.; Lin, H.; Royer, S.; Ohler, L.; Shaler, T. A.; Hunter, J. M.; Pollart, D. J.; Monforte, J. A.; Becker, C. H. Single Nucleotide Polymorphism Determination Using Primer Extension and Time-of-Flight Mass Spectrometry. *Electrophoresis* **1999**, *20*, 1258–1265.
10. Griffin, T. J.; Hall, J. G.; Prudent, J. R.; Smith, L. M. Direct Genetic Analysis by Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6301–6306.
11. Sauer, S.; Lechner, D.; Berlin, K.; Lehrach, H.; Escary, J.-L.; Fox, N.; Gut, I. G. A Novel Procedure for Efficient Genotyping of Single Nucleotide Polymorphisms. *Nucleic Acids Res.* **2000,** *28*, e13.
12. Huber, C. G.; Oberacher, H. Analysis of Nucleic Acids by On-Line Liquid Chromatography-Mass Spectrometry. *Mass Spectrom. Rev.* **2001**, *20*, 310–343.
13. Zhang, S.; Van Pelt, C. K.; Huang, X.; Schultz, G. A. Detection of Single Nucleotide Polymorphisms Using Electrospray Ionization Mass Spectrometry: Validation of a One-Well Assay and Quantitative Pooling Studies. *J. Mass Spectrom.* **2002**, *>37*, 1039–1050.
14. Muddiman, D. C.; Null, A. P.; Hannis, J. C. Precise Mass Measurment of a Double-Stranded 500 Base-Pair (309 kDa) Polymerase Chain Reaction Product by Negative Ion Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 1201–1204.
15. Oberacher, H.; Oefner, P. J.; Parson, W.; Huber, C. G. On-Line Liquid Chromatography-Mass Spectrometry: A Useful Tool for the Detection of DNA Sequence Variation. *Angew. Chem. Int. Ed.* **2001,** *40*, 3828–3830.
16. Doktycz, M. J.; Hurst, G. B.; Habibi-Goudarzi, S.; McLuckey, S. A.; Tang, K.; Chen, C. H.; Uziel, M.; Jacobson, K. B.; Woychik, R. P.; Buchanan, M. V. Analysis of Polymerase Chain Reaction-Amplified DNAP by Mass Spectrometry Using Matrix-Assisted Laser Desorption and Electrospray: Current Status. *Anal. Biochem.* **1995**, *230*, 205–214.
17. Null, A. P.; George, L. T.; Muddiman, D. C. Evaluation of Sample Preparation Techniques for Mass Measurements of PCR Products using ESI-FT-ICR Mass Spectrometry. *J. Am. Mass Spectrom.* **2002,** *13*, 338–344.
18. Smith, R. D.; Loo, J. A.; Edmonds, C. G.; Barinaga, C. J.; Udseth, H. R. New Developments in Biochemical Mass Spectrometry: Electrospray Ionization. *Anal. Chem.* **1990**, *62*, 882–889.
19. Stults, J. T.; Marsters, J. C. Improved Electrospray Ionization of Synthetic Oligodeoxynucleotides. *Rapid Commun. Mass Spectrom.* **1991**, *5*, 359–363.
20. Bayer, E.; Bauer, T.; Schmeer, K.; Bleicher, K.; Maler, M.; Gaus, H. J. Analysis of Double-Stranded Oligonucleotides by Electrospray. *Anal. Chem.* **1994**, *66*, 3858–3863.
21. Deroussent, A.; Le Caer, J.-P.; Gouyette, A. Electrospray Mass Spectrometry for the Purity of Natural and Modified Oligodeoxynucleotides. *Rapid Commun. Mass Spectrom.* **1995**, *9*, 1–4.
22. Ross, P. L.; Davis, P. A.; Belgrader, P. Analysis of DNA Fragments from Conventional and Microfabricated PCR Devices Using Delayed Extraction MALDI-TOF Mass Spectrometry. *Anal. Chem.* **1998**, *70*, 2067–2073.
23. Krahmer, M. T.; Johnson, Y. A.; Walters, J. J.; Fox, K. F.; Fox, A.; Nagpal, M. Electrospray Quadrupole Mass Spectrometry Analysis of Model Oligonucleotides and Polymerase Chain Reaction Products: Determination of Base Substitutions, Nucleotide Additions/Deletions, and Chemical Modifications. *Anal. Chem.* **1999**, *71*, 2893–2900.
24. Huber, C. G.; Buchmeiser, M. R. On-Line Cation-Exchange for Suppression of Adduct Formation in Negative-Ion Electrospray Mass Spectrometry of Nucleic Acids. *Anal. Chem.* **1998**, *70*, 5288–5295.
25. Hahner, S.; Schneider, A.; Ingendoh, A.; Mosner, J. Analysis of Short Tandem Repeat Polymorphisms by Electrospray Ion Trap Mass Spectrometry. *Nucleic Acids Res.* **2000**, *28*, e82.
26. De Bellis, G.; Salani, G.; Battaglia, C.; Pietta, P.; Rosti, E.; Mauri, P. Electrospray Ionization Mass Spectrometry of Synthetic Oligonucleotides Using 2-Propanol and Spermidine. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 243–249.
27. Apffel, A.; Chakel, J. A.; Fischer, S.; Lichtenwalter, K.; Hancock, W. S. New Procedure for the Use of High-Performance Liquid Chromatography-Electrospray Ionization Mass Spectrometry for the Analysis of Nucleotides and Oligonucleotides. *J. Chromatogr. A* **1997**, *777*, 3–21.
28. Huber, C. G.; Krajete, A. Analysis of Nucleic Acids by Capillary Ion-Pair Reversed-Phase HPLC Coupled to Negative Ion-Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **1999**, *71*, 3730–3739.
29. Oberacher, H.; Parson, W.; Mühlmann, R.; Huber, C. G. Analysis of Polymerase Chain Reaction Products by On-Line Liquid Chromatography-Mass Spectrometry for Genotyping of Polymorphic Short Tandem Repeat Loci. *Anal. Chem.* **2001**, *73*, 5109–5115.
30. Oberacher, H.; Oefner, P. J.; Hölzl, G.; Premstaller, A.; Huber, C. G. Resequencing of Multiple Single Nucleotide Polymor-

phisms by Liquid Chromatography-Electrospray Ionization Mass Spectrometry. *Nucl. Acids Res.* **2002**, *30*, e67.
31. Berger, B.; Hölzl, G.; Oberacher, H.; Niederstätter, H.; Huber, C. G.; Parson, W. Single Nucleotide Polymorphism Genotyping by On-Line Liquid Chromatography-Mass Spectrometry in Forensic Science with the Y-Chromosomal Locus M9. *J. Chromatogr. B* **2002**, *782*, 89–97.
32. Oberacher, H.; Huber, C. G.; Oefner, P. J. Mutation Screening by Ion-Pair Reversed-Phase High-Performance Liquid Chromatography-Electrospray Ionization Mass Spectrometry (ICEMS). *Hum. Mutat.* **2003**, *21*, 86–95.
33. Premstaller, A.; Oberacher, H.; Huber, C. G. High-Performance Liquid Chromatography-Electrospray Ionization Mass Spectrometry of Single- and Double-Stranded Nucleic Acids Using Monolithic Capillary Columns. *Anal. Chem.* **2000**, *72*, 4386–4393.
34. Huber, C. G.; Krajete, A. Sheath Liquid Effects in Capillary High-Performance Liquid Chromatography-Electrospray Mass Spectrometry of Oligonucleotides. *J. Chromatogr. A* **2000**, *870*, 413–424.
35. Oberacher, H.; Walcher, W.; Huber, C. G. Effect of Instrument Tuning on Detectability of Biopolymers in Electrospray Ionization Mass Spectrometry. *J. Mass Spectrom.* **2003**, *38*, 108–116.
36. Muddiman, D. C.; Cheng, X.; Udseth, H. R.; Smith, R. D. Charge-State Reduction with Improved Signal Intensity of Oligonucleotides in Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 706–797.
37. Liu, C.; Muddiman, D. C.; Tang, K.; Smith, R. D. Improving the Microdialysis Procedure for Electrospray Ionization Mass Spectrometry of Biological Samples. *J. Mass Spectrom.* **1997**, *32*, 425–431.
38. Walters, J. J.; Muhammad, W.; Fox, K. F.; Fox, A.; Xie, D.; Creek, K. E.; Pirisi, L. Genotyping Single Nucleotide Polymorphisms Using Intact Polymerase Chain Reaction Products by Electrospray Quadrupole Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2001**, *15*, 1752– 1759.
39. Limbach, P. A.; Crain, P. F.; McCloskey, J. A. Molecular Mass Measurement of Intact Ribonucleic Acids via Electrospray Ionization Quadrupole Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 27–39.
40. Bleicher, K.; Bayer, E. Various Factors Influencing the Signal Intensity of Oligonucleotides in Electrospray Mass Spectrometry. *Biol. Mass Spectrom.* **1994**, *23*, 320–322.
41. Griffey, R. H.; Sasmor, H.; Greig, M. J. Oligonucleotides Charge States in Negative Ionization Electrospray-Mass Spectrometry are a Function of Solution Ammonium Ion Concentration. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 155–60.
42. Huber, C. G. Micropellicular Stationary Phases for High-Performance Liquid Chromatography of dsDNA. *J. Chromatogr. A* **1998**, *806*, 3–30.
43. Greig, M. J.; Griffey, R. H. Utility of Organic Bases for Improved Electrospray Mass Spectrometry of Oligonucleotides. *Rapid Commun. Mass Spectrom.* **1995**, *9*, 97–102.