

Matrix-induced fragmentation of P3'-N5' phosphoramidate-containing DNA: high-throughput MALDI-TOF analysis of genomic sequence polymorphisms

Mikhail S. Shchepinov*, Mikhail F. Denissenko, Kevin J. Smylie, Ralf J. Wörl, A. Lorieta Leppin, Charles R. Cantor and Charles P. Rodi

Sequenom Inc., 3595 John Hopkins Court, San Diego, CA 92121, USA

Received June 4, 2001; Revised and Accepted July 20, 2001

ABSTRACT

Chemical and enzymatic approaches were used to produce polynucleotide fragments containing acid-labile internucleotide P3'-N5' phosphoramidate bonds, either in a surface-bound form or in solution. The primer extension reaction utilizing 5'-amino-5'-deoxynucleoside 5'-triphosphates generates polynucleotides that can be fragmented into short, easy-to-analyze pieces simply by being premixed with the acidic matrices typically used for MALDI-TOF mass spectrometry of nucleic acids. This leads to detection procedures that are simple, robust and easy to automate. Utilizing this approach, a polymorphic site in the human ADRB3 gene was interrogated. Primer extensions with phosphoramidate analogs of dNTPs allowed for unambiguous discrimination of all possible genotypes.

INTRODUCTION

Upon the completion of the first draft of the human genome (1,2) rapid, automated methods of single nucleotide polymorphism (SNP) analysis are now in even greater demand for applications in pharmacogenomics, target validation and identification of drug leads, as well as in agriculture, forensics and environmental work. Validating diagnostic candidate SNPs can be a challenging task because the frequency of the contributing polymorphism may be only slightly elevated in a disease group as compared to a control. Furthermore, the genetic basis of many diseases is often composed of not one, but multiple gene variants.

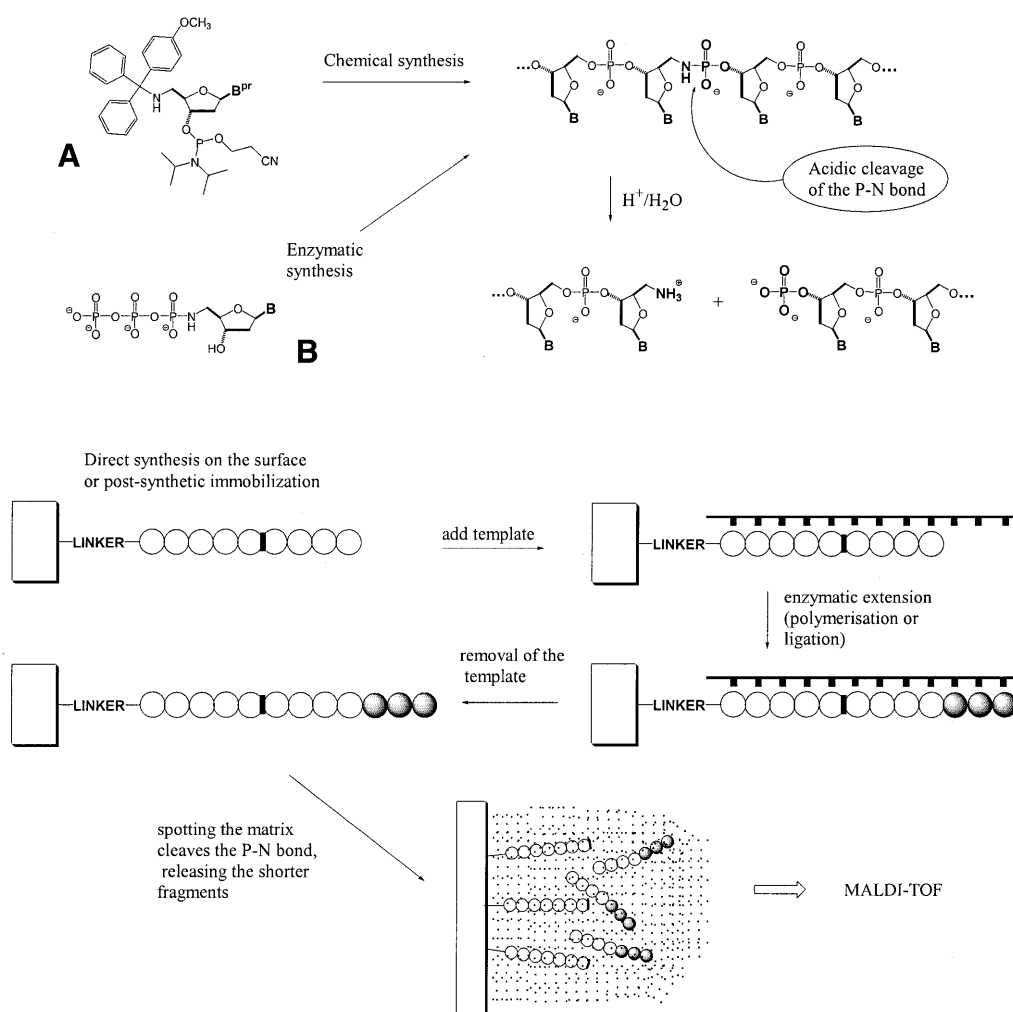
The MassARRAY system has recently been developed for testing large numbers of samples with unprecedented speed, accuracy and levels of automation by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) (3). First, a genotyping primer is annealed to the amplified target in a region immediately adjacent to the polymorphic site. The primer is then enzymatically extended by one to a few bases through the polymorphic site and terminated. This approach is known as Primer Oligo Base

Extension (PROBE) and is now referred to as the MassEXTEND method (4,5). The extended primer is robotically dispensed onto a silicon microchip at locations preloaded with matrix and the genotype is automatically determined using the mass values obtained with MALDI-TOF MS. This method of analysis generally does not cause any substantial fragmentation of DNA. Current resolution allows for unambiguous and precise detection of oligonucleotides of up to 50 bases in length (6), but shorter fragments are preferable because longer fragments produce lower intensity signals as compared to shorter fragments and this signal intensity further decreases when complex mixtures are analyzed. In some applications it is therefore desirable to cut DNA fragments, generated by polymerases or by other methods, into shorter pieces, which are easier to measure.

Sequence-specific shortening or fragmentation of a polynucleotide strand can be achieved by either enzymatic or chemical means. Enzymatic cleavage of DNA may include site-specific digestion using various ribo- and deoxyribonucleases or glycosylases. The obvious advantage of this method is generation of molecules with well-defined and mostly uniform 3'- and 5'-ends. Chemical cleavage is generally more cost-effective and robust, but less site-specific. Therefore, chemically modified internucleoside bonds positioned at defined locations would be ideal for specific DNA fragmentation and MALDI-TOF MS detection.

Of many chemically cleavable nucleotide modifications described, we were particularly interested in the so-called 'achiral' 5'-phosphoramidate analog introduced more than 30 years ago (7). Attractive features of this compound include its ability to exist in a triphosphate form, which can be utilized by some polymerases (8), and the fact that cleavage of the P-N bond occurs under acidic conditions (9). Importantly, the modification can also be easily introduced into synthetic oligodeoxyribonucleotides using standard automated oligonucleotide synthesis (10). Acidic deprotection used in oligonucleotide synthesis to remove 5'-protecting groups does not cleave the P-N bond of the nascent oligonucleotide, because prior to alkali deprotection of the oligos at the end of the synthesis, all the phosphate and phosphoramidate groups are still protected, most likely by methyl or cyanoethyl protecting groups. The

*To whom correspondence should be addressed. Tel: +1 858 202 9073; Fax: +1 858 259 5735; Email: misha@bioch.ox.ac.uk



Scheme 1. Chemical structure of (A) the 5'-deoxy-5'-amino-3'-deoxynucleoside phosphoramidite and (B) the 5'-P-N-3'-deoxynucleoside triphosphate. Schematic principle of the enzymatic extension of immobilized P-N primer on a template added in solution and acidic cleavage of the P-N bond.

phosphoramidates are thus present in the form of diesters, which are stable to acids.

The matrix routinely used in MALDI-TOF MS to analyze DNA (3-hydroxypicolinic acid, 3-HPA) is acidic. In practice, all enzymatic steps are usually carried out in solution and DNA samples are mixed with matrix prior to MS analysis (11). In those cases where the enzymatic steps preceding the MS, such as primer extension, are conducted on DNA immobilized to a solid support (DNA chips), the matrix is then dispensed atop these chips.

Here we describe a method whereby the DNA strands are fragmented at predetermined locations in the process of mixing the DNA samples with acidic matrix, thus reducing the number of steps and simplifying the process of analysis (Scheme 1).

MATERIALS AND METHODS

Oligodeoxyribonucleotides including 5'-biotinylated ones were obtained from IDT (Iowa City, IA). 7-Deaza-dGTP and 7-deaza-dATP were purchased from TriLink (San Diego, CA). 5'-Deoxycytidine-5'-aminonucleoside was custom synthesized

by Fidelity Systems (Gaithersburg, MD). The following DNA polymerases were used: HotStar (Qiagen, Valencia, CA), Sequenase v.2.0 (US Biochemical, Cleveland, OH), Vent, Vent(exo⁻) and Deep Vent(exo⁻) (New England Biolabs, Beverly, MA), Thermo Sequenase (Amersham Pharmacia Biotech, Piscataway, NJ), *Tth* (Roche, Mannheim, Germany), TaKaRa *Taq* (TaKaRa Biomedicals, Japan) and DyNAzyme I and II (Finnzymes, Espoo, Finland). 5'-MMTr-5'-deoxy-5'-aminothymidine 3'-phosphoramidite and other phosphoramidites were from Glen Research (Sterling, VA). dNTPs and ddNTPs were from Amersham Pharmacia Biotech. 3-HPA and all other chemicals and solvents were from Sigma/Aldrich (St Louis, MO).

Chemical synthesis and cleavage of a P-N bond in immobilized oligonucleotides

Silicon (100) chips (2 × 3 cm) were silanized as described (12) and oligonucleotides were directly synthesized on this surface using the physical masking method developed by Southern *et al.* (13). 5'-Monomethoxytrityl-5'-deoxy-5'-aminothymidine 3'-phosphoramidite (Scheme 1A) was incorporated at the

desired position in the nascent oligonucleotide (solid support)-(linker)-3'-T₄(p)T-NH-(p)T-T₈-5' using automated oligonucleotide synthesis (ABI 394 DNA Synthesizer; Applied Biosystems, Foster City, CA). Subsequent coupling of the next phosphoramidite to a deprotected 5'-amino group and standard oxidation with iodine generated the desired P-N bond with yields comparable to those for non-modified synthons. The chip was then immersed in a saturated ammonia solution (~30% in water) for 30 min at room temperature to deprotect the oligonucleotide. The 3-HPA matrix solution [200 nl of 350 mM 3-HPA in acetonitrile, 1.5% trifluoroacetic acid (TFA)] was placed atop the surface covered with this oligonucleotide and the chip analyzed by MALDI-TOF MS as described below.

Synthesis and cleavage of a P-N bond in oligonucleotides in solution

Oligonucleotides **1**, d(AGAAGGTGTCp-nTGC^UGGGAG), and **2**, d(AAAGCTGCGTGATGAp-nTGAAA), where 5'-P-N modified nucleosides are underlined, were synthesized using standard phosphoramidite chemistry. The 5'-deoxy-5'-aminonucleoside was incorporated as described above and both oligos were purified according to standard protocols. Oligonucleotides were purified by PAGE and dissolved in TE buffer, pH 8.5. For acidic hydrolysis of a P-N bond, oligomers were spotted onto a stainless steel MS target support and incubated with a MALDI matrix solution adjusted to 0.33% TFA for ~1 min. The cleavage products were analyzed using a linear time-of-flight Voyager DE mass spectrometer (PE BioSystems, Framingham, MA) with a nitrogen laser firing at 3 Hz.

PCR using P-N bond-containing oligonucleotide primers

Genomic DNA was extracted from the blood of anonymous donors using a Puregene DNA isolation kit (Gentra, Minneapolis, MN) and specific SNPs determined in separate experiments using a standard MassEXTEND (PROBE) assay (3). A short (42 bp) fragment of the MHTFR gene exon 4 (GenBank accession no. AF105980) was amplified using P-N bond-containing primers **1** and **2** and HotStar heat-activated DNA polymerase as described (5). The thermal cycling profile was as follows: 15 min at 95°C, two cycles of 10 s at 94°C, 30 s at 58°C and 40 s at 72°C, 45 cycles of 10 s at 85°C, 20 s at 58°C and 30 s at 72°C, followed by 2 min at 72°C. PCR product was analyzed on an agarose gel.

Synthesis and enzymatic incorporation into DNA of P-N bond-containing dTTP and dCTP

5'-Triphosphates of deoxythymidine and deoxycytidine (Scheme 1B) were synthesized from the corresponding 5'-deoxy-5'-aminonucleosides as described (8,9). The reaction was monitored by HPLC (data not shown). Triphosphates were methanol precipitated, redissolved in TE buffer, pH 8.5, and used in enzymatic reactions without further purification. Pre-annealed model extension template was prepared by annealing oligonucleotides d(CTGAAATCTCGGCTCGTGTCTCGCGTCAACCTTGGA) and d(GACGCGAGACACGAGCC). A genomic fragment of the lipoprotein lipase (LPL) gene (GenBank accession no. AF050163) was amplified in a three primer PCR reaction with a universal biotinylated sequence using primers d(biotin-AGCGGATAACAATTTTCACACAGG), d(ACCGCTGCAA-

CAATCTGG) and d(AGCGGATAACAATTTTCACACAGGA-GAACGAGTCTTCAGGTAC). Primers d(biotin-AGCGGATAACAATTTTCACACAGG), d(AATACCGCCAACACCAGTGG) and d(AGCGGATAACAATTTTCACACAGGTGGTCA-TGGTCTGGAGTCTC) were used for genomic amplification of the β 3-adrenergic receptor (ADRB3) fragment (GenBank accession no. X72861). Single-stranded DNA templates for extension reactions were generated using a standard MassEXTEND (PROBE) procedure (5). Extension primers d(TCTGGGCTA-TGAGATCA) and d(TGGTCATCGTGGCCATCGCC) were used on single-stranded templates LPL and ADRB3, respectively. A typical Sequenase 2.0 extension reaction (10 μ l) contained 25 mM Tris-HCl, pH 9.5, 20 mM MgCl₂, 50 mM NaCl, 5 mM DTT, 200 μ M dNTP or ddNTP, 2–10 mM 5'-deoxy-5'-amino-CTP (P-N CTP) or P-N dT, 10 pmol annealed extension template and 1–2 U Sequenase 2.0. The Vent(exo⁻) reaction (10 μ l) contained 25 mM Tris-HCl, pH 9.5, 20 mM MgCl₂, 5 mM DTT, 200 μ M dNTPs, 2 mM 5'-deoxy-5'-amino-TTP (P-N dTTP), 10 pmol annealed extension template, 2 U Vent(exo⁻). Sequenase and Vent(exo⁻) reactions were incubated at 37 and 72°C, respectively, then terminated by placing on dry ice. After a C18 ZipTip (Millipore, Bedford, MA) purification, extension products were analyzed by MS as described above. For acidic cleavage, a 3-HPA matrix was adjusted to 0.33–0.5% TFA or 0.1% HCl.

MALDI-TOF MS of DNA fragments

If not indicated otherwise, purified DNA samples were robotically dispensed onto a silicon chip (SpectroCHIP; Sequenom, San Diego, CA) and analyzed using a MassARRAY system (Sequenom) as described (5).

RESULTS AND DISCUSSION

Chemical incorporation of a P-N bond-containing thymidine nucleotide into DNA

A P-N modification of a phosphodiester bond can be introduced into synthetic oligonucleotides using standard automated oligonucleotide synthesis (10). The acidic deprotection used in this synthesis for removal of 5'-protecting groups does not cleave the P-N bond of the nascent oligomer because, prior to alkali deprotection of the oligonucleotides at the end of the synthesis, all phosphate and phosphoramidate groups are still alkyl-protected. The phosphoramidates are thus present in the form of diesters, which are stable to acids present during the synthesis. Initially, we have chosen the chemical synthesis strategy to introduce the P-N bond into DNA. First, we incorporated a single P-N bond into an oligonucleotide covalently attached to a solid support. The local areas of an on-chip synthesized P-N bond-containing thymidine oligonucleotide were exposed to acidified MALDI matrix with the expectation that cleavage of the fragment would occur while the matrix was crystallizing (usually ~2 min). The chip was then analyzed by MALDI-TOF MS. The resulting mass spectrum is shown in Figure 1. As expected, a peak at 2754 *m/z* corresponding to a shorter fragment (nonathymidyl-3'-phosphate, calculated mass 2755.8 Da) appeared as a result of acid hydrolysis at the P-N bond location indicated in Scheme 1. It is important to note that we did not detect any fragments longer than the 9mer, specifically the full-length 14mer (calculated mass 4195.8 Da) or the

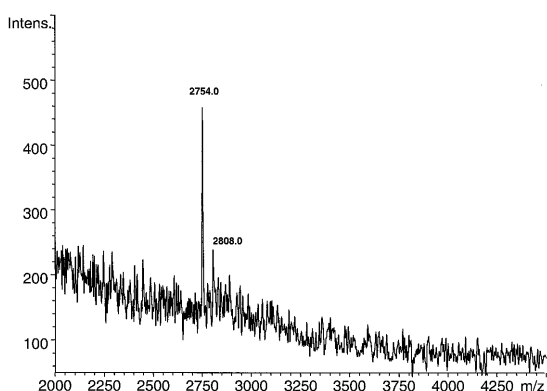


Figure 1. Selective release of the oligodeoxyribonucleotide fragment from the solid support by matrix-assisted P-N bond cleavage. Nonathymidyl-3'-phosphate [3'-(p)T₉] was selectively released from the surface by treatment with acidified matrix (3-HPA with 1.5% TFA) and detected by MALDI-TOF MS. 3'-(p)T₉ is the expected fragment from P-N bond cleavage of 3'-T₄(p)T-NH-(p)T-T₈-5', which was synthesized directly on a silicon chip (calculated mass 2755.8 Da, found 2754.0 *m/z*). The signal at 2808.0 *m/z* indicates incomplete removal of the 2-cyanoethoxy group during deprotection ($\Delta m = 54$ *m/z*).

second shorter fragment 5'-NH₂-T-(p)-T₄ (calculated mass 1458 Da). This proved that the cleavage was site-specific and that the oligonucleotide was covalently attached to the solid support. It should be emphasized, however, that experiments on solid supports showed generally weaker signal-to-noise ratios than those in solution. Minor impurities occasionally detected by MALDI-TOF MS were related to (*n* - 1) fragments (octathymidylate, calculated mass 2451.6 Da; data not shown) or incomplete removal of the cyanoethyl protecting group during ammonia deprotection (the 2808 *m/z* peak in Fig. 1).

To evaluate the cleavage properties of an internucleoside P-N bond in solution, oligonucleotides **1** and **2** were synthesized (see Materials and Methods), incubated with acidified 3-HPA MALDI matrix and analyzed by MALDI-TOF MS. Both primers contained single P-N bonds positioned towards the 3'-end. The P-N bond was successfully hydrolyzed in the course of co-crystallization of oligonucleotides and matrix (Fig. 2; data for oligonucleotide **1** not shown). Nearly complete cleavage was observed, although a residual amount of the original 6214 Da fragment was visible. Attempts to achieve complete cleavage by increasing the matrix acid concentration led to DNA depurination and appearance of additional shorter species (data not shown).

To test the stability of P-N bond-containing primers in the course of enzymatic thermal cycling, a 42 bp fragment of the MHTFR gene exon 4 was amplified from genomic DNA using oligonucleotides **1** and **2** as described in Materials and Methods. A PCR product of the correct size was obtained (data not shown). P-N bond-containing primers were fully compatible with thermal cycling and HotStar (*Taq*) DNA polymerase was able to synthesize DNA across the modified bonds. The P-N bond-containing primers are therefore likely to be suitable for DNA shortening applications involving products of thermal amplification.

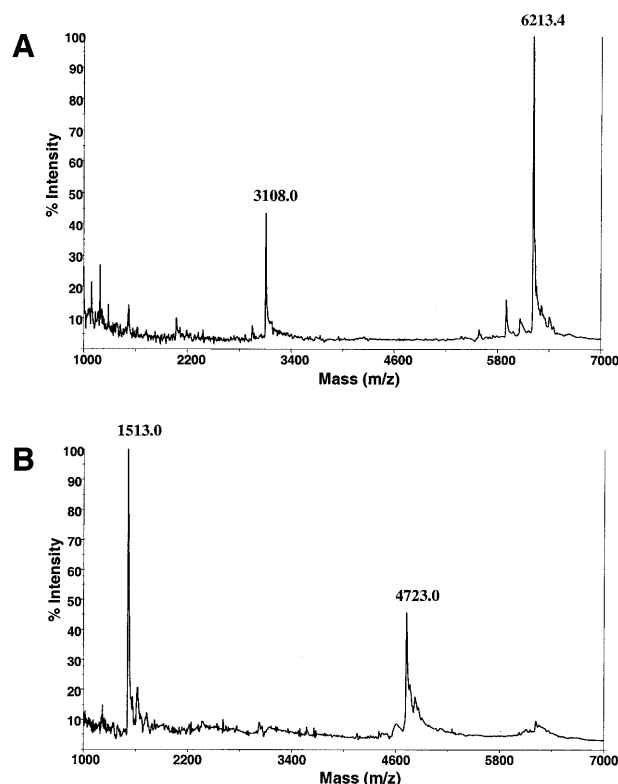


Figure 2. Cleavage of P-N bond-containing oligonucleotides in solution. Oligonucleotide **2** (see Materials and Methods) was subjected to cleavage in the presence of 3-HPA matrix acidified with 0.75% TFA followed by analysis by MALDI-TOF MS. (A) Oligonucleotide exposed to regular matrix. Note the smaller size peak at 3108 *m/z* corresponding to a double charged species of a 6213 *m/z* original fragment (calculated mass 6214 Da). (B) Oligonucleotide exposed to acidified matrix (calculated masses 1512 and 4722 Da).

Enzymatic incorporation of the P-N bond into DNA

The possibility of enzymatic incorporation of P-N bond-containing nucleotides into a newly synthesized DNA opens up the appealing option of using a simple acid cleavage scheme in sequencing-like reactions readable by MS. In order to obtain a complete read-out from any sequence, one needs to incorporate just two degradable purine or pyrimidine nucleotides, provided the template is read in both directions. A total of four reactions must be run if one degradable nucleotide is included, but only two reactions may be necessary (depending on the application) if polymerization is carried out with two degradable analogs. Modified pyrimidine triphosphates P-N dTTP and P-N dCTP were chosen for these experiments and were obtained as described (8,9). Based on the earlier discovery of Feldman and Thilo (14) that primary amines react with trimetaphosphates to yield corresponding amidotriphosphates, Letsinger *et al.* (9) synthesized 5'-amidotriphosphates in one step starting from 5'-deoxy-5'-aminonucleosides. Because the synthesis of 5'-aminonucleosides is quite straightforward, it makes amidotriphosphates even more available than standard 3'-deoxytriphosphates, which are synthesized in several steps from 5'-hydroxynucleosides. The P-N bond appeared to be sufficiently stable in both triphosphates and as an internucleoside

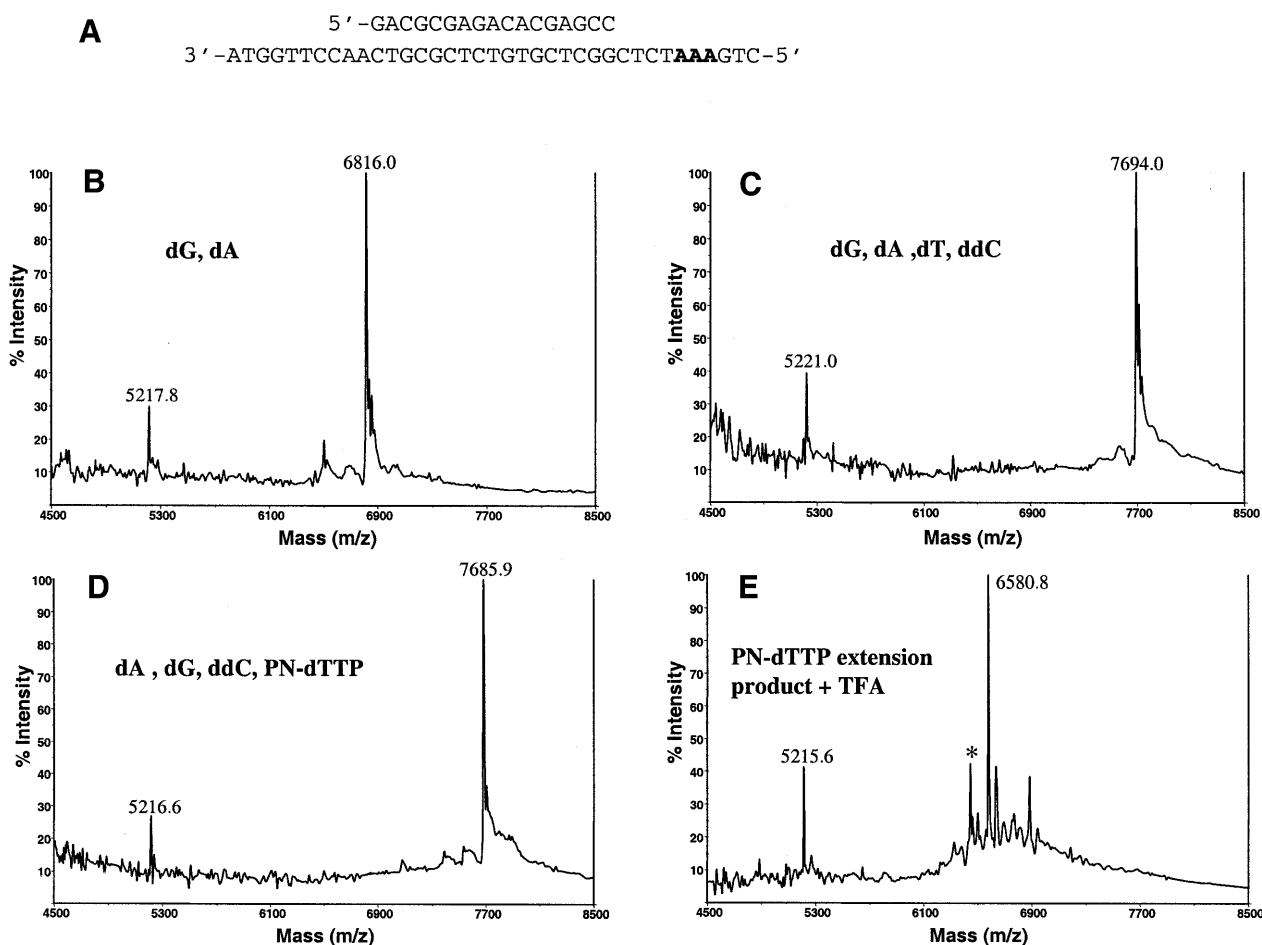


Figure 3. Sequenase-mediated primer extension on an artificial template. Polymerization was carried out as described in Materials and Methods. Masses of the full extension products are indicated in parentheses. (A) Double-stranded DNA template. (B) Incorporation of dATP and dGTP (calculated mass 6813.4 Da, measured mass 6816.0 *m/z*). (C) Incorporation of dATP, dGTP, ddCTP and dTTP (calculated mass 7686 Da, measured mass 7694.0 *m/z*). (D) Incorporation of dATP, dGTP, ddCTP and P-N dTTP (calculated mass 7683 Da, measured mass 7685.9 *m/z*). (E) Product of the reaction in (D) cleaved with 0.75% TFA-acidified 3-HPA matrix (calculated mass 6580.2 Da, measured mass 6580.8 *m/z*). The asterisk indicates a doublet (measured masses 6447.1 and 6464.0 *m/z*) resulting from depurination of a 6580.8 *m/z* extension product. Calculated mass of the unextended primer is 5215.4 Da, measured masses in (B), (C), (D) and (E) are 5217.8, 5221.0, 5216.6 and 5215.6 *m/z*, respectively. Measured masses differ from the calculated masses due to separate calibrations for each experiment.

bond. The bond was quite resistant to breakage in neutral and basic aqueous buffers, including those used during enzymatic synthesis and isolation of polynucleotides.

Several commercially available DNA polymerases were tested for their ability to incorporate 5'-triphosphates of 5'-aminonucleosides (Table 1). A preannealed oligonucleotide template was used (Fig. 3A) to determine polymerization efficiency. This double-stranded template was useful to test P-N dTTP incorporation in the upper strand, because the lower template strand had a higher mass (11308.4 Da) that would not interfere with measuring the shorter extension products. The template also had a stretch of three consecutive dA residues followed by a dG so that the extension could be terminated with ddCTP. The unextended upper primer of 5215.4 Da served as an internal mass calibrant. We found two 3'→5' exonuclease-deficient enzymes capable of incorporating P-N-modified triphosphates into DNA. Incorporation of both P-N dTTP and P-N dCTP by Sequenase 2.0 at 37°C was most efficient. Figure 3B shows that polymerization with only dATP and dGTP led to incorporation of an extra dA (calculated mass

Table 1. Incorporation of P-N dTTP by various DNA polymerases

DNA polymerase	Incorporation
Sequenase v.2.0	++++
Vent (exo-)	++
Deep Vent (exo-)	-
HotStar <i>Taq</i>	-
Thermo Sequenase	+
<i>Tth</i>	+
TaKaRa <i>Taq</i>	+
DyNAzyme I	+
DyNAzyme II	+

++++, excellent incorporation; ++, slow incorporation; +, poor incorporation; -, little or no incorporation.

6813.4 Da) immediately after incorporation over CTCT and upon encountering the triplet AAA. This proved the suitability

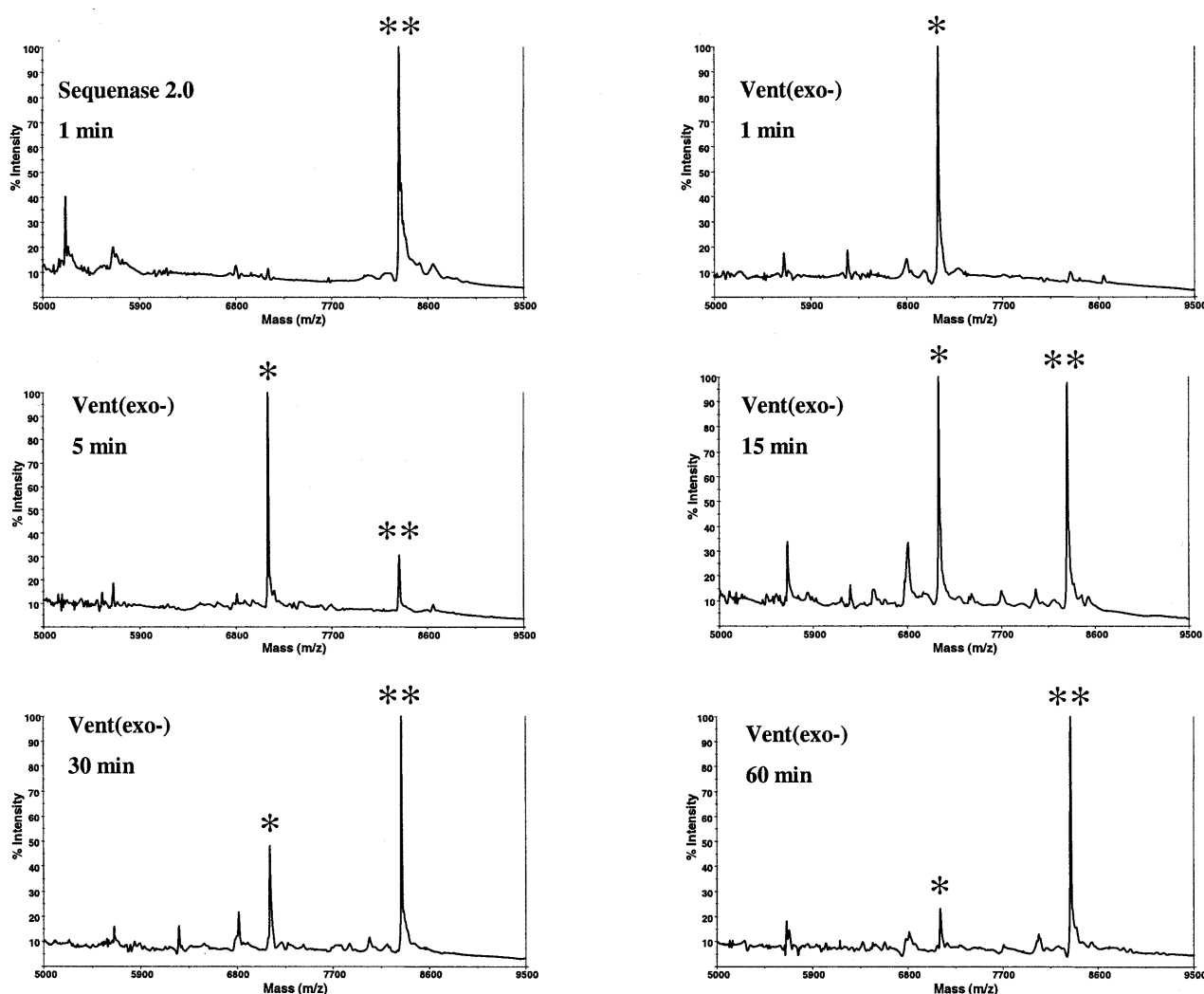


Figure 4. Vent(exo⁻)-mediated primer extension. Polymerization was carried out as described in Materials and Methods using an artificial primed template (Fig. 3A). No dideoxy terminator was used. Panels represent MALDI-TOF MS analysis of extension reactions processed using the regular 3-HPA matrix. Sequenase extension (upper left) is shown for comparison. Extension times are indicated. *, pausing peak (calculated mass 7106.6 Da) after addition of two consecutive dT; **, full extension product (calculated mass 8341.4 Da).

of the template for MALDI-TOF MS detection of extended products.

When polymerization was terminated at the first dC using a ddCTP terminator, no MS difference was seen between reactions that included modified P-N dTTP or its normal counterpart (Fig. 3C and D). As shown in Figure 3C, the measured mass difference between an extension product and an unextended primer is 2473 Da (calculated difference 2470.6). This reflects the combined molecular weights of natural nucleotide triphosphates incorporated by a polymerase. The respective measured mass difference in Figure 3D is ~3.7 Da smaller (calculated difference 3 Da smaller). Considering the mass resolution limit of MS (~1 Da), this confirms that three modified dT analogs were incorporated, each having a molecular weight 1 Da lower than their natural dT counterparts. Furthermore, the P-N dTTP extension product was successfully cleaved by exposure to acidified matrix (Fig. 3E). Sequenase efficiently polymerized over a stretch of three

adenine nucleotides and no pausing peaks were seen. All extensions were effectively completed in 1 min. It should also be noted that minor extra fragmentation peaks resulting from depurination under low pH conditions were detected. The asterisk in Figure 3E indicates depurination peaks 6447.1 and 6464.0 *m/z* obtained as a result of the loss of purine bases (both guanine and adenine) from a 6580.8 *m/z* extension product. A different profile of P-N dTTP incorporation was observed with Vent(exo⁻) polymerase. In contrast to Sequenase, after 1 min incubation at 72°C, strong pausing at the second dA in the AAA template stretch was apparent and the first full extension products appeared only after 5 min incubation (Fig. 4). Moreover, even after 1 h at this temperature, extension was not complete and a pausing peak resulting from addition of only two consecutive dTs in the polymerization product was seen. Therefore Vent(exo⁻) has stricter requirements for incorporation of the P-N nucleotide analogs, especially in contiguous stretches of identical bases. This lack of efficiency would

prevent the use of P-N-modified nucleotide triphosphates in PCR-like cycling reactions. On the other hand, effective incorporation of P-N dTTP and P-N dCTP analogs by Sequenase makes it possible to determine the complete sequence of an amplified DNA by annealing and extending primers complementary to either strand of the fragment by linear amplification.

Site-specific cleavage of P-N-modified polynucleotides

The simple approach that we have developed for acquiring site-specific cleavage information is based on a standard MassEXTEND assay (3,5). Single-stranded DNA templates were obtained by immobilizing the biotinylated PCR products on streptavidin-coated magnetic beads, then denaturing and washing out the non-biotinylated strand. An extension primer was then annealed to the single-stranded template and extended with Sequenase while still attached to the beads. Extension products were isolated and processed by MALDI-TOF MS using an acidic 3-HPA matrix. To expedite the fragmentation, an additional TFA treatment step may be introduced prior to matrixing. Preliminary experiments with fully extended product showed the appearance of concomitant depurination peaks. As is clear from Figures 1–3, while supporting effective breakage of the DNA strands at the desired sites, the acidic environment also causes some degree of depurination, which complicates interpretation of the results and is therefore undesirable. It is well known that when treated with acids, purine 2'-deoxynucleosides and 2'-deoxynucleotides hydrolyze to form free bases and sugars. Pyrimidine 2'-deoxynucleosides, such as dT and dC, are substantially more stable under these conditions. The effect is particularly pronounced in the case of adenosine. It was suggested earlier (15) that N-3 in the six-membered ring of purines (as well as the exocyclic 2-amino group of isocytidine) in the *syn* conformation helps to protonate the 4'-(ring)-oxygen of the ribose (step I), which triggers the hydrolysis of the N-glycosidic bond (step II). [It has been speculated that a specific type of depurination takes place during MALDI-TOF, which is possibly caused by laser irradiation as well as the presence of the matrix (16). Unlike acidic cleavage in solution, the depurinated product lacks a hydroxyl group at the 1' position, which causes the mass shift of 17 Da.] The rates of these steps are somewhat comparable. Therefore, reducing the rate of any of these two steps will decrease the total rate of depurination. To decelerate step I, purine analogs with a redistributed electron density pattern on the *endo* atoms of the rings but with the same exocyclic groups retaining the hydrogen bond pattern can be used. Commercially available examples include 7-deaza-adenosine and 7-deaza-guanosine (17), which have decreased electron densities at N-3. To reduce the rate of step II, the stability of the reaction intermediate carrying a positive charge at C-1' should be decreased. This can be achieved by introducing electron-withdrawing groups or atoms such as hydroxyl or fluorine at C-1' or C-2'. The simplest example is purine ribonucleotides, which are much more stable to acids compared to their 2'-deoxy analogs.

To reduce the rate of depurination, we used triphosphates of 7-deazapurines (A and G) instead of non-modified dATP and dGTP in Sequenase-mediated incorporation reactions. This combination dramatically improved the MS spectra (Fig. 5A). Extension reactions on an LPL template were performed with

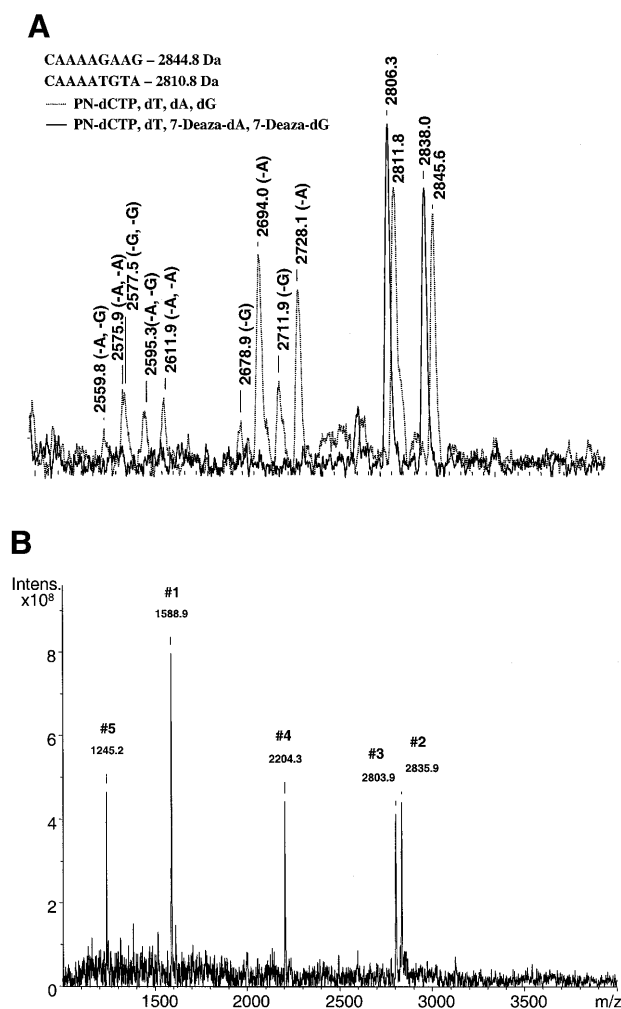


Figure 5. Sequenase-mediated primer extension on an immobilized single-stranded PCR product. (A) Suppression of depurination by deaza analogs. LPL extension products were obtained in reactions without (dotted line) or with (solid line) deaza-dA/GTP. Acidic cleavage of a 90mer P-N-modified LPL extension product was at 37°C. Loss of purine bases is indicated above each peak. Calculated masses are shown for fragments without deaza-dA/GTP. (B) Primer extension on a streptavidin bead-immobilized single-stranded LPL PCR product. The reaction was carried out as described in Materials and Methods using deaza-dGTP, deaza-dATP, dTTP and P-N dCTP. The acidic matrix contained 0.1% HCl. P-N dCTP (in bold) was incorporated into the sequence d(**TCTGGGCTATGAGATCAATAAAGT**CAGAGCCAAAAGAA-GCAGCAAAATGTACCTGAAGACTCGTTCTCCTGTGTGAAATTGTTA-TCCGCT). The extension primer is underlined. The following calculated fragmentation peaks are indicated by numbers: 1, CAGAG (1588 Da); 2, CAAAAGAAG (2836.8 Da); 3, CAAAATGTA (2804.8 Da); 4, CTGAAGA (2204.4 Da); 5, CGTT (1243.8 Da). The extension primer is located in fragment **TCTGGGCTATGAGATCAATAAAGT** (7491.8 Da; not shown).

deaza-dGTP, deaza-dATP, dTTP and P-N dCTP in place of normal dCTP and compared to reactions with no deaza compounds added. Extension products were cleaved using the acidified 3-HPA matrix at 37°C. In deaza-containing reactions, no concomitant depurination products were detected, yet the rate of incorporation and the completion of acidic cleavage were not compromised. Note that the masses of the deaza-dA/G-containing fragments are lower in proportion to the number of

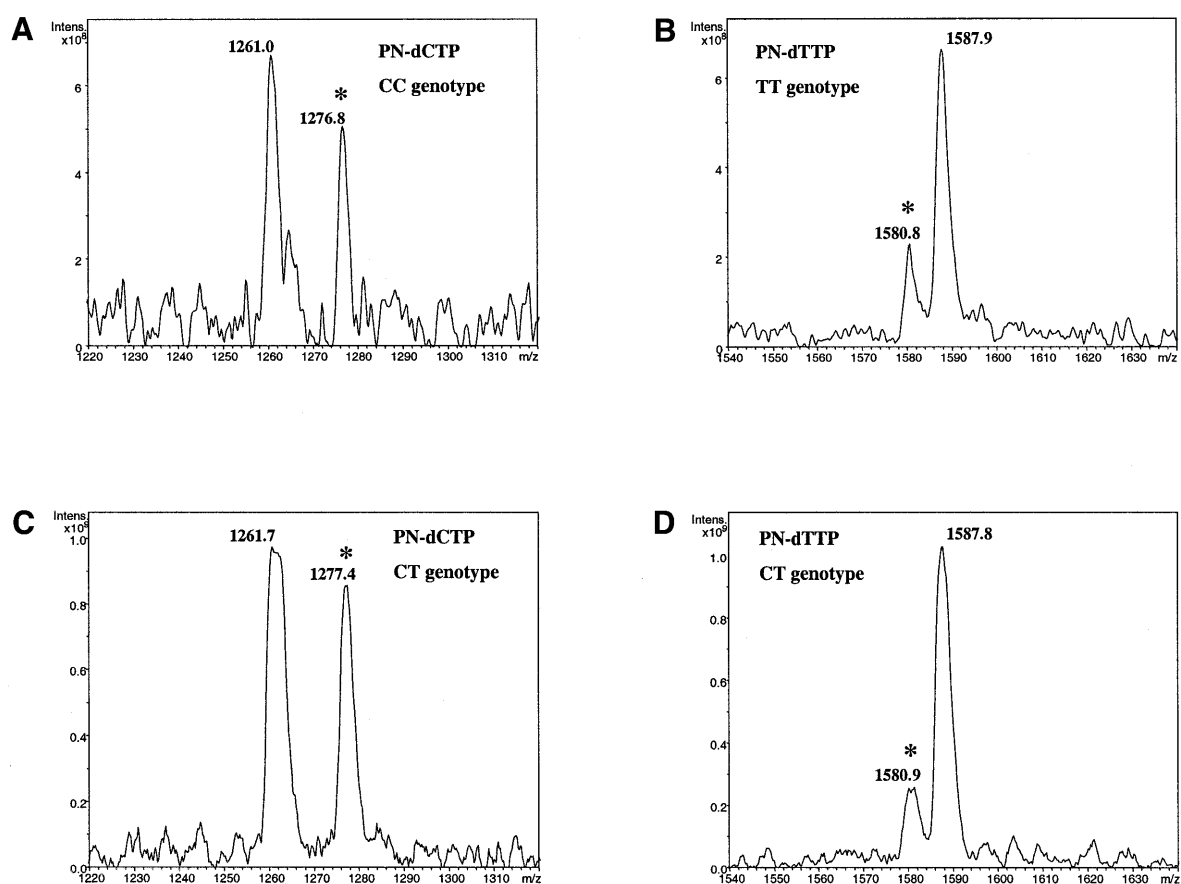


Figure 6. Genotyping by chemical cleavage of a P-N bond. A 174 bp human ADRB3 PCR product was used as template in primer extension reactions with P-N dCTP (A and C) and P-N dTTP (B and D) and deaza-dGTP and dATP. Modified analogs were incorporated into the sequence TGGTCATCGTGGCCATCGCC[T/C]GGACTCCGAGACTCCAGACCATGACCACCTGTGTGAAATTGTTATCCGCT, where a genotyping primer immediately upstream of a polymorphic site is underlined. Genotypes designated were determined in previous experiments using a PROBE assay. Asterisks indicate characteristic genotype peaks as explained in the text. Calculated mass of C allele 1275.8 Da; calculated mass of T allele 1580.0 Da.

deaza nucleotides in the fragment. This is due to a deaza nucleotide mass that is 1 Da less than the mass of a normal nucleotide.

Figure 5B shows a full-length acid cleavage pattern of a 90mer LPL extension product. Notably, all but one expected cleavage fragments are present. The most distal fragment d(CTGTGTGAAATTGTTAT) was not detected, conceivably because the biotin-streptavidin complex interferes with full extension. This is consistent with other primer extension experiments (results not shown). This indicates that the primer was extended, the P-N nucleotide analog was incorporated at the correct positions and the product was efficiently hydrolyzed at these positions. Thus, the primary sequence information (in this case C) was obtained by performing a simple treatment of the extension product with weak acid followed by highly accurate MS determination of resulting cleavage fragments. It should also be noted that minor extra fragmentation peaks resulting from depurination under low pH conditions were detected.

Acid cleavage-assisted genotyping on P-N-modified fragments

Based on the above result, a direct genotyping experiment was performed (Fig. 6). We have interrogated a polymorphic [T/C]

site in the human ADRB3 sequence (RefSNP ID rs#4994). Both the homozygous CC and TT and heterozygous CT genotypes were investigated. P-N dCTP and P-N dTTP triphosphates were incorporated in separate deaza-dA/GTP-complemented polymerization reactions and the products exposed to acidic matrix and measured by MS. Correct genotypes were unambiguously determined based on appearance of signature peaks at 1276 and 1580 m/z (calculated masses 1275.8 and 1580.0 Da). A peak at 1276 m/z is characteristic of the C allele and a peak at 1580 m/z is indicative of the T allele in P-N dCTP and P-N dTTP incorporation reactions, respectively. Additional peaks of 1260 (Fig. 6A and C) and 1587 m/z (Fig. 6B and D) corresponded to downstream cleavage products CAGA and TGAAA, respectively. This shows that in the case of C/T variation, incorporation of a P-N dC or P-N dT at the SNP site can introduce an allele-specific chemical cleavage site. Importantly, when a variable site sequence does not support allele-specific cleavage, the correct genotype can be deduced from the characteristic mass of the SNP-containing fragment. We did not attempt to perform extension reactions on the complementary strand, which may be necessary for genotyping certain sequences. Appropriate incorporation/cleavage schemes should be determined for each particular polymorphism.

Altogether, these experiments confirmed the high accuracy of the chemical cleavage method.

CONCLUSIONS

We describe a novel approach that greatly simplifies determination of DNA sequence information. It is achieved through base-specific cleavage followed by highly accurate MALDI-TOF MS analysis of the fragmentation spectrum. Base by base cleavage of the polynucleotides may be accomplished using a chemical base degradation method pioneered by Maxam and Gilbert (18), which is time-consuming and requires harsh chemicals, and a variety of similar multistep approaches. The DNA fragments obtained in most of these protocols are incompatible with MALDI-TOF MS unless purified either by precipitation or through affinity, size exclusion or reverse phase columns. Exploiting a different principle, Sanger *et al.* (19) introduced polymerization-based sequencing. This reaction benefits from enzymatic incorporation of dideoxy analogs of nucleotide triphosphates, which do not support subsequent polymerization. The present procedure appears as a combination of DNA breakage and polymerization methods. The procedure does not involve removal of the corresponding base, but instead targets and directly cuts the modified sugar-phosphate backbone. It is based on incorporation of acid-labile 5' P-N bond-containing nucleotide analogs into a polynucleotide, thus allowing selective cleavage at these defined locations. The breakage products generated have well-defined and stable 5'-amino and 3'-phosphate termini (Scheme 1) and may be easily detected by MALDI-TOF MS under conditions optimized for detection of nucleic acids. A minor loss of adenine and guanine takes place concurrently with P-N bond hydrolysis, but is completely suppressed when 7-deaza analogs of dA and dG are used. Oligo/polynucleotides containing P-N bonds retain many diagnostically useful properties of nucleic acids, such as complementarity and ability to hybridize and participate in enzymatic reactions, as well as stability of the P-N bonds towards enzymatic hydrolysis (7).

This method has great potential in high throughput genotyping due to its simplicity, robustness and compatibility with MS. The complete sequence of the region under interrogation may be obtained by running forward and reverse polymerization reactions followed by fully automated MS analysis on a chip (3,5). When mixtures of stable and cleavable triphosphates are used, such as, for example, 50% P-N dTTP and 50% unmodified dTTP, and because the modified phosphoramidite bond is cleaved at all locations, the whole spectrum of resulting fragments lies within the range of good MS resolution for oligonucleotides (300–15000 *m/z*). Moreover, multiple polymorphic sites within a defined region could be sequenced simultaneously in a single run. Heterozygous samples are determined in the same reaction and confirmed by the data from the opposite strand. The possibility of P-N bond incorporation into oligonucleotides attached to a solid support may suggest additional applications. Modified strands may be used as capture probes for hybridization and/or sequence-dependent extension and may be detached and analyzed by MALDI-TOF MS in a single step. These properties of P-N bond-based cleavage of DNA are invaluable for high throughput, large scale diagnostics of polymorphic sequences.

ACKNOWLEDGEMENTS

We wish to express our gratitude to Dr S. M. Gryaznov for fruitful discussions. We also thank the members of the Sequenom team for their excellent support.

REFERENCES

- Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860–921.
- Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A. *et al.* (2001) The sequence of the human genome. *Science*, **291**, 1304–1351.
- Tang, K., Fu, D.-J., Julien, D., Braun, A., Cantor, C.R. and Köster, H. (1999) Chip-based genotyping by mass spectrometry. *Proc. Natl Acad. Sci. USA*, **96**, 10016–10020.
- Little, D.P., Braun, A., O'Donnel, M.J. and Köster, H. (1997) Mass spectrometry from miniaturized arrays for full comparative DNA analysis. *Nature Med.*, **3**, 1413–1416.
- Buetow, K.H., Edmonson, M., MacDonald, R., Clifford, R., Yip, P., Kelley, J., Little, D.P., Strausberg, R., Köster, H., Cantor, C.R. and Braun, A. (2001) High-throughput development and characterization of a genomewide collection of gene-based single nucleotide polymorphism markers by chip-based matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Proc. Natl Acad. Sci. USA*, **98**, 581–584.
- Fu, D., Tang, K., Braun, A., Reuter, D., Darnhofer-Demar, B., Little, D.P., O'Donnel, M.J., Cantor, C.R. and Köster, H. (1998) Sequencing exons 5 to 8 of the *p53* gene by MALDI-TOF mass-spectrometry. *Nat. Biotechnol.*, **16**, 381–384.
- Letsinger, R.L. and Mungall, W.S. (1970) Phosphoramidate analogs of oligonucleotides. *J. Org. Chem.*, **35**, 3800–3802.
- Letsinger, R.L., Wilkes, J.S. and Dumas, L.B. (1975) Incorporation of 5'-amino-5'-deoxythymidine 5'-phosphate in polynucleotides by use of DNA polymerase I and a ϕ X174 DNA template. *Biochemistry*, **15**, 2810–2816.
- Letsinger, R.L., Wilkes, J.S. and Dumas, L.B. (1972) Enzymatic synthesis of polydeoxyribonucleotides possessing internucleotide phosphoramidate bonds. *J. Am. Chem. Soc.*, **94**, 292–293.
- Caruthers, M.H. (1991) Chemical synthesis of DNA and DNA analogues. *Acc. Chem. Res.*, **24**, 278–284.
- Wu, K.J., Steding, A. and Becker, C.H. (1993) Matrix-assisted laser desorption time-of-flight mass spectrometry of oligonucleotides using 3-hydroxycyclohexanecarboxylic acid as an ultraviolet-sensitive matrix. *Rapid Commun. Mass Spectrom.*, **7**, 142–146.
- Guo, Z., Guilfoyle, R.A., Thiel, A.J., Wang, R. and Smith, L.M. (1994) Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports. *Nucleic Acids Res.*, **22**, 5456–5465.
- Southern, E.M., Maskos, U. and Elder, J.K. (1992) Analyzing and comparing nucleic acid sequences by hybridization to arrays of oligonucleotides: evaluation using experimental models. *Genomics*, **35**, 217–227.
- Feldman, V. and Thilo, E.Z. (1964) Zur Chemie der kondensierten Phosphate und Arsenate. XL. Über das N-Methyl- und das N-Äthylamidotriphosphat, $[P_3O_9NHCH_3]^{4-}$ und $[P_3O_9NHC_2H_5]^{4-}$. *Z. Anorg. Allg. Chem.*, **327**, 159–172.
- Dekker, C.A. (1960) Nucleic acids: selected topics related to their enzymology and chemistry. *Annu. Rev. Biochem.*, **29**, 453–486.
- Nordhoff, E., Kirpekar, F. and Roepstorff, R. (1996) Mass spectrometry of nucleic acids. *Mass Spectrom. Rev.*, **15**, 67–138.
- Siegert, C.W., Jacob, A. and Köster, H. (1996) Matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry for the detection of polymerase chain reaction products containing 7-deazapurine moieties. *Anal. Biochem.*, **243**, 55–65.
- Maxam, A.M. and Gilbert, W. (1977) A new method for sequencing DNA. *Proc. Natl Acad. Sci. USA*, **74**, 560–564.
- Sanger, F., Air, G.M., Barrell, B.G., Brown, N.L., Coulson, A.R., Fiddes, C.A., Hutchison, C.A., Slocombe, P.M. and Smith, M. (1977) Nucleotide sequence of bacteriophage phi X174 DNA. *Nature*, **265**, 687–670.