Facile method for automated genotyping of single nucleotide polymorphisms by mass spectrometry

Sascha Sauer¹,², David H. Gelfand³, Francis Boussicault², Keith Bauer³, Fred Reichert³ and Ivo G. Gut²,*

¹Freie Universität Berlin, Fachbereich Biologie, Chemie, Pharmazie, Takustrasse 3, 14195 Berlin-Dahlem, Germany,
²Centre National de Génotypage, Bâtiment G2, 2 Rue Gaston Crémieux, 91057 Evry Cedex, France and
³Roche Molecular Systems, Inc., 1145 Atlantic Avenue, Alameda, CA 94501, USA

Received November 27, 2001; Revised and Accepted January 16, 2002

ABSTRACT

In the future, analysis of single nucleotide polymorphisms (SNPs) should become a powerful tool for many genetic applications in areas such as association studies, pharmacogenetics and traceability in the agro-alimentary sector. A number of technologies have been developed for high-throughput genotyping of SNPs. Here we present the simplified GOOD assay for SNP genotyping by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI). The simplified GOOD assay is a single-tube, purification-free, three-step procedure consisting of PCR, primer extension and phosphodiesterase II digestion followed by mass spectrometric analysis. Due to the application of charge-tag technology, no sample purification is required prior to the otherwise very impurity-sensitive MALDI analysis. The use of methylphosphonate-containing primers and ddNTPs or α-S-ddNTPs together with a novel DNA polymerase derived from Thermotoga maritima for primer extension allow the fluent preparation of negatively charge-tagged, allele-specific products. A key feature of this polymerase is its preference for ddNTPs and α-S-ddNTPs over dNTPs. The simplified GOOD assay was run with automatic liquid handling at the lowest manageable volumes, automatic data acquisition and interpretation. We applied this novel procedure to genotyping SNPs of candidate genes for hypertension and cardiovascular disease.

INTRODUCTION

With the near completion of the human genome sequencing project, the emphasis of genomic research has shifted towards extraction of information on gene function. Analysis of the relationship between DNA variants and phenotypes, such as disease status or quantitative trait variables, is one of the most powerful tools to obtain such information. Genetic linkage and association or linkage disequilibrium mapping will increasingly rely on the characterization of dense maps of single nucleotide polymorphisms (SNPs) in very large patient and population cohorts (1–3). In the future, SNP analysis should also become a powerful tool for many applications in areas such as pharmacogenetics and traceability in the agro-alimentary sector (4). Methods for high-quality and low-cost SNP genotyping that are amenable to very high throughput are, thus, of great importance (5). Many methods such as DNA microarrays, gel-based and plate-reader based assays for SNP genotyping have been introduced (6–10).

In principle, mass spectrometry provides one of the most attractive solutions for SNP genotyping because it can be used to obtain direct and rapid measurement of DNA. In particular, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI) has revolutionized the analysis of biomolecules (11). A number of methods using MALDI have been developed for high-throughput SNP genotyping (12–14). Nevertheless, most SNP genotyping methods using highly impurity-sensitive MALDI analysis require stringent purification procedures, such as magnetic bead separation or reversed-phase binding. In general, purification is expensive and cumbersome in automation (5,10). Here we present the simplified GOOD assay for SNP genotyping using MALDI, which represents a significant advance over previously described procedures (15,16). The simplified GOOD assay is a single-tube, purification-free, three-step method consisting of PCR, primer extension and phosphodiesterase II digestion immediately followed by MALDI analysis. The key to it lies in charge tagging—a chemical modification strategy—which improves the detection sensitivity 100-fold and renders the allele-specific products insensitive to impurities such as the preparation buffers in MALDI. This is achieved by conditioning products to carry either excess positive or single excess negative charge (17–19). Here we describe two crucial improvements to the standard GOOD assay: first, the inclusion of extension primers containing methylphosphonate groups on their 3′-ends; secondly, the application of a novel DNA polymerase, Tma 31 FS, derived from Thermotoga maritima, for the primer extension reaction. The inclusion of methylphosphonates means that the alkylation reaction with toxic methyl iodide of the standard GOOD assay can be avoided and, thus, the simplified GOOD assay is void of chemistry. Tma 31 FS DNA polymerase accepts methylphosphonate-containing primers and preferentially incorporates ddNTPs over dNTPs. We demonstrate here that
 MATERIALS AND METHODS

Oligonucleotides for PCR were synthesized by MWG Biotech (Ebersberg, Germany) and were all HPSF purified. Oligonucleotides containing methylphosphonates were synthesized by Eurogentec (Liège, Belgium) and PAGE purified. dNTPs, ddNTPs and phosphodiesterase II (from calf spleen) were purchased from Roche Diagnostics (Mannheim, Germany). Platinum™ Taq DNA polymerase high fidelity was obtained from Gibco BRL (Karlsruhe, Germany). α-S-ddNTPs were obtained from Biolog (Bremen, Germany). Chemical reagents were purchased from Aldrich (Steinheim, Germany). Tna 31 FS DNA polymerase was provided by Roche Molecular Systems (Alameda, CA).

α-Cyano-4-hydroxy-cinnamic acid methyl ester used as MALDI matrix for the simplified GOOD assay is available from Bruker Saxonia Analytik GmbH (Leipzig, Germany). The BasePlate robot from The Automation Partnership (Royston, UK) was used for liquid handling of reaction mixtures and for liquid handling at the lowest manageable volumes, automatic thermocycling procedures were carried out in an Eppendorf Gradient Thermocycler (Eppendorf, Germany) or a Primus Thermocycler from MWG Biotech (Ebersberg, Germany). A BasePlate robot from The Automation Partnership (Royston, UK) was used for liquid handling of reaction mixtures and for MALDI target preparation.

PCR

β-2-Adrenergic receptor. Human DNA (20 ng), purified as described by Sambrook et al. (20), was mixed with 3.5 pmol of the forward primer (5'-CTC GCG GCC CGC AGA GCC) and 3.5 pmol of the reverse primer (5'-GTT GGT GAC CGT CTG CAG ACG CTC), 40 mM Tris base (pH 8.8), 32 mM (NH₄)₂SO₄, 50 mM KCl, 2 mM MgCl₂, 40 µM dNTPs and 0.2 U Taq DNA polymerase in a 20 µl volume. The reaction was denatured at 95°C for 2 min, then thermostated for 20 s at 95°C, 30 s at 68°C and 30 s at 72°C, repeating the cycle 30 times.

Platelet/endothelial cell adhesion molecule (PECAM) G+58A. The forward primer (5'-ATG TTC CGA GAA GAA CAGmpAmpT) and 3.5 pmol of each) were mixed together with 5 ng genomic DNA and reverse primer (5'-GCA GGG CAG GTT CAT AAA TAA G) (2.5 pmol of each) were mixed together with 5 ng genomic DNA, 0.3 µl Platinum™ Taq DNA polymerase high fidelity (10×) buffer, 1 mM MgSO₄, 20 µM dNTPs and 0.1 U Platinum™ Taq DNA polymerase high fidelity in a 3 µl volume. The reaction was denatured for 4 min at 94°C, then thermostated for 30 s at 94°C, 45 s at 65°C and 30 s at 72°C, repeating 30 times.

Primer extension

For the analysis of SNPs in the β-2-adrenergic receptor gene, 20 pmol of the primer for SNP position 298 (5'-CCC GCC GTG GTG CCGmpCmpC) (mp, methylphosphonate) or 20 pmol of the primer for SNP position 325 (5'-CGC GCA GTC TGG CAGmpGmpT) were mixed with 40 mM Tris base (pH 8.8), 2 mM MgCl₂, 0.2 mM MnCl₂, 100 µM dNTPs or α-S-ddNTPs (ddCTP and ddTTP in the case of SNP 298, and ddGTP and ddATP in the case of SNP 325) and 1 U Tma 31 FS DNA polymerase in 10 µl. These 10 µl were added to the preceding PCR. An initial denaturing step of 1 min at 95°C was used followed by 35 cycles of 10 s at 95°C, 30 s at 58°C and 15 s at 72°C.

In the case of the SNP G+58A in the PECAM gene, 5 pmol of primer 5'-ATG TTC CGA GAA GAA CAGmpAmpT and 0.5 U Tma 31 FS DNA polymerase in 2 µl were added to the 3 µl PCR using the same reaction conditions as described for the primer extension of the β-2-adrenergic receptor gene.

RESULTS

Recently, we presented the first procedure for SNP genotyping using MALDI detection that does not require purification, termed the GOOD assay (referred to as the standard GOOD assay) (15,16). It requires five reagent additions into a single PCR tube or microtiter plate and incubations or thermostating. Its major drawback is the requirement of toxic methyliodide.

The principle of the simplified GOOD assay is illustrated in Figure 1. The procedure starts with a PCR to generate a sufficient amount of template and to reduce the complexity of genomic DNA, thus reducing the risk of mispriming. Immediately following this, primer extension with methylphosphonate containing primers and ddNTPs or α-S-ddNTPs yields allele-specific products. The methylphosphonate bridges of the primers resist the phosphodiesterase II digestion, which is used in the third step to remove the unmodified part of the primers. The resulting products for MALDI analysis have methylphosphonate linkages that are charge-neutral and one negative charge-carrying linkage deriving from the ddNTF or the α-S-ddNTF, rendering the allele-specific products negatively charge-tagged. The crude reaction mixture is transferred onto a MALDI target plate and analyzed. The simplified GOOD assay was applied to genotyping SNPs such as the β-2-adrenergic receptor gene and the PECAM gene (22,23). The analysis of SNPs 298 and 325...
in the β-2-adrenergic receptor gene is shown in Figures 2 and 3. SNP 298 is a C/T polymorphism, while SNP 325 is a G/A polymorphism. Efficient and reliable incorporation during primer extension of all four ddNTPs opposite according bases was observed. The correct analysis of genotypes was confirmed by DNA sequencing using 20 different patient DNAs.

The simplified GOOD assay was adapted to 384 microtiter plates employing a liquid-handling robot. Thereby, the easy automation of the procedure was demonstrated on the SNP G+58A in the PECAM gene, which is a G/A polymorphism. The starting reaction volume of the PCR was reduced to 3 µl. The final volume after phosphodiesterase II digestion was 7 µl. Samples (384) were prepared in one run with four robot movements of a 96-tip head for each of the three liquid dispensings. Microtiter plates were sealed and manually transferred to thermocyclers for the PCR and primer extension reaction, or an incubator for the phosphodiesterase digestion. MALDI targets were prepared by applying α-cyano-4-hydroxy-cinnamic acid methyl ester in a thin layer with the liquid-handling robot. Samples were applied on top of the dry matrix with the same robot. High reproducibility in automatic measurements was observed, accumulating on average only 10 laser shots per spectrum. Even so, spectra with signal-to-noise ratios higher than 20:1 and near baseline isotopic resolutions were obtained in the linear detection mode. Spectra were recorded and alleles called automatically using the GenoTools software (21). Three typical results of the 384 are shown in Figure 4A. In >95% of preparations results were obtained on first pass, which is in a similar range compared with the standard GOOD assay. Missing results were due to PCR failures, probably because of insufficient integrity of genomic DNA templates. The observed genotype frequencies were in good agreement with predictions under the assumption of the Hardy–Weinberg equilibrium.

The simplified GOOD assay demonstrated in Figure 4A was carried out with a primer containing two methylphosphonates at its 3′ end. We obtained respective results with a primer containing three methylphosphonates instead of two, thereby shifting masses of allele-specific products 311 Da; from 1177 Da to 1488 Da and from 1193 Da to 1504 Da (Fig. 4B). Additionally, α-S-ddNTPs in place of ddNTPs were used for primer extension under these conditions with comparable...
efficiency of incorporation (Fig. 4C). This provides an added degree of freedom for establishing SNP genotyping multiplexes.

**DISCUSSION**

The ability to multiplex reactions is of great importance to reduce reagent cost per SNP genotype, analysis time, and in order to make good use of the multi-channel detection capability of MALDI mass spectrometers. Different combinations of ddNTPs and α-S-ddNTPs and different numbers of methylphosphonate linkages allow the separation of alleles. This has interesting implications for the multiplexing of assays, as it adds degrees of freedom. This is similar to that recently demonstrated by another MALDI SNP genotyping method using primer extension, with the advantage that α-S-ddNTPs are less expensive than the fluorescent dideoxy nucleotides used there (24). Haff and coworkers showed that a 12-fold multiplex in the primer extension and subsequent MALDI analysis is possible (25). Nevertheless, the quality of the enzymatic reactions in a multiplex assay depends on the SNPs that are combined. Surrounding DNA sequences have a significant influence on the quality of the PCR. The trade-off for establishing an experiment for SNP genotyping with a high degree of multiplexing is the time required for optimization. The extent of multiplexing at the PCR level and of the primer extension reaction of the simplified GOOD assay is currently under investigation.

Key to the simplified GOOD assay is the Tma 31 FS DNA polymerase. In contrast with other commercially available DNA polymerases (e.g. Taq, Thermosequenase, DeepVent), this novel DNA polymerase readily extends primers containing several methylphosphonate linkages on their 3′ ends and efficiently incorporates ddNTPs and α-S-ddNTPs over dNTPs. This gives the simplified GOOD assay the added advantage that residual dNTPs of the PCR do not have to be removed. In the standard GOOD assay this was done by shrimp alkaline phosphatase digestion. Thus one liquid-handling step, the cost of reagents and time is saved. At a 5:1 ratio of dNTPs to ddNTPs, equal signal intensities of the primer extension were observed, yet under the conditions of the simplified GOOD assay the ratio is 1:5 (this experiment was carried out on synthetic templates and is not described in detail). Due to this preference Tma 31 FS DNA polymerase might also be the DNA polymerase of choice for other methods using primer extension after a PCR if the method, like the simplified GOOD assay, has a means of discriminating between the primers used for the PCR and the primer extension.

The simplified GOOD assay presented here is based on the same charge-state concept for DNA detection by MALDI as the standard GOOD assay, but requires fewer steps (three instead of five simple reagent additions). It also allows highly precise MALDI analysis straight from a crude reaction mixture. Simplification is achieved by the introduction of methylphosphonate groups, which are already charge neutral, in place of phosphorothioates. This has the advantage that the potentially toxic methylating reagent of the standard GOOD assay is avoided. Due to this, as is demonstrated here, the simplified GOOD assay is even easier to implement for high-throughput applications. It provides a genotyping method with low consumption of DNA and reagents, high accuracy of the results and complete flexibility on the choice of SNPs. The simplified GOOD assay lends itself to further miniaturization and could potentially be executed on a microfluidic device. We are currently implementing the procedures described here in a production line for high-throughput genotyping.
ACKNOWLEDGEMENTS

We thank Laetitia Dorel for technical assistance, and Mark Lathrop and Noah Christian for critical reading of this manuscript. This work was kindly supported by the French Ministry of Research (MENRT). S.S. was funded by the ‘Stiftung Stipendienfonds’ of the German Chemical Industry.

REFERENCES


