Discussion

The standard GOOD assay

In this work several novel procedures for genotyping SNPs termed the GOOD assays are presented.[124-126] Generally, the GOOD assays implement sensitivity increasing DNA modifications into a procedure for generating allele-specific products and subsequent analysis by MALDI mass spectrometry. The standard GOOD assay was successfully applied to genotyping a large number of SNPs in candidate genes for case-control studies at the Centre National de Génotypage. In contrast to other methods available for SNP analysis using mass spectrometry, the GOOD assay does not require any purification and was therefore much more amenable for automation. The implemented DNA modification chemistry contributes a significantly lower part to the cost of the analysis than expensive purification procedures such as magnetic beads, reverse-phase materials (e.g. "Zip-Tips") or DNA purification kits.[92,143,144] The standard GOOD assay has the added advantage over gel-based and DNA array procedures that each found allele gives an absolute, measured mass and therefore does not suffer from secondary structure formation. This significantly contributes to the ease of allele calling.

The standard GOOD assay starts with a PCR that actually fulfils two functions.[145,146] Firstly it generates a sufficient amount of template for the allele-specific processing, and secondly it reduces the sequence complexity of the template decreasing the risk of mispriming. There are very few SNP genotyping procedures that do not require the amplification of genomic DNA by PCR. The only published PCR free method is the Invader assay, which in turn might require large amounts of genomic DNA. For the GOOD assay several DNA polymerases and corresponding buffers can be applied making it straightforward for the development of genotyping assays of SNPs discovered by different methods such as DNA sequencing and denaturing HPLC.

The GOOD assay procedure was adapted to allow the use of low quality DNA without sacrificing robustness. This is important as it allows this method to be extended to streamlined genetic fingerprinting of domestic animals. The quality of SNP

genotyping experiments strongly depends on the quality of the DNA as was observed in large-scale applications of the GOOD assay performed at the CNG. Presumably low DNA integrity did not allow efficient PCR amplification. In most cases DNA is extracted from whole blood using commercially available kits. Because as little as 0.5-2.0 nanograms of genomic DNA are sufficient for genotyping a SNP by the GOOD assay, these DNA extractions tend to last a quite long time. The cost of extraction is spread over the number of SNPs that are genotyped. However, for example in agricultural applications it is inconceivable to use expensive and cumbersome extraction procedures. Therefore a system of taking tissue sample during ear-tagging of cattle has been adapted to the GOOD assay. Tissue samples are digested with Proteinase K and this preparation is immediately used for the PCR of the GOOD assay without any isolation of the DNA. The Proteinase K digest by-products apparently do not disturb the GOOD assay. As an example genotyping of SNP 129 in the prion protein of cows and humans is presented. The GOOD assays for SNP genotyping of bovine DNA could be applied to study large numbers of cattle in order to determine genetic susceptibilities. Alternatively, the basis is provided for establishing a technology platform for traceability.

The GOOD assay for positive ion mode detection was standardised to analyse pooled DNA samples. The obtained results were in a similar range as was shown for analogous procedures.[147] For example it was determined for SNPs in the β-2-adrenergic receptor gene that the GOOD assay can be used to identify one allele in ten individuals if pools are set up carefully. A high correlation of signal intensity and allele frequency was observed (R^2-value = 0.982-0.999). The allele frequency was determined with a relative error of 6.5-12.0 %, while an allele frequency of 50 % could be predicted not exceeding a fidelity of 4.8 % in the best case of SNP 390 in β-2-adrenergic receptor gene.[132] The mentioned values are still too high for pooled large-scale SNP genotyping. The main problem remained the accurate DNA concentration measurement, which was also the main obstacle for all other techniques.[148]

Microsatellite genotyping benefits from the large number of alleles each marker can provide. Yet, they require analysis by gel-based methods and do not provide information about coding changes. SNPs on the other hand give only binary information. The information content of SNPs can be increased if the haplotype of

multiple SNPs within a region is analysed physically.[149] The GOOD assay as with other assays for SNP genotyping employing PCR could be extended for haplotyping by integrating allele-specific PCR.[150,151] This way haplotypes within PCR fragments can be measured directly.

SNP	1	2	3
genotype	C/G	T/A	C/G
haplotype 1	C	T	C
:			
:			
haplotype 8	G	A	G

Figure 4.1. n SNPs in close proximity give rise 2^n haplotypes. With three SNPs these are 8 haplotypes with 36 possible genotypes.

dNTPs of the PCR had to be removed if the following primer extension was done with Thermosequenase. The dNTPs were enzymatically degraded by shrimp alkaline phosphatase. Later, it was shown that by replacing Thermosequenase with Tma 31 FS DNA polymerase, which has a significantly higher specificity for ddNTPs, this step could be omitted. This enzyme has a significant advantage over common DNA polymerases. At a 5:1 ratio of dNTPs to ddNTPs equal signal intensities of the primer extension were observed, yet under the conditions of the GOOD assays the ratio is approximately 1:5. 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 GOOD assays, has a way of discriminating between the primers used for the PCR and the primer extension.

In the next step of the GOOD assay a primer extension is done with a specifically tailored set of α-S-dNTPs and α-S-ddNTPs or ddNTPs. A primer extension reaction is used to generate allele-specific products. A primer is chosen upstream of the SNP that is to be genotyped. Primers can be placed on either strand of the PCR product. Nevertheless the positioning of a primer is restricted. Therefore it could be difficult to optimise the primer extension reaction for certain SNPs. Even if zwitterionic products with a net charge of 0 were well detectable, extension primers are still preferably synthesized with functionalities that will result in the final products being +1 or -1 in net charge. The last three bases at the 3'-end of the primer are connected with two phosphorothioate bridges for positive ion-mode analysis. The middle base has an amino-modification that allows the attachment of the positive charge-tag. First, the main limitation of the GOOD assay for positive ion mode detection was the commercial availability of only T^{NH2} for the introduction of a charge tag near the 3'-end of the extension primer. It limited the positioning of the primer for the extension reaction. Using a set of novel propargylamino-modified phosphoramidites, A^{NH2}, G^{NH2}, C^{NH2} and U^{NH2}, introduced in primer synthesis, full flexibility of positioning of a primer for the primer extension reaction was achieved. These phosphoramidites are now commercially available. By introducing a positive charge-tag, reagents with different masses could be attached to the amino-modified nucleobases. This allows mass tagging in order to shift masses for the mass spectrometric analysis of multiplex reactions. For the negative ion-mode, the nucleosides at the 3'-end of the primer are connected with phosphorothioate bridges. These linkages fulfil two functions. They are quantitatively charge neutralisable by alkylation and they inhibit the complete digestion of the primer in the following step of the procedure. For the primer extension reaction primers are added together with a specifically selected set of α-S-dNTPs, α-S-ddNTPs and ddNTPs. These substrates are readily accepted by a number of DNA polymerases. Their addition to the primer results in the formation of further phosphorothioate bridges. Typically, primers are placed immediately next to the position of the SNP and extended only with α-S-ddNTPs or ddNTPs. This results in the most homogenous signal patterns. The efficiency of a primer extension with ddNTPs and α-S-ddNTPs is similar. These substrates can be used to shift product masses as was demonstrated recently in another SNP genotyping

method using primer extension and MALDI-MS with the advantage that α-S-ddNTPs are less expensive than the fluorescent dideoxy nucleotides used there.[84]

However, if primers with more than three phosphorothioates were used no efficient extension was observed. By synthesizing phosphorothioates into oligonucleotides, stereocenters are introduced that cause significantly reduced hybridisation efficiency. Therefore it is not possible to use primers containing only phosphorothioates in order to omit the removal of the unmodified part of the extension primer. The efficiency of hybridisation and elongation of extension primers containing two phosphorothioates was increased by the use of oligonucleotides with the "good" stereo configuration as is shown in figure 3.18. HPLC purification of these primers particularly if more than three phosphorothioates are used might be a solution to increase the hybridisation and thus the primer extension efficiency, and might also increase the variety of primer design.

Another limitation of the use of phosphorothioates for primer extension is the isotopic distribution of sulphur. While ^{32}S has a relative abundance of 100 ^{34}S has a striking relative abundance of 4.442.[152] Hence the more sulphur atoms that are integrated in a DNA molecule for mass spectrometric analysis the more complex in terms of isotopic distribution the resulting spectrum is, giving dramatic peak broadening and sensitivity decrease of these oligonucleotides. Therefore it is in principal impossible to use alkylation of several phosphorothioates integrated in large DNA fragments, for example for sequencing or microsatellite analysis. Only mini-sequencing of some nucleobases is feasible.[124]

The following step of the GOOD assay consists in the digestion of a large part of the primer. The primer does not contribute to the information content of the allele-specific product. It serves only to recognise the complementary sequence on the PCR product. Therefore it can be removed and the size of the products to be analysed is significantly reduced. Thus the molecular weights of the products are shifted into a mass range where the detection sensitivity and resolution of the MALDI mass spectrometer is best. A disadvantage of the use of phosphodiesterase II might be in some cases the described sequence dependency of primer digestion.

The last step of the GOOD assay is the alkylation. The phosphorothioate bridges are charge neutralised by the alkylation reaction. The reaction conditions are chosen so that

the selectivity of the reaction is optimised for the addition of methyl groups to phosphorothioate bridges, while no alkylation of the bases takes place. The addition of the alkylating agent results in the generation of two phases. The allele-specific products reside in the upper aqueous phase. A part of this phase is taken and diluted in a sample solution. From there samples are transferred onto a MALDI target applying a thin layer preparation and then analysed. The product masses lie in a mass range of 1,000 to 2,000 Da.

As was shown recently, charge tagging and charge neutralisation of the phosphorothioate DNA backbone result in a dramatic increase in sensitivity and decrease in susceptibility to form adducts with constituents.[111,114] Due to the increased detection sensitivity of modified products the reaction mixture can be diluted. Through the dilution in the GOOD assay, the concentration of reaction constituents is decreased without the allele-specific products falling beneath the detection threshold. The biphasic system of the alkylation reaction additionally helps to remove some inhibitory reagents for the sample preparation (i.e. detergents from enzymes). The modified allele-specific products are detectable without requiring purification prior to MALDI analysis. The bandwidth of reagents that are tolerated by the process is quite large, so that different PCR conditions (DNA polymerases and buffers) can be coupled to the procedure. Nevertheless, the alkylation reaction of the GOOD assay can induce three disadvantages. The reaction can sometimes lead to under-alkylated or over-alkylated byproducts, which might be explained by different nucleobase compositions of products. Some nucleobases might be more susceptible to alkylation than others. More importantly, the methyliodide used for the alkylation is a very toxic chemical, which could be problematic for large-scale applications. Furthermore, miniaturisation is limited because of the alkylation reaction implicating a phase separation. Therefore this assay has to be performed in a microliter scale.

A current limitation of the standard GOOD assay as for alternative methods using MALDI-MS rests the low degree of multiplexing. On the other hand data accumulation by mass spectrometry is rapid. The trade-off for establishing an experiment with a high degree of multiplexing is the time that has to be invested for optimisation. Applications of SNP genotyping, like for traceability of cattle and their meat, merit an effort for the optimisation of reactions with a high degree of multiplexing, as these reactions might be

used millions of times afterwards. Here it would be a great benefit to establish stable multiplexes of tens of SNPs.

Mass spectrometric aspects of the GOOD assay

The nature of the matrix chosen for the MALDI analysis in the GOOD assay does not correspond to traditionally used matrices. It is a non-acidic matrix. As such, it might act as a discriminator for the modified products against other components of the procedure.

In the range of 1,000 to 2,000 Da, into which the products of the GOOD assay come to lie, isotopic resolution can be achieved, and the resolution m/Δm is generally higher than 1,000. Mass spectrometric detection ideally lends itself to the detection of multiple SNPs in a single preparation, as the resolution power easily allows the distinction of peaks separated by as little as 4 m/z in this mass range. The smallest mass difference between two natural nucleobases (T and A) is 9 m/z. For the positive charge-tag version the bandwidth for multiplexing can be increased by attaching charge-tag reagents of varying masses to primers of different SNPs. Furthermore, it has been shown that it is possible to measure positively and negatively charge-tagged products of the same sample. Allele-specific and positively charge-tagged DNA products of the GOOD assay were only detectable in the positive ion mode while negatively charge-tagged DNA molecules were only detectable in the negative ion mode. This unique feature of DNA detection by MALDI-MS allows, in contrast to any other MALDI-MS based method for SNP analysis, multiplexing by switching the ion mode. The same experiments performed with synthetic molecules have shown that positively charge-tagged DNA is also detectable in the negative ion mode but with bad sensitivity. It is probable that the concentration in these experiments was higher than in real GOOD assays and thus these (contradictory) results were obtained. α-Cyano-4-hydroxy-cinnamic acid methyl ester turned out to be the ideal matrix for charge-tagged DNA compounds with either one positive or one negative charge. In contrast to common DNA matrices such as 3-hydroxypicolinic-acid this matrix does not protonate DNA in solution, which is

discussed below. Therefore, by switching the ion mode in the mass spectrometer, the potential for multiplexing was doubled. However the mass spectrometer is already a good multi-channel detector using one ion mode.

The introduction of ddNTPs as supplementary substrates instead of α-S-ddNTPs for primer extension with positively charge-tagged primers leads to a cheap alternative for shifting masses in experiments where a certain degree of multiplexing should be achieved. The resulting products were zwitterions with a net charge of 0 that do not correspond to regular charge-tagged DNA molecules. These DNA molecules were detectable in the positive ion mode with a sensitivity that was empirically slightly lower than with regular, positively charge-tagged molecules containing one positive charge. Interestingly, using the negative ion mode these products were hardly detectable, which confirms recent studies where charge-neutral DNA molecules were only detectable in the positive ion mode but not in the negative ion mode.[111] The proton affinity of the phenolic group of the α-cyano-4-hydroxy-cinnamic acid methyl ester is low enough and the abundance of it during desorption high enough to protonate the zwitterionic product during the desorption/ionisation process. This coincides with the fact that positively charge-tagged DNA molecules are detectable with a sensitivity slightly higher than negatively charge-tagged DNA molecules. Negatively charge-tagged DNA in turn tends to result in higher sensitivities than zwitterionic DNA. In solution the pK_a of α-cyano-4-hydroxy-cinnamic acid methyl ester (>7) is higher than the pK_a of the phosphate group (<0). However, the highest pK_a represented in DNA is found on the nucleobases (about 4). This is still significantly below where one would expect the matrix to transfer a proton in solution. Thus one would not expect protonation of the zwitterionic species during matrix preparation. Nevertheless as it is shown in this work protonation/deprotonation after desorption plays a crucial role for charge-tagged DNA as well. A model DNA oligomer containing two positive charge-tags was only detected with a single charge, which is clear proof of deprotonation in the gas-phase. DNA can be conditioned for MALDI detection by charge-tagging but - and this is a new insight for the charge-state concept - protonation or deprotonation is apparently still an important process. Furthermore it might be proposed that the excess single charge of the product detected by MALDI-MS might not derive from the fixed charge but is rather the result of the interaction of the analytes with the matrix. The used matrix (α-cyano-4-

hydroxy-cinamic acid methyl ester) contains a phenolic group that could protonate well in the gas-phase. Interestingly corroborating this theory, proteins usually prepared with α-cyano-4-hydroxy-cinamic acid are measured with a similar performance by the α-cyano-4-hydroxy-cinamic acid methyl ester. Additionally α-cyano-4-methoxy-cinamic acid methyl ester was used. This matrix is completely unable to protonate, as there are no free protons at all. Unfortunately it did not crystallise well during MALDI preparation, which affected significantly the desorption process. Such a matrix with good crystallisation behaviour would definitely help to elucidate the ionisation process of charge-tagged molecules during MALDI. A clear proof for protonation could not be stated but the fact that deprotonation takes place makes it very probable that also protonation is of importance for the desorption of charge-tagged DNA molecules.

Common DNA preparations used in all other mentioned SNP genotyping methods using MALDI-MS result in "sweet spots".[153] Compared to this, thin layer preparation gives less spot-to-spot variation, better mass accuracy and good resolution. Using α-cyano-4-hydroxy-cinnamic acid methyl ester with thin layer preparation for modified DNA products of the GOOD assay introduced a significant improvement of the reproducibility of the sample preparation in an automated set-up, while simple liquid-handling systems can be used to deposit samples onto the MALDI target plate.

The simplified GOOD assay

The simplified GOOD assay is based on the same charge-state concept for DNA detection by MALDI-MS as the standard GOOD assay, but requires fewer steps. Due to this the simplified GOOD assay is easier to implement for high-throughput applications. Simplification is achieved by the introduction of methylphosphonates, which are already charge neutral, in place of phosphorothioates, removing one liquid-handling step from the standard GOOD assay. Charge-tagging of primers for the GOOD assay for positive ion mode detection and the alkylation are circumvented. A further advantage is that the potentially toxic methyliodide of the standard GOOD assay is avoided and no problems like over- and under-alkylation can occur. The required single negative charge

is introduced with the terminating base of the allele-specific primer extension reaction. The cost for extension primers is significantly reduced by the use of methylphosphonate instead of amino-modified phosphoramidites and phosphorothioate-containing oligonucleotides as used in the standard GOOD assay in the positive ion mode version.

Key to the simplified GOOD assay is the Tma 31 FS DNA polymerase, which also helped to avoid the shrimp alkaline phosphate digestion of the standard GOOD assays. In contrast to 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 preferably incorporates ddNTPs over dNTPs. Residual dNTPs of the PCR have no longer to be removed saving a second liquid-handling step and the cost of reagents. Tma 31 FS DNA polymerase probably will be commercially available soon.

The ability to multiplex reactions is important to reduce reagent cost per SNP genotype, analysis time and make good use of the capability of MALDI mass spectrometers. Two strategies to shift product masses of alleles are possible.

By using oligonucleotides containing for example three instead of two methylphosphonate bridges a multiplex could be generated in each base window. As for phosphorothioates similar restriction because of the integration of stereocenters were observed so that not more than three methylphoponate linkages could be employed for efficient primer extension.

Further a mixture of for example three of the four ddNTPs and one α-S-ddNTPs can be used to increase the panoply of allele masses in a multiplex. Haff and coworkers showed that a 12-fold multiplex in the primer extension and subsequent MALDI analysis is possible.[88] The simplified GOOD assay requires a single reaction vial through three reaction steps, makes use of the highly discriminative nature of DNA polymerases for the allele distinction and the high resolution of mass spectrometric detection and automatic allele-calling. Amongst the procedures using MALDI-MS, the simplified GOOD assay has the major advantage that it is executed as a short sequence of liquid-handling, thermocycling and incubation steps. Like its precursor the simplified GOOD assay does not require purification steps, which are cumbersome and expensive in large-scale applications. It provides a genotyping method with low consumption of

DNA and reagents, high fidelity of results, and complete flexibility on the choice of SNPs. The simplified GOOD assay lends itself to further miniaturisation and could potentially be executed on a microfluidic device.

Comparison of three different enzymatic allele-differentiation methods

An enzymatic procedure for the distinction of alleles increases significantly the specificity compared to simple DNA hybridisation. DNA backbone modifications required for DNA charge-tagging such as amino-modified nucleobases, phosphorothioates and methylphosphonates were compatible with enzymes such as DNA polymerases, DNA ligases and DNA flap-endonucleases. As is shown in this thesis some of them readily accept a variety of DNA modifications without loosing their efficiency and specificity.

Primer extension is the most commonly used procedure for the generation of allele-specific products because it is robust and flexible.[28,29] Additionally, a small amount of genomic DNA is required, design of experiments is easy to perform and similar reaction conditions can be used for many different primers. Among the procedures presented in this thesis primer extension using a DNA polymerase was clearly the most successful approach. The GOOD assay procedures culminated in a three-step procedure termed the simplified GOOD assay employing Tma 31 FS DNA polymerase.

DNA ligases were evaluated to replace DNA polymerases for GOOD assays. They may show similar allele-specificity as DNA polymerases but significantly less amplification efficiency.[154] Therefore in many protocols using a ligase for point mutation detection, a PCR is done before to generate enough product for detection.[155,156] Or in other approaches after the ligase reaction, rolling circle amplification takes place to generate a sufficient amount of product for detection.[157] Analogous to described protocols, a DNA ligation on a PCR amplicon with oligonucleotides containing phosphorothioates instead of usual phosphates at the adjacent linkages of the ligation site was successfully established. Thus some DNA ligases also accept this backbone

modification. But in order to get rid of the unmodified part of the used oligonucleotides, two nucleases had to be applied, such as the 5' and the 3' phosphodiesterase.

LCR of genomic DNA with oligonucleotides containing phosphorothioates and direct analysis by mass spectrometry suffers from insufficient amplification efficiency. Therefore it was conceivable to use PCR for amplification and a ligation reaction like LCR for the generation of allele-specific products. Jurinke et al. showed the analysis of LCR products obtained from a pBluescript KII phagemid DNA template and subsequent MALDI-MS analysis.[139] The analysis of mutations on the human genome with its much higher complexity was never shown. Other groups that tried to employ LCR on human genomic DNA for subsequent mass spectrometric mutation analysis failed.[158] At least 20 fmol/µl of DNA are required on the sample for MALDI-MS analysis but a LCR only generates products in the attomolar to femtomolar range. Therefore it seems to be impossible to apply LCR without additional amplification, even with charge-tag technology.

While the enzymatic digestion of LCR products from both ends was efficient according to gel analysis, no expected negatively charge-tagged products after alkylation were detected by MALDI-MS. As is shown in figure 3.28, by the combination of phosphodiesterase I and II, the double stranded synthetic DNA was digested. However, the achieved efficiency may be too low, particularly because of insufficient turnover of phosphodiesterase II. A synthetic DNA containing phosphorothioates and mimicking a LCR product with an abundance ca. 10-fold higher than what a LCR can yield had to be employed to generate enough product for MALDI detection. Already established GOOD assay protocols for alkylation and sample preparation were applied. If 5' and 3' exonucleases were used, interference might explain the inefficient digestion. Or the digestion was too aggressive (particularly with the more aggressive phosphodiesterase I) so that sometimes the phosphorothioate linkages were digested as well. The ideal solution would have been the integration of modifications for specific chemical cleavage of oligonucleotides. Currently there are no chemicals available that could be applied universally for this purpose.

The combination of the ligation reaction, exonucleolytic digestion, alkylation and subsequent analysis by MALDI-MS as is done for GOOD assays turned out to be quite

difficult. In fact the procedure would have become more complicate and more expensive than the GOOD assay using primer extension. Therefore this approach was not followed up. Nevertheless the knowledge that ligases could be applied for ligating DNA-backbone modified oligonucleotides rests an interesting insight, which might have an impact on technology development in molecular biology in the future.

The "FEN-GOOD assay" approach combines charge-tag and Invader technology. In the first step (allele-specific) amplification is required to provide sufficient template for the following allele-specific cleavage by a flap endonuclease. Therefore this procedure would be the most straightforward needing theoretically only two steps before MALDI-analysis assuming that resulting flap DNA oligomers are already charge-tagged after cleavage. The products of the shown experiments were regularly charge-tagged DNA molecules and measured by MALDI-MS as was done for the GOOD assays. Nevertheless the analysis was quite time-consuming as generally more than 100 laser shots had to be applied to yield good spectra, which is ca. 10-fold more than for GOOD assays. Additionally, spot-to-spot variation made the analysis more tedious. As this was also observed in experiments using stringent purified samples, probably traces of water-soluble components deriving from the flap-endonuclease batch affected the sample preparation quality.[104]

The proof-of-principle for the combination of Invader and charge-tag technology was shown on synthetic templates. The used flap-endonucleases accept probe molecules containing phosphorothioates at their 5'-ends. Assays with probes containing methylphosphonates were also successful. Sufficient sample purity was achieved by phase separation limiting the potential for miniaturisation.

An advantage of the flap endonuclease reaction is that in contrast to primer extension thermocycling is not required, which is discussed in the following chapter. A further advantage could be the possibility of higher multiplexing as the base composition of the flaps could be varied with a high degree of freedom. Nevertheless this was not shown by another mass spectrometric approach of the Invader assay using unmodified DNA where this should have been possible.[103,104] As already stated in the introduction the optimisation and multiplexing of the Invader assays seems to be significantly more difficult than primer extension.

A further obstacle of the FEN-GOOD approach is that the (engineered) flap endonucleases are currently only commercially available together with very expensive SNP genotyping kits developed by Third Wave Technologies (Madison, USA) in-house. It is currently not possible to compare the price of this enzyme with DNA polymerases. The employed flap endonuclease deriving from *Methanococcus janaschii* is sufficiently allele-specific for the distinction of SNPs of synthetic templates. It was also tested for generating allele-specific products on PCR products. A reaction sequence comprising a PCR, a flap-endonuclease reaction and an alkylation is feasible leading to a three-step procedure. Recently, several studies of the kinetic behaviour of flap-endonucleases particularly for Invader assays have been published.[141] According to these engineered variants of natural flap-endonucleases show better allele discrimination. The improvement of allele-specificity using a number of (engineered) flap-endonucleases requires further investigation.[142] However, the novel procedures shown here open a new field as the reaction sequence is established. A collaboration with specialists of the Invader technology providing expertise in assay design and special engineered enzymes and specialists of DNA charge-tag technology is aspired to profit from the insights presented here. Although there was a material transfer launched between the patent holder of Invader technology and flap-endonucleases, Third Wave Technology (Madison, USA), and the CNG no real collaboration took place. It is assumed that because of economic interest of the company scientific progress was impeded as no engineered, highly specific flap-endonuclease enzymes were delivered. Cloning and genetic engineering, expression and purification of the respective enzymes showing better allele-specificity would have to be established independently in the future.

Automation of the standard GOOD assay

Genotyping by the standard GOOD assay was evaluated in large-scale projects at the CNG on a number of SNPs of candidate genes for cardiovascular disease in different populations. A general success-rate of about 75-95 % was achieved. If no results were obtained in most cases the PCR reactions did not work probably because of low

integrity of template DNA. The observed genotype frequencies were in good agreement with the assumption of Hardy-Weinberg equilibrium. Furthermore, the quality of allele-output of the GOOD assays was confirmed by genotyping CEPH (Centre d'Etude Polymorphisme Humain) families and subsequent analysis of obtained heredity patterns.

It was the aim of this work to provide an efficient and affordable method for the high-throughput analysis of SNPs. The main criterion was that the procedure could be directly transferred to robots. The GOOD assays have been implemented on the BasePlate robot from the Automation Partnership (Cambridge, UK) that is a very rapid liquid-handling robot. This robot can be applied for accurate liquid-handling down to the microliter range, also in a 96-tip format and is therefore a very powerful tool for high-throughput genotyping.

The entire process of the GOOD assays requires only a microtiter plate into which via successive reaction steps appropriate reagents are dispensed. Of the starting PCR volume only a small amount is used for the MALDI sample preparation. With a PCR machine that is capable of handling smaller sample volumes reliably, reagent consumption could be reduced dramatically. From an engineering point of view a drawback of the presented GOOD assays and a lot of molecular biological procedures is the necessity for thermocycling, which makes large-scale genotyping difficult and expensive. For example, devices such as Peltier elements are applied in PCR machines.

Commercially available MALDI mass spectrometers are capable of recording 20,000 spectra per day. A single system could be used to generate 100,000 genotypes from up to 20,000 different individuals. With the next generation of MALDI mass spectrometers such as the "Autoflex" from Bruker Daltonik, on which the GOOD assay was applied in β-tests, around five times more spectra could be recorded. Furthermore, a Twister robot from Zymark (Hopkinton, MA) can be coupled to the front end of this mass spectrometer for automatic target loading. Data accumulation of 384 samples can be fully automated using AutoXecute and allele calling can then be performed online by the Genotools SNP manager software. This software is suitable for the use of Biflex/Reflex/ and Autoflex mass spectrometers from Bruker Daltonik (Bremen, Germany). Currently, the analysis of a single offline spectrum analysis in Biflex- and Reflex mass spectrometers takes 2-3 seconds. So in theory a 384-well target could be

analysed in 13-20 minutes. As the data acquisition of spectra is automated for the Autoflex mass spectrometer and coupled to the AutoXecute software, the system is able to analyse online the SNP genotypes during recording of a new spectrum.

The future challenge for high-throughput SNP genotyping by the GOOD assays will be the control of samples and the streamlining of a production line. Further developments have to be implemented, particularly a laboratory information management system (LIMS) including barcodes of DNA samples and respective allele output formats must be constructed to monitor the enormous data flux. Methods like the GOOD assays are fairly easy to control as these are facile and repetitive procedures where no samples have to be transferred and no formats have to be condensed and deconvoluted. Therefore technologies like the MassArray assay [91] that apply pooled DNAs are difficult to track when there is a failure in one or more PCRs resulting in missing points that can only be regenerated by repeating the whole experiment.

Are the GOOD assays economical enough for large-scale SNP genotyping?

Surprisingly as it may seem to discuss the economical power of a procedure in an academic script like this, but as was defined an objective of this thesis, novel ways for DNA analysis such as SNP genotyping should finally lead to better and cheaper than existing methods. Many potential users avoided starting large-scale SNP genotyping because of existing technologies costing too much, approximately around one US-dollar per SNP analysis.[54] MALDI-MS provides one of the most attractive solutions for SNP genotyping because it can be used to obtain direct and rapid measurement of DNA. A major part of the price (estimated 50 %) of technologies applying MALDI-MS such as the MassArray from Sequenom (San Diego, CA) [92] or the PinPoint assay from Perseptive (Framingham, MA) [93] stems from purification or separation methods used in the reaction sequence.[159] As the GOOD assays do not require any purification procedures the price would be consequently at least ca. 50 % lower than of mentioned procedures. All of the MALDI technologies for SNP genotyping require amplification, usually a PCR or for example in the case of the Invader assay the first round of the

"squared Invader reaction". The 2nd step generally consists in a purification procedure while during the 2nd step of the GOOD assays allele-specific products are directly generated due to Tma 31 FS DNA polymerase. The 3rd step of the other methods is generally the generation of allele-specific products while in the GOOD assays the unmodified part of the primer is digested off. The 4th step of the other methods consists of stringent purification of the DNA products for MALDI-analysis, while for the simplified GOOD assay MALDI-target preparation could be done directly. In the case of the previously developed standard GOOD assay an alkylation reaction had to be done before target preparation. The cost for the phosphodiesterase used for the removal of the unmodified part of extension primers and the reagents for the alkylation reaction is much lower than that of purification materials such as magnetic beads or reversed-phase material. The DNA modifications for the extension primers do not contribute significantly to the final price of an assay. Ca. 60 % of the cost derives from the two DNA polymerases that are required for the GOOD assay and for many other procedures for SNP genotyping as well. The material cost of singleplex GOOD assays is low compared to mentioned technologies. Commercially available SNP genotyping assays from Sequenom, generally regarded as the most advanced technology platform, currently cost at least 1 US-Dollar per SNP analysis.[92] As was stated, a multiplex factor of three for easy GOOD assay development seems to be a reasonable number thereby lowering the price per SNP analysis threefold.

To the calculated price for materials the cost of robots and mass spectrometers has to be added. The utilised BasePlate robot (The Automation Partnership, Cambridge, UK) has a price of ca. 250,000 US-Dollar and an Autoflex mass spectrometer (Bruker Daltonik, Bremen, Germany) of ca. 150,000 US-Dollar. With the current analysis time ca. 10,000 SNPs could be genotyped per day on one Autoflex mass spectrometer. Assuming consequently a daily number of 10,000 SNP analyses per day and 1,000 working days these two instruments would contribute only 4 Cents to the price per single SNP analysis. Of course other instrumentation such as thermocyclers and incubation ovens has to be considered. Nevertheless, the contribution to the final cost rests negligible. More important is the staff cost. SNP assays have to be developed by technicians before they can be analysed in high-throughput. The longer the development lasts the higher would be the contribution to the price of the procedure. Alternatively,

similar protocols of a single PCR condition could be applied for every SNP in the human genome with a success rate of ca. 70 %.[160] Two persons could do the high-throughput SNP analysis of 10,000 SNPs, one doing the lab work, while the other is undertaking the analysis and quality control of the data. Per year the cost of these two persons would be ca. 100,000 US-Dollar thereby contributing ca. 2 Cents to the final cost of the GOOD assays. However, all these calculations rest sketchy. At the moment it is too early for calculating a definite final price of the assay as the production line of the GOOD assay is still under development. Nevertheless the mentioned numbers could give an idea of the magnitude of cost of the GOOD assays compared to other technologies for SNP genotyping demonstrating its highly promising potential.