Application of the C4'-Alkylated Deoxyribose Primer System (CAPS) in Allele-Specific Real-Time PCR for Increased Selectivity in Discrimination of Single Nucleotide Sequence Variants

Björn Tews^{1,a}, Jochen Wilhelm^{1,b}, Daniel Summerer², Michael Strerath², Andreas Marx², Peter Friedhoff¹, Alfred Pingoud¹ and Meinhard Hahn^{1,a,*}

¹ Institut für Biochemie, FB 08, Justus-Liebig-Universität Giessen, Heinrich-Buff-Ring 58, D-35392 Giessen, Germany

² Kekulé-Institut für Organische Chemie und Biochemie, Universität Bonn, Gerhard-Domagk-Strasse 1, D-53121 Bonn, Germany

*Corresponding author

This study describes a quantitative real-time PCR-based approach for discrimination of single nucleotide sequence variants, called CAPS (C4' alkylated primer system). To increase the discrimination potential of DNA polymerases against competing sequence variants of single nucleotides, 3'-terminally modified primers were designed carrying a methyl residue bound to the C4' of the thymidylate deoxyribose. In a model sequence system positional dependencies of modified thymidylate (at -1, -2, -3) were tested for their influence on discrimination. Highest discrimination factors were obtained with the modification at the ultimate 3'-position. In a comparison between *Taq* and *Pwo* DNA polymerases, substantial better results were obtained by *Taq* DNA polymerase. In contrast to conventional PCR methods for discrimination of sequence variants, achieving a maximum discrimination potential of about 20, CAPS is capable of obtaining sequence-specific amplifications of a desired target among discriminated templates with a dynamic range of 1:100. Therefore, CAPS is a method able to quantitatively discriminate two sequence variants only differing in a single base (e.g., SNP alleles or point mutations). The range of applications of this easy to perform, fast and reliable technique reaches from medical diagnostics, transplantation medicine, molecular and cell biology to human genetics. Target-

ing of SNPs assures a universal exertion of this method, since these markers are gender-independent, highly abundant and ubiquitous.

Key words: Discrimination / Modification / Real-time PCR / Sequence variants.

Introduction

Today there is an increasing demand for quantitative methods that are able to sensitively discriminate different DNA sequence variants presented together in a mixed (i.e., chimeric) sample and which should feature a very high dynamic range, meaning that a single copy of a certain sequence **B** should be detectable among as many copies as possible deriving from a different sequence **A**. These methods, thereafter described by the term 'chimerism analysis', are required for answering questions raised in molecular and cell biology, human genetics, oncology, haematology, transplantation medicine and population biology.

Established PCR-based methods used for chimerism analysis (e.g., Zhang *et al.*, 1999; Formankova *et al.*, 2000; Jone *et al.*, 2000; Zetterquist *et al.*, 2000; Fehse *et al.*, 2001; Sasanakul *et al.*, 2001) are able to reach dynamic ranges of about 1:20, i.e., one copy of variant **B** can be quantified among 20 copies of variant **A**. Recently, the combination of real-time PCR techniques with the amplification refractory mutation system, called ARMS (Newton *et al.*, 1989; Wu *et al.*, 1989), was described for chimerism analysis (Whitcombe *et al.*, 1998). This combination has permitted a much more reliable interpretation of experimental results because of the possibility of real-time measurement of PCR-product accumulation. The concept of ARMS is to benefit from the loss of PCR-amplification in the presence of a primer-template mismatch. Using real-time PCR with ARMS for the analysis of a specific polymorphism, defined by the presence/absence of short insertion/deletion sequences, it was possible to achieve a dynamic range of 1:3000 (Wilhelm *et al.*, 2002). Unfortunately, this kind of polymorphisms is infrequent in mammalian genomes.

On the genomic level, single nucleotide polymorphisms (SNPs) represent the most simple form of possible DNA sequence differences between two individuals. SNP features are: (i) high genomic density: one SNP per 1000 bp (Sachidanandam *et al.*, 2001; Venter *et al.*, 2001), (ii) no stringent correlation with diseases in the overwhelming majority of

Present addresses:

^a Abteilung Molekulare Genetik (B060), Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, D-69120 Heidelberg, Germany

^b Institut für Pathologie, FB 11, Justus-Liebig-Universität Giessen, Langhansstr.10, D-35385 Giessen, Germany

cases, (iii) gender-independence, (iv) lack of preferential allele amplification or allelic drop-out effects and (v) the number of known SNPs continuously increases (currently more than 1.4 million SNPs are known; Sachidanandam *et al.*, 2001). Therefore, allelic SNP variants should serve as suitable targets for ARMS in real-time PCR.

Recently, we reported that the specificity of a DNA polymerase is significantly increased with primers containing thymidine having its 4'-hydrogen substituted by small alkyl groups. They are ideally suited since potential effects on hydrogen-bonding patterns and conformations of the nucleotides are minimized (Summerer and Marx, 2001, 2002). These findings were carried forward to the development of a PCR system with increased sensitivity (Strerath and Marx, 2002). Based on these investigations, we have now developed a real-time PCR model

system for ARMS using primers carrying such a modified base in different positions. We used this model to systematically elucidate which positions and mismatch combinations lead to the best discrimination results between two mixed sequence variants only differing in a single base. We were able to quantitatively discriminate sequences differing only in a single base with a dynamic range of at least two orders of magnitude.

Results

Construction and Synthesis of Model Targets

To systematically analyze the properties of CAPS, model template DNAs containing defined SNPs within identical

Table 1 Sequences of the Artificial Target Constructs Used in the Real-Time PCR Experiments.

5'-GTGGTGCGAA ATTTCTGACA GAC <u>1234</u> CAG ACGCACAC AACTGTGTTT ACTAGCAACC TCAACAGAC ACCATGGTGC ACCTGACTCC TGAGGAGAAG TCTGCCGTTA CTGCCCTGTG GGGCAAGGTG AACGTGGATG AAGTTG-3'					
Template	1	2	3	4	Length (bp)
1.1	A	T	G	A	146
1.2	A	C	G	A	146
1.3	A	T	A	A	146
1.4	A	G	A	A	146
1.5	A	G	G	A	146
1.6	A	C	G	C	146
1.7	A	T	A	C	146
1.8	C	T	G	A	146
1.9	C	T	A	A	146
1.10	A	A	G	A	146

In the different target constructs 1.2 to 1.10, positions 1 to 4 are varying compared to the reference template 1.1. Exemplarily, the primer binding site of the series 27X (see Table 2) is shown by underlined nucleotides, while the binding site of the primer PC04 is indicated by a dashed line. Primers of the series 26X (alternatively 25X) miss one (two) nucleotides at the 3'-terminus (see Table 2). The dotted line indicates the binding site (PC03 sequence) of the template generating primers (PC03-1.1 to PC03-1.10) on human genomic DNA. These template-generating primer variants contain the solid plus dotted underlined sequence (56 nt).

Table 2 Sequences of the Discriminating Primers Used for the Amplification of the Target Constructs.

Name	Sequence 5'-3'	Length (bp)
251	5'-GTG GTG CGA AAT TTC TGA CAG ACA T-3'	25
252	5'-GTG GTG CGA AAT TTC TGA CAG ACA T*-3'	25
253	5'-GTG GTG CGA AAT TTC TGA CAG ACA T**-3'	25
261	5'-GTG GTG CGA AAT TTC TGA CAG ACA TG-3'	26
262	5'-GTG GTG CGA AAT TTC TGA CAG ACA T*G-3'	26
263	5'-GTG GTG CGA AAT TTC TGA CAG ACA T**G-3'	26
271	5'-GTG GTG CGA AAT TTC TGA CAG ACA TGA-3'	27
272	5'-GTG GTG CGA AAT TTC TGA CAG ACA T*GA-3'	27
273	5'-GTG GTG CGA AAT TTC TGA CAG ACA T**GA-3'	27

Thymidine residues carrying a methyl group at C4' of deoxyribose are labelled by one asterisk, those carrying an ethyl group by two.

sequence context had to be generated. In total, ten different template DNAs (1.1–1.10, see Table 1) were generated by conventional PCR, using human genomic DNA as template. The primer sequences (see Tables 1 and 2)

Initial PCR Cycle

Fig. 1 Generation of the Artificial Targets.

The scheme shows the upstream primer (template primer), which binds with the sequence of **PC03** to the human gene *HBB*. This primer exists in 10 variants (**PC03–1.1–1.10**, see Table 1). The protruding artificial sequence (striped bar) contains the polymorphic bases **1**, **2**, **3** and **4**; see Table 1). The downstream primer is always **PC04**.

were based on the primers **PC03/PC04** (Wittwer *et al.*, 1989), which amplify a part of the human β -globin gene (Figure 1). The different variants of the polymorphic site were attached 5' to the forward primer **PC03** (see Table 1). As a result, the variants of the 146 bp long model template DNA can be amplified with one of the ten non-discriminating template generating primers (**PC03–1.1** to **PC03–1.10**) and the reverse primer **PC04** (Figure 1). The sequences of the different artificial target constructs are shown in Table 1. All ten targets were pure as indicated by their UV absorbance spectra, the presence of only one single band on a 15% polyacrylamide gel as well as a clear single peak in the melting curve analysis (data not shown).

Discrimination of Different 3' Terminal Primer/Target Mismatches

To get a clue of the potential of SNP discrimination in real-time PCR due to the modified primer thymidine, several primer/template mismatch situations (see Tables 2 and 3) were tested. In a first series of experiments, both,

Table 3 Tested Primer/Template Situations.

A				
25X	5'–GTG GTG CGA AAT TTC TGA CAG AC	A T	–3'	
26X	5'–GTG GTG CGA AAT TTC TGA CAG AC	A TG	–3'	
27X	5'–GTG GTG CGA AAT TTC TGA CAG AC	A TGA	–3'	
		1 2 3 4		
1.1	3'–CAC CAC GCT TTA AAG ACT GTC TG	T ACT	GTC TGC GTG TGT...–5'	
1.2	3'–CAC CAC GCT TTA AAG ACT GTC TG	T GCT	GTC TGC GTG TGT...–5'	
1.3	3'–CAC CAC GCT TTA AAG ACT GTC TG	T ATT	GTC TGC GTG TGT...–5'	
1.4	3'–CAC CAC GCT TTA AAG ACT GTC TG	T CTT	GTC TGC GTG TGT...–5'	
1.5	3'–CAC CAC GCT TTA AAG ACT GTC TG	T CCT	GTC TGC GTG TGT...–5'	
1.6	3'–CAC CAC GCT TTA AAG ACT GTC TG	T GCG	GTC TGC GTG TGT...–5'	
1.7	3'–CAC CAC GCT TTA AAG ACT GTC TG	T ATG	GTC TGC GTG TGT...–5'	
1.8	3'–CAC CAC GCT TTA AAG ACT GTC TG	G ACT	GTC TGC GTG TGT...–5'	
1.9	3'–CAC CAC GCT TTA AAG ACT GTC TG	G ATT	GTC TGC GTG TGT...–5'	
1.10	3'–CAC CAC GCT TTA AAG ACT GTC TG	T TCT	GTC TGC GTG TGT...–5'	
B				
	25X	26X	27X	
	1 2	3	4	
1.1	M M	M	M	
1.2	M ⊗	M	M	
1.3	M M	⊗	M	
1.4	M ⊗	⊗	M	
1.5	M ⊗	M	M	
1.6	M ⊗	M	⊗	
1.7	M M	⊗	⊗	
1.8	⊗ M	M	M	
1.9	⊗ M	⊗	M	
1.10	M ⊗	M	M	

Part (A): The upper block shows the primer sequences (**25X**, **26X**, **27X**; plus strand), the lower block the minus strand of corresponding templates (**1.1–1.10**). Part (B): Schematic presentation of matching (M) and mismatching (⊗) base pairs.

In the primer names, the abbreviation 'X' stands for 1, 2 or 3 (e.g., **251** unmodified primer, **252** C4'-methylated T, **253** C4'-ethylated T); **T** represents the unmodified, methylated or ethylated thymidine, respectively. The box displays the relevant varying sequence positions, indicated by **1**, **2**, **3** or **4**, as shown in Figure 1 and Table 1.

Fig. 2 Comparison of the Discrimination Potential for Match and All Three Possible Mismatch Combinations of the Unmodified Primer **251** versus the Modified Primer **252**.

(A) Fluorescence signal curves (channel 1, RFU: relative fluorescence units) of real-time PCR experiments using the unmodified primer **251**. (B) The same experiments as shown in (A), but using the 3'-terminally modified primer **252** (T*: methylation at C4' of deoxyribose of thymidylate) instead of **251**. The insets show the structural formula of either the unmodified (A) or modified (B) thymidylate. The 3'-terminal primer/template base situation is also shown. All reactions were performed under standard conditions.

the terminally unmodified and modified primers **251** and **252** were used to amplify four different targets (**1.1**, **1.2**, **1.5**, and **1.10**). All targets had the same sequences except at position 2. The amplification curves are shown in Figure 2. All four targets could be well amplified with the unmodified primer **251**: the amplification curves have C_T values (the C_T value is the cycle at which the amplification curve of a reaction reaches a fluorescence intensity above background) between 10 and 15 (Figure 2A). These results were confirmed in 10 repetitions. The amplification with the methyl C4' modified primer **252** yielded an amplification curve with a C_T value of 12 for the matching target DNA (Figure 2B), also verified by 10-fold

Fig. 3 Fluorescence Signal Curves of Real-Time PCR Experiments with Low Template Concentrations.

The modified primer **252** was used either with matching target **1.1** or mismatching target **1.2**. The reaction was performed under standard conditions except for the use of low copy numbers of template. The signal is measured in channel 1 (RFU: relative fluorescence units).

Fig. 4 Discrimination Potential of *Taq* DNA Polymerase versus *Pwo* DNA Polymerase.

Comparison of the mismatch discrimination potential of *Taq* (white bars) and *Pwo* DNA polymerase (dark bars) using different primer series and mismatch target **1.2**. Numbers above the columns indicate the efficiencies (ϵ) of the amplification reactions. The relevant sequences of the different primer/template pairs, containing the modified thymidylate residue (bold italics), are shown. Reactions were performed under standard conditions. Primer series **27X** shows no detectable amplification with *Pwo* DNA polymerase. The data are the average of triplicate experiments.

repetition. The amplification curves of the other targets have C_T values of more than 25. The increasing order of C_T values is T/G < T/C < T/T (primer:template). Non-competitive experiments revealed discrimination factors (for definition of terms, see Materials and Methods) of $D = 8000$ (T/T mismatch), $D = 7000$ (T/C mismatch), $D = 800$ (T/G mismatch) (data not shown). These results could be confirmed even when using low template copy numbers (10^4 and 10 copies) which is exemplarily shown for primer **252** with the templates **1.1** and **1.2** in Figure 3.

In case of amplification reactions with modified primers carrying an ethyl residue at C4' of the deoxyribose of the terminal thymidylate (**2X3**, i.e., **253**, **263** or **273**, respectively), discriminations could not be evaluated due to very low amplification efficiencies throughout, no matter which template was used (data not shown).

Similar experiments with additional mismatch situations (C4'-methylated primers **2X2** with templates **1.3**, **1.4**, **1.6**, **1.7**, **1.8** and **1.9**) resulted in far worse discrimination potentials (data not shown) compared to those of the 5'-terminally methylated primer **252** with templates **1.2**, **1.5** and **1.10** (i.e., 5'-terminal single mismatch situation, see Table 3).

Comparison of the Discrimination Potential of *Taq* and *Pwo* DNA Polymerase

To elucidate whether the discrimination can be further im-

proved using a thermostable DNA polymerase with 3'→5' proofreading activity, a quantitative discrimination analysis was carried out. Amplification reactions were performed using *Taq* and *Pwo* DNA polymerase, the primers **25X** (i.e., **251**, **252** or **253**, resp.), **26X**, and **27X**, and the template **1.2** as mismatch target. Here, in all cases the mismatch situation is T/G, which has been shown to be the least discriminated mismatch in this assay (see Figure 2). The calculated discrimination factors and corresponding amplification efficiencies for all combinations are shown in Figure 4. When using primer **25X** (3'-terminal mismatch), the discrimination by *Taq* DNA polymerase was 4- to 100-fold higher than by *Pwo*. Using primers **26X** (penultimate mismatch) and *Pwo*, it was vice versa with a 3- to 40-fold higher discrimination, but very poor amplification efficiencies were observed. The observed amplification efficiencies of *Taq* DNA polymerase were always higher (1.60–1.84) than those of *Pwo* (1.22–1.41). The *Pwo* DNA polymerase could not even amplify the template **1.1** with the primers **27X** (pre-penultimate mismatch situation) under our assay conditions. Due to these poor amplification efficiencies, *Pwo* DNA polymerase is not suited for quantitative assays.

Titration Experiments

The limit for the quantitative analysis of a low abundant sequence variant in the DNA sample can be estimated

Fig. 5 Titration Experiments.

C_T -values (as defined in Materials and Methods) for different concentrations of perfect matching template (**1.1**) in a background of a constant amount of mismatching template (10^7 copies of **1.2**, **1.5** or **1.10**). Primer pair **252/PC04** was used for all amplification reactions under standard conditions. Panel (A) shows the perfect matching situation (T/A) in absence of mismatching templates. Panel (B) shows the T/G mismatch situation (background **1.2**). Panel (C) presents the T/C mismatch situation (background **1.5**) and panel (D) the T/T mismatch situation (background **1.10**). Values are the average of triplicate experiments.

from titration series. In these titration experiments, low amounts of matching target ('amplified DNA') are mixed and, thereafter, amplified in competition to a high amount of mismatch target ('background DNA'). The obtained C_T values for the amplification curves of a titration series should show a linear correlation to the logarithm of the matching target concentration. This linear correlation will be lost for concentrations of amplified DNA below the discrimination limit.

The discrimination limits were determined for all possible 3'-terminal mismatches (C4'-methylated primer **252** and targets **1.2**, **1.5**, and **1.10** as background DNA, see Table 3) in triplicates using *Taq* DNA polymerase (Figure 5). As a control, the titration experiment was carried out using only the amplified DNA (matching target **1.1**) in different concentrations (Figure 5 A). A linear correlation between C_T values and log copy numbers was observed in all cases. The range of linearity of the C_T values versus log matching template concentration was 1.5 orders of magnitude in case of the background DNA **1.2** leading to the weak discriminated T/G-mismatch. The amplified DNA could be quantified in the presence of a more than 30-fold excess of background DNA. In case of other background DNAs (**1.5** and **1.10**), the linearity was at least two orders of magnitude. Initial concentrations of the amplified DNA higher than 10^7 copies resulted in amplification curves with C_T values below 10. Due to the limited amount of data points in the (pre-)exponential phase, these curves were not used for a quantitative analysis.

Melting Curve Analysis

In real-time PCR experiments using SYBR-Green I as ds-DNA-specific dye, the obtained PCR products can be characterized by a melting curve analysis. Products of the same length but different sequences can be well distinguished by melting curve analysis. Examples of melting curves are shown in Figure 6. The correct 146 bp product of the matching target **1.1** amplified with any forward primer **2X2** (and **PC04** as reverse primer) melts at 88.3°C. All amplifications of mismatch target DNAs with **2X2/PC04** produce a final PCR product of the same length and melting behavior (see melting peak at 88.3°C for template **1.10** in Figure 6). Only the amplification of template **1.10**, i.e., the case of the best discriminated mismatch T/T, produces minor amounts of a shorter unspecific product melting at 84.9°C (Figure 6).

Discussion

Comparison of Established Methods and CAPS

Sensitive quantitative methods for discrimination of two different DNA sequence variants are indispensable tools for many questions raised in biological research and medical diagnostics. At the present time, methods used for chimerism analysis are usually based on the analysis of DNA via PCR. Therefore, polymorphic markers differ-

Fig. 6 Melting Curves.

The melting temperature T_m is the temperature at which 50% of the strands of a double-stranded PCR product are dissociated. It can be derived from a diagram in which $-dF/dT$ (i.e., negative change of fluorescence with respect to temperature) is plotted versus the temperature T , in which the characteristic melting process of a product is shown as a distinct peak with its maximum indicating the T_m value. Here, the melting curves of PCR products **252/1.1** and **252/1.10** (strongest mismatch discrimination; T/T mismatch) are shown. Amplification was carried out under standard conditions using 10^7 copies of initial target. All PCR-products show the same melting temperature at 88.3°C. This indicates the presence of identical PCR products, which was additionally verified by polyacrylamide gel electrophoresis (data not shown). Only in case of combination **252/1.10** small amounts of primer-dimers ($T_m = 84.9^\circ\text{C}$) were detected. Control: identical reaction mixture in the absence of template molecules.

ing in size, prevalently STRs, serve as targets (Nollet *et al.*, 2001; Thiede *et al.*, 2001). Nevertheless, two main problems could not be solved so far. Firstly, the methods are not sensitive enough. Conventional allele-specific PCR-methods targeting SNPs can only achieve a maximal discrimination factor of about 20 (Kleeberger *et al.*, 2000; Nollet *et al.*, 2001), which is due to the appearance of slippage products (Mital *et al.*, 1999) and allelic-drop-out effects. Secondly, methods offering a high discrimination potential, such as real-time quantitative Y chromosome-specific PCR (QYCS-PCR; Zhang *et al.*, 1999; Fehse *et al.*, 2001), only have a limited range of application (e.g., male-in-female sex mismatched chimerisms). CAPS, our newly developed method, is able to discriminate between closely related sequences without any limitation concerning application: one copy of allele **B** can be quantified among 100 copies of allele **A**.

Unmodified versus Modified Primers

By introduction of C4'-alkyl substituents in thymidines at the appropriate positions steric strain at the 3'-end of the primer is increased. Thus, we speculate that the flexibili-

ty at the 3'-primer end is decreased through these modifications, resulting in less DNA polymerase propensity for extension of geometrically altered mismatched primer template complexes (Summerer and Marx, 2001; Kool, 2002). While this might hold true for the C4'-methyl modifications, C4'-ethyl modifications cause exceeding interference, rendering the primer incompetent for efficient extension.

Comparing the amplification reactions of the perfect matching template with unmodified and modified (methylated) primers, primer extension occurs in both cases. DNA polymerases synthesize products on every given template in the presence of an unmodified primer, no matter which primer-template mismatch situation is given. In case of a methylated primer, the accumulation of product from the non-matching template is negligible. Significant amounts of product are only generated on perfect matching template, which implies that the DNA polymerase discriminates between different templates, depending on the presence of a mismatch (Figure 2). The usage of C4'-ethylated primers leads to insufficient amplification results due to very low PCR amplification efficiencies.

Discrimination Potential: Dependence on Mismatch-Position and Type of DNA Polymerase

As discussed above, the best discrimination occurs when using a 3'-terminal C4'-methylated primer. Furthermore, the position of the modified nucleotide has an essential influence on the discrimination-potential of DNA-polymerases lacking 3'→5' proof-reading activity (*Taq* DNA polymerase) or not (*Pwo* DNA polymerase). Obviously, mismatches of the terminal primer-nucleotide have the most powerful effect on discrimination while using *Taq* DNA polymerase. If the terminal mismatch is T/T, the best discrimination values for *Taq* DNA polymerase ($D = 8000$) can be obtained, followed by $D = 7000$ for a terminal T/C mismatch. T/G can be considered as the mismatch with the lowest discrimination potential ($D = 800$). For this 3'-terminal modified primer (i.e., **251** and **252**; see left third of Figure 4), *Pwo* DNA polymerase does not lead to better results in discrimination, but rather shows weaker discrimination potentials and throughout poorer amplification efficiencies than *Taq* DNA polymerase in combination with terminal mismatch variants.

Titration Experiments

Titration experiments were performed simulating conditions where CAPS could be applied. The results indicate that even with the mismatch combination having shown the weakest discrimination potential in former non-titration experiments (T/G; see Figure 2), the assay was capable to quantify the desired template in the presence of more than 30 times more background DNA. *Taq* DNA polymerase is able to discriminate between strong mismatch combinations (T/T, T/C) in titration experiments of at least 2 orders of magnitude. Remarkably, we found no

further FRET-based real-time PCR methods for quantitative SNP-based chimerism analysis in the literature; we only found one publication (Solinas *et al.*, 2001) for classical allelotyping with a discrimination potential of approximately five.

Conclusions

We have developed CAPS, a system for detection and quantification of SNPs by allele-specific real-time PCR. Comparing CAPS with data from literature of other set-ups, it provides the highest dynamic range for quantitative analysis of a target allele in the presence of excess amount of non-target alleles. Furthermore, the application is not restricted by the lack of appropriate markers due to desiderative appearance, because SNPs occur with the highest genomic density of all known markers (Sachidanandam *et al.*, 2001; Venter *et al.*, 2001). A set of only 10 biallelic markers with allele frequencies of 0.5 would guarantee to have at least one informative marker in 99% of the cases or individuals tested (Wilhelm *et al.*, 2002). In addition, these markers are gender-independent. Especially if a rapid-cycle real-time PCR machine like the LightCycler is used for analysis, results can be obtained in a very short time.

We believe that CAPS is a method that can be used as a routine diagnostic tool for monitoring of bone-marrow and solid organ transplantations (Lagaaij *et al.*, 2001). Furthermore, questions raised in human genetics or molecular and cell biology can also be easily answered with CAPS.

Materials and Methods

If not further specified, all conventional methods were done following the instructions of Current Protocols in Molecular Biology (Ausubel *et al.*, 2002) and Molecular Cloning (Sambrook *et al.*, 2001).

Template DNA Preparation

Human genomic RNA-free DNA was isolated from lymphocytes of fresh EDTA-treated blood using the QIAamp DNA Blood reagent set (Qiagen, Hilden, Germany). PCR products were purified with the QIAquick PCR Purification reagent set (Qiagen). DNA quality and concentration were determined by recording ultraviolet absorbance spectra between 220 and 320 nm using a Hitachi U-3000 Spectrophotometer (Hitachi, London, UK) and the relation $1 A^{260} = 50 \mu\text{g/ml dsDNA}$.

PCR

All non-modified primers were obtained from Applied Biosystems, Weiterstadt, Germany. Modified primers (Table 2) were synthesized as described (Summerer and Marx, 2001, 2002). All oligonucleotide sequences were checked with the program Oligo 5.0 (National Biosciences Inc., Plymouth, USA) for absence of false priming sites, formation of primer dimers and secondary structures.

Standard PCR for generating the target DNA variants (Figure 1

and Table 1) was carried out in a GeneAmp 2400 PCR System (Applied Biosystems) using a 30 µl volume of the reaction mixture containing 0.5 g/l BSA (Roche Diagnostics, Mannheim, Germany), 3.0 mM MgCl₂, 0.5 µM each primer, 0.2 mM each of dATP, dCTP, dGTP and dTTP (Promega, Mannheim, Germany), 0.5 U Taq DNA-polymerase (Eppendorf, Hamburg, Germany) and 20–30 ng human genomic DNA in 1× PCR buffer (Roche Diagnostics) which consisted of 10 mM Tris-HCl, 1.5 mM MgCl₂, 50 mM KCl, pH 8.3 (20°C).

The amplification protocol consists of an initial denaturation step at 95°C for 30 s, followed by 11 amplification cycles at 95°C for 30 s, 60°C (–0.5°C/cycle until 55°C at cycle 11) for 20 s and 72°C for 20 s and 35 amplification cycles at 95°C for 30 s, 55°C for 30 s and 72°C for 60 s with a final extension step of 72°C for 2 min. Purified DNA was checked by UV absorbance as above.

The LightCycler system (Roche Diagnostics) was used for real-time PCR amplification. All reactions were carried out in a total volume of 10 µl per capillary. Each reaction mixture contained 0.5 g/l BSA (Roche Diagnostics), 5.9 mM MgCl₂ (Promega), 0.5 µM of each primer, 0.2 mM each of dATP, dCTP, dGTP and dTTP (Promega), 0.5 U Taq DNA polymerase (Eppendorf), a 1:10⁵ dilution of SYBR-Green I stock solution (Roche Diagnostics) and 1 µl of DNA template (~10⁷ copies of one variant of the targets 1.1–1.10, see Table 1) in 1× PCR buffer (see above).

The amplification protocol consists of an initial denaturation step at 95°C for 30 s, followed by 50 amplification cycles at 95°C for 0 s, 58°C for 2 s and 72°C for 10 s (constant temperature ramp of 20°C/s). Measurements were taken at the end of the extension phase at 72°C. After PCR, each amplification reaction was checked for the presence of non-specific products by melting curve analysis during a linear temperature ramp from 72°C to 95°C with 0.1°C/s and native polyacrylamide gel electrophoresis.

Non-Competitive Amplification

To examine which mismatch combination yields to the best discrimination, we tested which C_T-values are obtained when the different mismatching templates are amplified with Taq DNA-polymerase. The perfect-matching template (1.1) was used in two different concentrations, 10⁷ and 10⁴ copies, for the calculation of amplification efficiency of each run (see chapter 'Data Evaluation, Calculation of Efficiency' below). Each mismatching template (1.2–1.10) was tested with a concentration of 10⁷ copies per reaction. To elucidate whether the system works also with low copy numbers of the templates, the reaction was also performed using 10⁴ and 10 copies of each template.

Competitive Amplification (Titration)

To simulate the situation in chimeric samples, the perfect-matching template (1.1) was serially diluted 1:2 from 10⁷ to 6×10² copies and mixed with a constant amount (10⁷ copies) of mismatching template (1.2, 1.5 or 1.10).

Data Evaluation

The raw data were collected with the LightCycler Run software, Ver. 3.0 and evaluated with SoFAR, Ver. 1.0, which is a software for the analysis of real-time PCR data (Wilhelm *et al.*, 2003).

Calculation of Efficiency The amplification efficiency ε ($1 \leq \varepsilon \leq 2$) is calculated from the ΔC_T value of two samples **a** and **b** with known but different target concentrations (here: 10⁷ and 10⁴

copies). Assuming an exponential product accumulation, the C_T values of the samples can be derived from $N_{C_T} = N_0 \cdot \varepsilon^{C_T}$, where N₀ is the initial target concentration and N_{C_T} is the concentration of amplified product after C_T cycles of amplification. Because the threshold value used for the determination of the C_T values is the same for both samples, the product concentrations at the C_T values of both samples **a** and **b** are equal: $N_{C_{Ta}} = N_{C_{Tb}} = N_{C_T}$.

Therefore, the ratio of the initial concentrations of the samples **a** and **b** can be written as:

$$\frac{N_{0,a}}{N_{0,b}} = \frac{N_{C_T} \cdot \varepsilon^{-C_{Ta}}}{N_{C_T} \cdot \varepsilon^{-C_{Tb}}} = \varepsilon^{C_{Tb} - C_{Ta}} = \varepsilon^{\Delta C_T} \quad (1)$$

Equation (1) can be solved for ε :

$$\varepsilon = \left(\frac{N_{0,a}}{N_{0,b}} \right)^{1/\Delta C_T} \quad (2)$$

For the different template concentrations (10⁷ and 10⁴ copies) used in this work, equation (2) simplifies to:

$$\varepsilon = 10^{3/\Delta C_T} \quad (3)$$

Calculation of Discrimination When using a mismatching template **A** in real-time PCR, it will probably happen, that the polymerase will erroneously elongate a hybridized mismatch primer, specific for template **B**, despite the given mismatch situation. This leads to an elongation product with the sequence of template **B**, which is a perfect matching template in the following cycles. Therefore, due to mispriming, PCR product of type **B** is generated and amplified and leads to an increase in fluorescence signal.

Hence, the discrimination factor *D* is defined as the ratio between the initial concentration of mismatching template **A**, N_{0,A}, and the amount of erroneously quantified template **B**, corresponding to a fictive initial concentration of N_{0,B} (which can be derived from the measured C_T-value):

$$D = \frac{N_{0,A}}{N_{0,B}} \quad (4)$$

Knowing the amplification efficiency ε , which is assumed being identical for both, the amplification of the matching and mismatching template, the discrimination factor can be derived from the C_T values determined from a mismatching template **A** compared to a matching template **B** similar to equation (1):

$$D = \frac{N_{0,A}}{N_{0,B}} = \frac{N_{C_T} \cdot \varepsilon^{-C_{TA}}}{N_{C_T} \cdot \varepsilon^{-C_{TB}}} = \varepsilon^{C_{TB} - C_{TA}} = \varepsilon^{\Delta C_T} \quad (5)$$

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft (Graduiertenkolleg *Molekulare Biologie und Pharmakologie* to J.W.), and grant MA 2288/3-1 to A.M., the Volkswagen Foundation to A.M. and the Fonds der Chemischen Industrie (A.P.).

References

- Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (2002). *Current Protocols in Molecular Biology* (New York, USA: John Wiley & Sons Inc.).
- Fehse, B., Chukhlovina, A., Kuhlcke, K., Marinets, O., Vorwig, O.,

- Renges, H., Kruger, W., Zabelina, T., Dudina, O., Finckenstein, F.G. *et al.* (2001). Real-time quantitative Y chromosome-specific PCR (QYCS-PCR) for monitoring hematopoietic chimerism after sex-mismatched allogeneic stem cell transplantation. *J. Hematother. Stem Cell Res.* 10, 419–425.
- Formankova, R., Honzatkova, L., Moravcova, J., Sieglova, Z., Dvorakova, R., Nadvornikova, S., Vitek, A., Lukasova, M., Stary, J., and Brdicka, R. (2000). Prediction and reversion of post-transplant relapse in patients with chronic myeloid leukemia using mixed chimerism and residual disease detection and adoptive immunotherapy. *Leuk. Res.* 24, 339–347.
- Jone, C.M., Akel, N., and Killeen, A.A. (2000). Evaluation of chimerism in DNA samples by PCR amplification of D1S80 with detection by capillary electrophoresis. *Mol. Diagn.* 5, 101–105.
- Kleeberger, W., Rothamel, T., Glockner, S., Lehmann, U., and Kreipe, H. (2000). Laser-assisted microdissection and short tandem repeat PCR for the investigation of graft chimerism after solid organ transplantation. *Pathobiology* 68, 196–201.
- Kool, J. (2002). Active site tightness and substrate fit in DNA replication. *Annu. Rev. Biochem.* 71, 191–219.
- Lagaaij, E.L., Cramer-Krijnenburg, G.F., van Kemenade, F.J., van Es, L.A., Bruijn, J.A., and van Krieken, J.H. (2001). Endothelial cell chimerism after renal transplantation and vascular rejection. *Lancet* 357, 33–37.
- Mital, M.K., Curtis, A., Spencer, V., Barge, D., and Skinner, R. (1999). Delayed engraftment and mixed chimerism after HLA-identical sibling donor BMT in Fanconi anaemia. *Bone Marrow Transplant.* 24, 201–204.
- Newton, C.R., Heptinstall, L.E., Summers, C., Super, M., Schwarz, M., Anwar, R., Graham, A., Smith, J.C., and Markham, A.F. (1989). Amplification refractory mutation system for prenatal diagnosis and carrier assessment in cystic fibrosis. *Lancet* 2, 1481–1483.
- Nollet, F., Billiet, J., Selleslag, D., and Criel, A. (2001). Standardisation of multiplex fluorescent short tandem repeat analysis for chimerism testing. *Bone Marrow Transplant.* 28, 511–518.
- Sachidanandam, R., Weissman, D., Schmidt, S.C., Kakol, J.M., Stein, L.D., Marth, G., Sherry, S., Mullikin, J.C., Mortimore, B.J., Willey, D.L. *et al.* (2001). A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 409, 928–933.
- Sambrook, J., and Russel, D.W. (2001). *Molecular Cloning. A Laboratory Manual*, 3rd Edition (Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory Press).
- Sasanakul, W., Hongeng, S., Chuansumrit, A., Chaiyaratana, W., Pakakasama, S., and Hathirat, P. (2001). The usefulness of X-linked polymorphic loci as gene markers to track X allele and chimerism in a post-allogeneic peripheral blood stem cell transplant patient with Wiskott-Aldrich syndrome. *J. Med. Assoc. Thai.* 84, 379–384.
- Solinas, A., Brown, L.J., McKeen, C., Mellor, J.M., Nicol, J.T.G., Thelwell, N. and Brown, T. (2001). Duplex Scorpion primers in SNP analysis and FRET applications. *Nucleic Acids Res.* 29, e96.
- Strerath, M., and Marx, A. (2002). Tuning PCR specificity by chemically modified primer probes. *Angew. Chem. Int. Ed.* 41, 4766–4769.
- Summerer, D., and Marx, A. (2001). DNA polymerase selectivity: sugar interactions monitored with high-fidelity nucleotides. *Angew. Chem. Int. Ed.* 40, 3693–3695.
- Summerer, D., and Marx, A. (2002). Differential minor groove interactions between DNA-polymerase and sugar backbone of primer and template strands. *J. Am. Chem. Soc.* 124, 910–911.
- Thiede, C., Bornhauser, M., Oelschlagel, U., Brendel, C., Leo, R., Daxberger, H., Mohr, B., Florek, M., Kroschinsky, F., Geissler, G. *et al.* (2001). Sequential monitoring of chimerism and detection of minimal residual disease after allogeneic blood stem cell transplantation (BSCT) using multiplex PCR amplification of short tandem repeat-markers. *Leukemia* 15, 293–302.
- 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.
- Whitcombe, D., Brownie, J., Gillard, H.L., McKechnie, D., Theaker, J., Newton, C.R., and Little, S. (1998). A homogeneous fluorescence assay for PCR amplicons: its application to real-time, single-tube genotyping. *Clin. Chem.* 44, 918–923.
- Wilhelm, J., Reuter, H., Tews, B., Pingoud, A., and Hahn, M. (2002). Detection and quantification of insertion/deletion-variations by allele-specific real-time-PCR: application for genotyping and chimerism analysis. *Biol. Chem.* 383, 1423–1433.
- Wilhelm, J., Pingoud, A., and Hahn, M. (2003) SoFAR: software for fully automatic evaluation of real-time PCR data. *Biotechniques* 34, 324–332.
- Wittwer, C.T., Fillmore, G.C., and Hillyard, D.R. (1989). Automated polymerase chain reaction in capillary tubes with hot air. *Nucleic Acids Res.* 17, 4353–4357.
- Wu, D.Y., Ugozzoli, L., Pal, B.K., and Wallace, R.B. (1989). Allele-specific enzymatic amplification of beta-globin genomic DNA for diagnosis of sickle cell anemia. *Proc. Natl. Acad. Sci. USA* 86, 2757–2760.
- Zetterquist, H., Mattsson, J., Uzunel, M., Nasman-Bjork, I., Svenberg, P., Tammik, L., Bayat, G., Winiarski, J., and Ringden, O. (2000). Mixed chimerism in the B cell lineage is a rapid and sensitive indicator of minimal residual disease in bone marrow transplant recipients with pre-B cell acute lymphoblastic leukemia. *Bone Marrow Transplant.* 25, 843–851.
- Zhang, J., Tong, K.L., Li, P.K., Chan, A.Y., Yeung, C.K., Pang, C.C., Wong, T.Y., Lee, K.C., and Lo, Y.M. (1999). Presence of donor- and recipient-derived DNA in cell-free urine samples of renal transplantation recipients: urinary DNA chimerism. *Clin. Chem.* 45, 1741–1746.