2032–2034 *Nucleic Acids Research, 1997, Vol. 25, No. 10* *© 1997 Oxford University Press*

Direct DNA sequence determination from total genomic DNA

Christian Kilger* and Svante Pääbo

Institute of Zoology, University of Munich, PO Box 202126, D-80021 Munich, Germany

Received February 11, 1997; Revised and Accepted April 1, 1997

ABSTRACT

It is possible to perform a combined amplification and sequencing reaction ('DEXAS') directly from complex DNA mixtures by using two thermostable DNA polymerases, one that favours the incorporation of deoxynucleotides over dideoxynucleotides, and one which has a decreased ability to discriminate between these two nucleotide forms. During cycles of thermal denaturation, annealing and extension, the former enzyme primarily amplifies the target sequence whereas the latter enzyme primarily performs a sequencing reaction. This method allows the determination of single-copy nuclear DNA sequences from amounts of human genomic DNA comparable to those used to amplify nucleotide sequences by the polymerase chain reaction. Thus, DNA sequences can be easily determined directly from total genomic DNA.

Cycle sequencing' (1) generates a sequence reaction during repeated cycles of thermal denaturation and extension of an oligonucleotide primer using a thermostable DNA polymerase. One of its advantages is that it requires drastically less template DNA than other DNA sequencing approaches. However, since it does not exponentially amplify the target sequence, as is achieved by the polymerase chain reaction (PCR) (2), it has hitherto been impossible to use cycle sequencing procedures to directly determine nucleotide sequences from complex genomes. Attempts to develop such methods have yielded two-stage protocols, e.g. 'coupled amplification and sequencing' or CAS (3,4), where in a first stage a sequencing template is generated by PCR, and in a second stage sequencing as well as additional amplification of the target sequence is achieved. Recently, a further development, 'direct exponential amplification and sequencing' or DEXAS (5), achieved the amplification of the template DNA as well as the sequencing reaction simultaneously in a single reaction, using two primers, a mixture of deoxynucleotides (dNTP) and dideoxynucleotides (ddNTP), and a thermostable DNA polymerase carrying the F667Y mutation (6) (e.g. ThermoSequenase™, Taquenase™, AmpliTaqFS™), which allows the enzyme to incorporate ddNTPs with an efficiency approaching that of dNTPs. Of importance for the generation of readable sequences by DEXAS is that unequal ratios of the two primers are used and that these are of sufficient length (>25mers) to allow a high temperature to be maintained throughout the reaction. However, in spite of these measures, impractically high amounts of genomic DNA (>0.8 μg) are needed to determine single-copy sequences in the human genome by DEXAS (5).

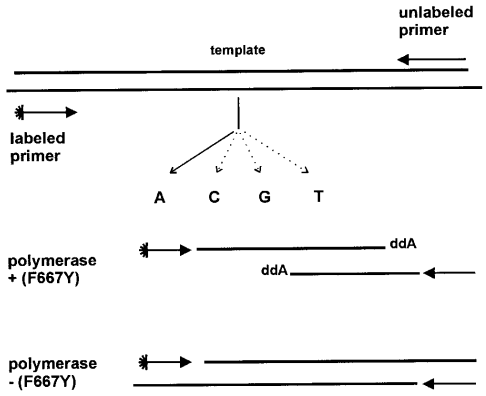

Figure 1. Schematic illustration of DEXAS and the 'two polymerase' concept. Each reaction tube contains (among other components) template DNA, one labeled and one unlabeled primer, dNTPs, one of the four ddNTPs, one DNA polymerase carrying a mutation reducing its discrimination against ddNTPs (here F667Y), and one DNA polymerase which does not. The former enzyme will incorporate dNTPs as well as ddNTPs, thereby creating a sequencing ladder, while the latter polymerase will incorporate preferentially dNTPs, thus creating mainly full-length products, which supply the reaction with additional sequencing template.

We speculated that the requirement for large amounts of template DNA in DEXAS might be due to the fact that many extension products are terminated through the incorporation of a ddNTP rather than extended to the other primer binding site, thereby stifling the exponential amplification of the target sequence. Therefore, we investigated whether a combination of two DNA polymerases that differ in their ability to incorporate ddNTPs might solve this problem by uncoupling the production of templates from the production of sequencing ladders (Fig. 1).

Two oligonucleotides spanning a 382 bp region of the CCR5 gene (7) were synthesized, one of which (CCR5-2) was labeled at the 5′-end with fluorescein, whereas the other (CCR5-1) was unlabeled. Three pairs of DEXAS reactions were set up, each containing 6 pmol of the labeled primer, 3 pmol of the unlabeled primer, 2 μl ThermoSequenase™ reagents, containing thermostable DNA polymerase with the F667Y mutation (6), reaction buffer, dNTPs, ddNTPs as well as 500, 120 and 60 ng of whole genomic

*To whom correspondence should be addressed. Tel: +49 89 590 2521; Fax: +49 89 590 2474; Email: kilger@zi.biologie.uni-muenchen.de

Downloaded from http://nar.oxfordjournals.org/ at Max Planck Institut Fuer Evolutionaere Anthropologie on June 23, 2015

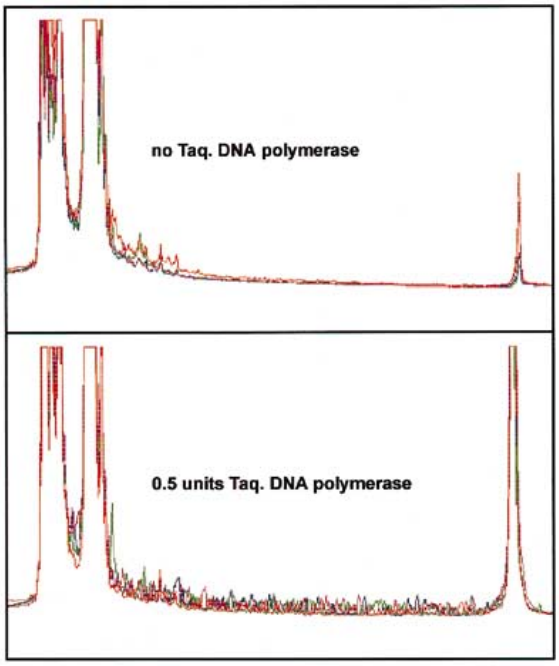

Figure 2. Sequencing reactions from 120 ng of template DNA with and without *Taq* DNA polymerase. The following oligonucleotides were used: (CCR5-1) 5′-GGC TGG TCC TGC CGC TGC TTG TCA T-3′; (CCR5-2) 5′-CTG CTC CCC AGT GGA TCG GGT GTA AAC-3′. Two 24 µl mixtures containing 6 pmol labeled and 3 pmol unlabeled primers, 120 ng whole genomic DNA and either 0.5 U of AmpliTaq™ DNA Polymerase (Perkin-Elmer, Norwalk, CT, USA) or no extra *Taq* polymerase were prepared, and aliquots of 6 µl added to 2 µl reagent mix (A, C, G and T, the supplier of the ThermoSequenase™ kit does not provide information on the contents of these solutions). Sequencing reactions were performed in a thermal cycler with heatable lid (MJ Research, Watertown, MA). The reactions were stopped by the addition of 5 µl of formamide containing 20 mM EDTA (pH 7.4) and 6 mg/ml dextran blue, followed by denaturation for 4 min at 95°C. Sequence reactions were analyzed on an A.L.F.™ (Pharmacia Biotech, Uppsala, Sweden). In all cases, HydroLink Long Ranger™ (FMC, Rockland, ME) gels and 30 cm glass plates were used. Gel conditions were as recommended by the supplier.

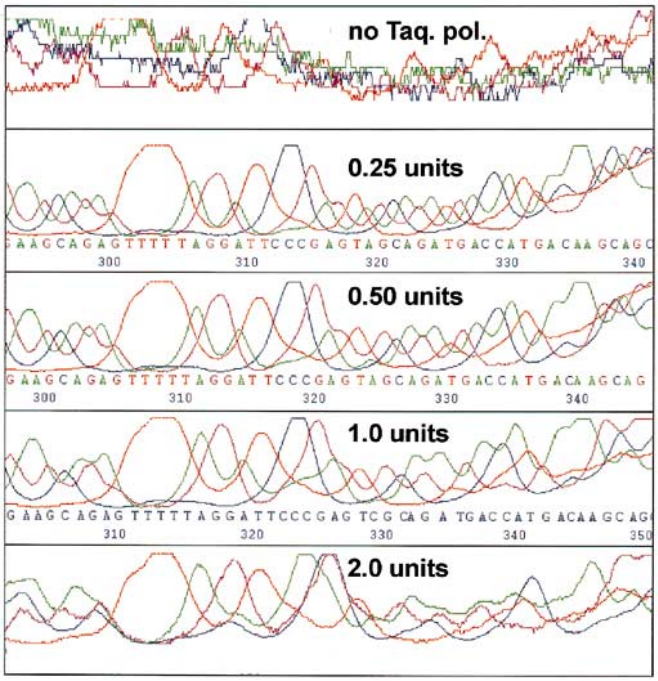

Figure 3. The effect of adding various amounts of *Taq* DNA polymerase to DEXAS reactions. Aliquots of 60 ng of whole genomic DNA were subjected to a DEXAS reaction using 6 pmol of a FITC labeled primer (CCR5-2) and 3 pmol of an unlabeled primer (CCR5-1). The region depicted in the windows is 20 bp away from the end of the DEXAS product. The amounts of *Taq* DNA polymerase added to the reactions are indicated. The A.L.F.™ software was not able to process a sequence in cases where 0 or 2 U of *Taq* DNA polymerase had been added.

DNA, respectively. One of the two reactions in each pair was in addition supplied with 0.5 U of unmodified *Taq* DNA polymerase. The reactions were incubated at 95°C for 3 min to allow complete denaturation of the template DNA before 45 cycles, each composed of 30 s at 68°C and 40 s at 95°C, were performed. The reactions were terminated by the addition of formamide, heated at 95°C for 4 min, and loaded on an A.L.F.™ sequencing device.

In all cases where *Taq* DNA polymerase had been added to the reactions, the signals towards the end of the sequencing reaction that correspond to full length amplification products were enhanced (Fig. 2). In addition, the artefacts within the first 40 bases of the DEXAS reactions (5) were reduced in intensity and the signal strength was higher and more uniform throughout the reaction than where only ThermoSequenase™ was used. When 500 ng of genomic DNA was used as a template, both reactions allowed the calling of approximately equally many bases (354 and 344, respectively) with similar numbers of ambiguities (12 and 8, respectively). In contrast, when 120 ng of genomic DNA was used to initiate the reactions, the A.L.F.™ software was able to process only a few bases when only ThermoSequenase™ had been added whereas 340 bases with five ambiguities were called when *Taq* DNA polymerase was also present (Fig. 2). Similarly, when 60 ng of genomic DNA was used, no readable signal was obtained without *Taq* DNA polymerase whereas 348 bases with six ambiguities were called in its presence. Thus, *Taq* DNA polymerase drastically increases the ability of DEXAS to determine DNA sequences from small amounts of genomic DNA.

In order to determine the optimal amount of *Taq* DNA polymerase necessary to enhance the DEXAS reactions, five reactions were performed to which 0.0, 0.25, 0.50, 1.0 and 2.0 U of *Taq* DNA polymerase were added. The results (Fig. 3) show that whereas no readable sequencing reaction is generated in the absence of *Taq* DNA polymerase, a clear and processable signal up to 20 bp away from the end of the template is generated in the presence of 0.25 or 0.50 U of *Taq* DNA polymerase. Where 1.0 U had been added a readable signal was generated but a number of full stops as well as an increase in the background resulted in an increase in ambiguous bases called. Where 2.0 U of *Taq* DNA polymerase were added, the sequence reaction was clearly inferior. In subsequent experiments, the combination of 2 µl of ThermoSequenase™ and 0.5 U of *Taq* DNA polymerase has been used to routinely sequence a segment of the p53 gene directly from total genomic DNA as well as inserts in plasmid vectors from the total DNA released from bacterial colonies. In general, DEXAS performs well on amplification products from single copy genes of up to at least 700 bp. The requirement to use different concentrations of the two primers seems less stringent than when a single DNA polymerase is used (5). Furthermore, DEXAS is routinely used to determine a 521 bp segment of the human mitochondrial control region using 0.8 µl (units not defined) of AmpliTaqFS™

Downloaded from http://nar.oxfordjournals.org/ at Max Planck Institut Fuer Evolutionaere Anthropologie on June 23, 2015

(Perkin-Elmer, Norwalk, CT, USA), 0.5 U *Taq* DNA polymerase (Perkin-Elmer, Norwalk, CT, USA), a 10× buffer consisting of 160 mM Tris–HCl (pH 9.5), 40 mM $MgCl_2$, 2.8% Tween-20™, 0.26% NP-40™, 0.56 mM 2-mercaptoethanol, and nucleotide mixes consisting of 1 mM dATP, 1 mM dCTP, 1 mM dGTP and 1 mM dTTP, respectively, where the A, C, G and T mixes contain 2.5µM of the respective dideoxynucleotides.

In conclusion, the combination of *Taq* DNA polymerase and a DNA polymerase with a limited ability to discriminate between dNTPs and ddNTPs allows the DEXAS procedure to determine DNA sequences of single-copy genes in mammalian genomes directly from amounts of total genomic DNA that are comparable to those conventionally used to initiate PCR amplification. Thus, amplification and determination of target sequences can now be routinely performed in one single reaction.

ACKNOWLEDGEMENTS

We are indebted to Karin Bauer, Beate Stiening and Walter Schartau for expert technical help and oligonucleotide synthesis.

REFERENCES

1 Murray,V. (1989) *Nucleic Acids Res.* **17**, 8889.
2 Mullis,K.B. and Fallona,F. (1987) *Methods Enzymol.* **155**, 335–339.
3 Ruano,G. and Kidd,K. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 2815–2819.
4 Sarkar,G. and Bolander,M.E. (1995) *Nucleic Acids Res.* **23**, 1269–1270.
5 Kilger,C. and Pääbo,S. (1997) *Biol. Chem.* **378**, 99–105.
6 Tabor,S. and Richardson,C.C. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 6339–6343.
7 Raport,C.J., Gosling,J., Schweickart,V.L., Gray,P.W. and Charo,I.F. (1996) *J. Biol. Chem.* **271**, 17161–17166.

Downloaded from http://nar.oxfordjournals.org/ at Max Planck Institut Fuer Evolutionaere Anthropologie on June 23, 2015