Terminal transferase-dependent PCR: a versatile and sensitive method for in vivo footprinting and detection of DNA adducts

Jun-ichiro Komura and Arthur D. Riggs*

Biology Department, Beckman Research Institute of the City of Hope, Duarte, CA 91010, USA

Received October 30, 1997; Revised and Accepted February 11, 1998

ABSTRACT

We report here a new, sensitive and versatile genomic sequencing method, which can be used for in vivo footprinting and studies of DNA adducts. Starting with mammalian genomic DNA, single-stranded products are made by repeated primer extension; these products are subjected to homopolymeric ribonucleotide tailing at the 3′ termini with terminal deoxynucleotidyl transferase and then ligated to a double-stranded linker having a complementary 3′ overhang, and used for PCR. This terminal transferase-dependent PCR (TDPCR) method can generate band signals many-fold stronger than conventional ligation-mediated PCR (LMPCR). A UV photofootprint in the mouse Xist gene promoter can be easily detected using TDPCR. No special enzymes or chemical reagents are needed to convert DNA adducts into strand breaks. Any lesion that blocks primer extension should be detectable.

INTRODUCTION

Only a few methods are useful for quantitatively displaying DNA lesions and chromatin structure in mammalian cells at single-nucleotide resolution (1,2). One such method is ligation-mediated PCR (LMPCR) (3–5), which is commonly used because it combines nucleotide-level resolution with the sensitivity of PCR. LMPCR has been used successfully in this laboratory and others for numerous in vivo studies of mammalian cells, especially for the detection of protein–DNA interactions (footprints) (1), for the analysis of cytosine methylation (3) and for the mapping of DNA damage (6,7). In the conventional method of LMPCR (Fig. 1, left), substrate genomic DNA is cut at the sites of altered bases either with a specific enzyme or by use of chemical reagents (e.g. Maxam–Gilbert cleavage). The resulting single-strand breaks are converted to blunt-ended termini by extension from a gene-specific primer and are ligated to a double-stranded linker. The sequences between the linker and a second (nested) gene-specific primer are amplified by PCR and the products are visualized as radioactive sequence ladders.

LMPCR is very sensitive but has some limitations. First, it measures directly only nicks or breaks in DNA. Secondly, the 5′ end of the template molecule must be phosphorylated (or phosphorylatable) and ligatable, because it must undergo blunt-end ligation after primer extension. Thirdly, only molecules in which primer extension has continued to the end of the template strand are able to participate in blunt-end ligation. Thus, prematurely terminated molecules are invisible to LMPCR. Finally, several hundred bands usually are seen in every LMPCR ladder, and each band must start from at least one genomic template molecule. Therefore, in practice, thousands of molecules of genomic DNA are needed to avoid poor quantitation and missing bands due only to statistical sampling fluctuations.

The Terminal transferase-Dependent PCR (TDPCR) method reported here (Fig. 1, right) depends on cohesive-end ligation to the 3′ ends of DNA molecules resulting from primer extension, followed by controlled ribonucleotide tailing by terminal deoxynucleotidyl transferase (8). TDPCR provides an alternative to LMPCR that should eliminate or lessen each of the limitations described above. TDPCR does not require ligatable 5′ ends, detects prematurely terminated molecules, can sample each template molecule many times and thus can be more sensitive than LMPCR. The method should aid footprint and chromatin structure experiments that need high sensitivity and should enable use of new footprinting or DNA-damaging agents whose products on DNA cannot easily be converted to ligatable 5′ termini. A photofootprint in the mouse Xist gene promoter is shown to be clearly detected by TDPCR.

MATERIALS AND METHODS

Cells and DNA preparation

Culture of BML-2 cells (9), DNA isolation, UV irradiation of cells or DNA and separation of the expressed Xist allele from the silent allele were performed as described by Komura et al. (10). Cleavage of UV-irradiated DNA at the sites of cyclobutane pyrimidine dimers using photolyase and T4 endonuclease V, or at the sites of pyrimidine–pyrimidone (6–4) photoproducts by treatment with piperidine, was carried out as described (11,12). Maxam–Gilbert cleavage of DNA was done as described (13), using 80 µg of genomic DNA.

*To whom correspondence should be addressed. Tel: +1 626 301 8352; Fax: +1 626 358 7703; Email: ariggs@smtplink.coh.org
Figure 1. Schematic outline of LMPCR and TDPCR. A DNA lesion in the starting DNA is indicated by a small diamond.

Primers and linkers

The gene-specific primers A1, A2 and A3 for the mouse Xist promoter region are the same as previously described (10). Linker α and the linker-primer (upper strand of α; Fig. 2) are the standard oligonucleotides usually used for LMPCR (5). Linkers β, γ, δ and ε are related to α as shown in Figure 2. Linkers γ, δ and ε were synthesized by the DNA synthesis shared resource facility of the City of Hope with an aminopentyl blocking group at their 3′ termini (14). For TDPCR the 5′ terminus of the lower oligonucleotide of each linker was phosphorylated by incubation in a 100 µl reaction consisting of 22.2 µM oligonucleotide, 50 U T4 polynucleotide kinase (New England Biolabs), 1× buffer supplied by the manufacturer and 1 mM ATP, at 37°C for 2 h. After inactivation of the enzyme by incubation at 65°C for 20 min, 11 µl of a 200 µM solution of the upper oligonucleotide was added. The mixture was heated to 95°C and allowed to cool gradually.

Terminal deoxynucleotidyl transferase-dependent PCR (TDPCR)

During the set up of reactions and between heating steps, manipulations were performed on ice unless stated otherwise. Genomic DNA (0.3–0.5 µg) was linearly amplified in a 30 µl reaction consisting of 1.2 U of Vent (exo−) DNA polymerase (New England Biolabs), 1× ThermoPol Buffer (New England Biolabs), extra 4 mM MgSO4, 3.3 mM Tris (pH 7.5), 0.3 mM EDTA, 250 µM each dNTP and 20 nM primer A1. Temperature cycles, which varied from 1 to 30 in frequency, were for 1 min at 95°C (5 min at 95°C for the first cycle), 3 min at 47°C, and 2 min at 72°C. After thermal cycling, each sample was incubated at 95°C for 2 min, then 20 µl was transferred into another tube containing 80 µl of a solution composed of 2.5 M ammonium acetate, 2.5 mM EDTA and 40 µg glycogen, and precipitated with 250 µl of ethanol. The precipitate was dissolved in 10 µl terminal deoxynucleotidyl transferase (TdT) mix consisting of 10 U TdT (Gibco BRL), 2× buffer supplied by the manufacturer and 4 mM rGTP, the sample was incubated at 37°C for 15 min. DNA was precipitated by addition of 80 µl of 2.5 M ammonium acetate
and 2.5 mM EDTA followed by 250 μl of ethanol. The precipitate was dissolved in 15 μl of 1/10 TE. After the addition of 10.5 μl of ligation solution [143 mM Tris (pH 7.5), 29 mM MgCl₂, 29 mM DTT and 2.9 mM ATP], 3 μl of 20 μM linker (either α, β, γ, δ or ε) and 1.5 μl of T4 DNA ligase (Promega, 3 U/μl), the mixture was incubated at 17°C overnight. After the direct addition of 70 μl of exo⁻ solution (4 U Vent exo⁻, 1.43 μl Thermopol buffer, extra 2.9 mM MgSO₄, 0.36 mM each dNTP, 0.29 μM ATP, 0.29 μM primer A2 and 0.29 μM linker primer), the sample was subjected to PCR using 23 cycles of 1 min at 95°C (4 min at 95°C for the first cycle), 2 min at 61°C and 1 min at 72°C. After the thermal cycling, 1 U of AmpliTaq DNA polymerase (Perkin Elmer) was added and the sample was further incubated at 72°C for 8 min prior to phenol–chloroform extraction and ethanol precipitation. This booster step is to ensure that all molecules have an extra nucleotide at the 3′ end (10). For LMPCR, shadow bands are sometimes seen if this step is omitted (10). Electrophoresis, electroblotting, hybridization and autoradiography were performed as described (5,10), using primer A3 to make the ³²P-labeled hybridization probe.

**Ligation-mediated PCR (LMPCR)**

Unless indicated otherwise, LMPCR was done exactly as described (10), using the same DNA samples as previously (10). T4 endonuclease V and photolyase were kindly provided by Dr R.S.Lloyd (University of Texas, Galvenston) and Dr A.Sancar (University of North Carolina, Chapel Hill), respectively.

**RESULTS**

Figure 1 shows schematically the TDPCR and LMPCR procedures. The TDPCR procedure begins with repeated primer extensions, producing multiple copies of the template strands. To enable PCR amplification of the newly synthesized, single-stranded products, a linker must be ligated to the 3′ ends. To accomplish this, we adapted the procedure of Schmidt and Mueller (8), which uses controlled ribo-tailing by TdT for cDNA cloning. As commonly used, TdT adds long deoxynucleotide tails to the 3′ terminus of single-stranded DNA molecules. However, the use of a ribonucleotide, such as rGTP, rather than a deoxyribonucleotide, limits the tailing to just a few residues. Under the conditions we use, an average of only three nucleotides are added. This homopolymeric rG tail is ligated to a double-stranded DNA linker with a complementary 3′ overhang of three cytosines. Preliminary experiments showed that a linker with an overhang of three Cs was better than two Cs (data not shown). It should be noted that in TDPCR only the lower, newly synthesized strands participate in the PCR step, while in LMPCR only the upper, old template strands participate.

Figure 3 shows data for TDPCR and LMPCR done in parallel, with the same amount of starting material in each reaction. DNA isolated from mouse cells and subjected to C-specific cleavage according to Maxam–Gilbert, C-specific cleaved DNA (C) or non-treated control DNA (N). Primer extension (ext.) was performed 1–30 times, as indicated. Several linkers (α, β, γ, δ, ε; Fig. 2) were compared. For this experiment, the temperature and time of incubations were the same for LMPCR and TDPCR, and are as described in Materials and Methods for the TDPCR method, with the exception that the 95°C denaturation step after first primer extension was not done for LMPCR, since this would preclude blunt-end ligation. The nucleotide position numbers indicate the number of bases upstream of the major transcription start site.

Figure 3, lanes 5–8, shows an investigation of four types of linkers with the same sequences but different 3′-blocking patterns (linkers β, γ, δ and ε; Fig. 2). Linker γ gives the strongest bands and lowest background, perhaps because the blocking of the lower oligonucleotide prevents any remaining TdT activity from adding deoxynucleotides during the first step of the PCR reaction while the tube is warming. The linkers δ and ε, which have the upper 3′-end blocked, give poor results. This could be due either to inhibition of ligation or perhaps to competition against the linker–primer during PCR.

Figure 3, lanes 3, 4, 6 and 9, illustrates the effect of repeating the first primer extension 1, 3, 10 or 30 times. More cycles give stronger signals. To quantitatively measure signal to noise and compare TDPCR with LMPCR, we measured in Figure 3 the intensity of several bands in both treated and non-treated DNA. The results of this analysis are shown in Figure 4. The signal intensity by TDPCR with one cycle of primer extension was 65% of that by conventional
Figure 4. Signal intensity as a function of the number of the cycles of primer extension of TDPCR. The intensity of the Maxam–Gilbert C bands in Figure 3 at positions –160, –166, –168 and –174 (LMPCR) and –158, –164, –166 and –172 (TDPCR) was measured by the use of a PhosphorImager (Molecular Dynamics). For non-treated control DNA, intensity at the corresponding areas was measured, even when at lower cycle numbers a band was not detectable. The average ± SD of the bands measured in the treated and untreated lanes is plotted as a function of cycle number. Solid symbols represent the intensity of the C bands in the treated-sample lanes; open symbols represent the intensity in the non-treated control lanes. Circles, TDPCR; triangles, conventional LMPCR.

For TDPCR, however, the signal increased from 1 to 30 cycles of first primer extension, reaching seven times that of LMPCR. The background in the lanes of non-treated control DNA did not increase significantly from 1 to 10 cycles, but did increase from 10 to 30 cycles. These results suggest that some breakage or damage to the template takes place during the thermal cycling and this becomes significant by 30 cycles. Under the conditions we used, 10 cycles of primer extension yielded the highest signal to noise, 28 times background.

It should be noted that the background in the C lanes of TDPCR (Fig. 3, bands at non-C sites in lanes 3 and 4) is somewhat higher than that in the C lane of conventional LMPCR (lane 2) and much higher than for non-treated DNA. This has been a consistent result even for DNA specifically cleaved at the other bases (data not shown), so we hypothesize that the chemical reactions and/or cleavage with piperidine may leave some DNA damage that stops primer extension. Any form of damage which stops the polymerase at the primer extension step can be detected by TDPCR. Whatever the explanation, chemically cleaved DNA, though useful to provide information on sequence and location, is not a preferred substrate for TDPCR.

Figure 5 shows results of a UV photofootprint experiment, again comparing TDPCR (lanes 1–4) with LMPCR (lanes 5–10). The two major types of damage induced by UV are cyclobutane pyrimidine dimers and pyrimidine–pyrimidone (6–4) photoproducts (15). Their formation is influenced by the sequence context and by chromatin structure (15), and UV has proven to be an excellent in vivo footprinting agent (6,12,16). We previously used UV photofootprinting and LMPCR to examine protein–DNA interactions at the promoter of the expressed Xist allele in BML-2 cells (10).

Conventional LMPCR requires cleavage at DNA lesions, so for this study, as previously (10), the DNA samples were cleaved with T4 endonuclease V at the sites of cyclobutane dimers (lanes labeled dimer), or with piperidine at the sites of (6–4) photoproducts (lanes labeled 6–4). Lanes labeled DNA, purified DNA was UV irradiated; lanes labeled Cells, cells in tissue culture were UV irradiated; lane C, DNA subjected to Maxam–Gilbert C-specific cleavage; lane T+C, DNA subjected to Maxam–Gilbert T and C-specific cleavage. The nucleotide position numbers indicate the number of bases upstream of the major transcription start site. The arrow shows nucleotide position –102, the location of a photofootprint.
nucleotide –102, indicated by arrow); the same footprint is seen by LMPCR (lanes 5 and 6). Thus, photofootprints can be visualized by TDPCR.

**DISCUSSION**

Genomic sequencing, as originally described by Church and Gilbert (19) and used for *in vivo* footprinting (20), has been improved and extended in several ways (21,22). These include prior enrichment of the target sequence (23), linear amplification by repeated extension from a gene-specific primer (16,24) and exponential amplification by PCR (3,4). Grimaldi et al. (18) successfully combined linear and exponential amplification by using T4 RNA ligase to join a single-stranded linker to the single-stranded products of primer extension. This is a useful procedure and important results were obtained for cisplatin adducts (18). However, DNA is not a preferred substrate of RNA ligase, the ligation of single-stranded DNA molecules is a slow and inefficient reaction, and the apparent $K_m$ for DNA is high (25). Using this method with Maxam–Gilbert-cleaved DNA as a substrate, even 30 cycles of first primer extension gives a signal considerably less than LMPCR (GP, Pfeifer, personal communication). To circumvent this difficulty, we modified the 3′ termini of primer extension products by ribo-tailing with TdT and confirmed the finding of Schmidt and Mueller (8) that the homopolymeric tails can be efficiently ligated to a double-stranded linker by cohesive-end ligation using T4 DNA ligase. The TDPCR method reported here consistently produces stronger signals than LMPCR. For UV-treated DNA, the signal to noise ratio is excellent, as is necessary for study of DNA adduct formation and repair (26). At 10 cycles of primer extension, the TDPCR signal was four times stronger than that of LMPCR and the bands were 28-fold above background. At 30 cycles the signal was seven times greater than LMPCR, although the signal to noise ratio decreased to only 8-fold over background, suggesting that a limit to the number of useful cycles was being approached.

LMPCR requires DNA single-strand breaks with 5′-phosphates for ligation. Some footprinting or DNA-damaging agents, such as DNsase I (27) and bleomycin (28), make this type of terminus. $N^7$-methylguanines (produced by dimethyl sulfate) (4) and (6–4) photoproducts (6) can be converted to strand breaks by treatment with piperidine. Cyclobutane dimers (11) and some oxidized bases (29) can be recognized by lesion-specific endonucleases, but these reagents require a special source to which many may not have access. Many other DNA-damaging agents have not been usable with LMPCR due to the absence of adequate cleavage methods. Recently, *Escherichia coli* UvrABC endonuclease was used in LMPCR experiments to cleave DNA at the sites of benzo[a]pyrene adducts (30,31). This enzyme recognizes a wide spectrum of bulky adducts (32); however, even under optimal conditions, UvrABC endonuclease incises the DNA with low efficiency and only a fraction of the adducts are converted to strand breaks (33).

To detect lesions in the template DNA, TDPCR employs a polymerase rather than an endonuclease or a chemical cleavage reagent. The use of the polymerase stop assay (16,17) makes TDPCR a versatile method. Without using a special damage-recognizing enzyme or reagent, TDPCR should be useful for mapping many kinds of damage in addition to the thymidine dimers detected in this study; all that is required is that the damage stops the polymerase. The polymerase stop assay may preferentially detect the lesions with the most biological significance, for it has been noted that highly mutagenic adducts stop DNA polymerase more strictly than less mutagenic adducts (34). Lesions produced by ionizing radiation have been difficult to study *in vivo* at single-nucleotide resolution. Radiation induces many types of damage, including base damage and strand breaks with various termini (15,35). Some of the lesions cannot be detected with presently available enzymes or chemicals, but should be amenable to assay by TDPCR.

**ACKNOWLEDGEMENTS**

This work was supported by National Institutes of Health grants GM50575 and CA69449.

**REFERENCES**