Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 9436–9440, December 1988
Biochemistry

DNA sequencing with *Thermus aquaticus* DNA polymerase and direct sequencing of polymerase chain reaction-amplified DNA

(thermophilic DNA polymerase/chain-termination/processivity/automation)

MICHAEL A. INNIS*, KENNETH B. MYAMBO, DAVID H. GELFAND, AND MARY ANN D. BROW

Department of Microbial Genetics, Cetus Corporation, 1400 Fifty-Third Street, Emeryville, CA 94608

Communicated by Hamilton O. Smith, September 8, 1988 (received for review August 15, 1988)

ABSTRACT **The highly thermostable DNA polymerase from *Thermus aquaticus* (*Taq*) is ideal for both manual and automated DNA sequencing because it is fast, highly processive, has little or no 3′-exonuclease activity, and is active over a broad range of temperatures. Sequencing protocols are presented that produce readable extension products >1000 bases having uniform band intensities. A combination of high reaction temperatures and the base analog 7-deaza-2′-deoxyguanosine was used to sequence through G+C-rich DNA and to resolve gel compressions. We modified the polymerase chain reaction (PCR) conditions for direct DNA sequencing of asymmetric PCR products without intermediate purification by using *Taq* DNA polymerase. The coupling of template preparation by asymmetric PCR and direct sequencing should facilitate automation for large-scale sequencing projects.**

DNA sequencing by the Sanger dideoxynucleotide method (1) has undergone significant refinement in recent years, including the development of additional vectors (2), base analogs (3, 4), enzymes (5), and instruments for partial automation of DNA sequence analysis (6–8). The basic procedure involves (*i*) hybridizing an oligonucleotide primer to a suitable single- or denatured double-stranded DNA template; (*ii*) extending the primer with DNA polymerase in four separate reaction mixtures, each containing one α-labeled dNTP, a mixture of unlabeled dNTPs, and one chain-terminating ddNTP; (*iii*) resolving the four sets of reaction products on a high-resolution polyacrylamide/urea gel; and (*iv*) producing an autoradiographic image of the gel, which can be examined to infer the DNA sequence. The current commercial instruments address nonisotopic detection and computerized data collection and analysis. The ultimate success of large-scale sequencing projects will depend on further improvements in the speed and automation of the technology. These include automating the preparation of DNA templates and performing the sequencing reactions.

One technique that appears to be ideally suited for automating DNA template preparation is the selective amplification of DNA by the polymerase chain reaction (PCR) (9). With this method, segments of single-copy genomic DNA can be amplified >10 million-fold with very high specificity and fidelity. The PCR product can then either be subcloned into a vector suitable for sequence analysis or, alternatively, purified PCR products can be sequenced (10–13).

The advent of *Taq* DNA polymerase greatly simplifies the PCR procedure because it is no longer necessary to replenish enzyme after each PCR cycle (14). Use of *Taq* DNA polymerase at high annealing and extension temperatures increases the specificity, yield, and length of products that can be amplified and, thus, increases the sensitivity of PCR for detecting rare target sequences. Here we describe other properties of *Taq* DNA polymerase that pertain to its advantages for DNA sequencing and its fidelity in PCR.

MATERIALS

Enzymes. Polynucleotide kinase from T4-infected *Escherichia coli* cells was purchased from Pharmacia. *Taq* DNA polymerase, a single subunit enzyme with relative molecular mass of 94 kDa (specific activity, 200,000 units/mg; 1 unit corresponds to 10 nmol of product synthesized in 30 min with activated salmon sperm DNA), was purified from *Thermus aquaticus*, strain YT-1 (ATCC no. 25104), according to S. Stoffel and D.H.G. (unpublished data). More recently, *Taq* DNA polymerase (GeneAmp) was purchased from Perkin–Elmer Cetus Instruments. The polymerase (5–80 units/μl) was stored at −20°C in 20 mM Tris·HCl, pH 8.0/100 mM KCl/0.1 mM EDTA/1 mM dithiothreitol/autoclaved gelatin (200 μg/ml)/0.5% Nonidet P-40/0.5% Tween 20/50% (vol/vol) glycerol.

Nucleotides, Oligonucleotides, and DNA. 2′-Deoxy-, and 2′,3′-dideoxynucleotide 5′-triphosphates (dNTPs and ddNTPs) were obtained from Pharmacia. 7-Deaza-2′-deoxyguanosine 5′-triphosphate (c^7GTP) was from Boehringer Mannheim. dATP[α-^{35}S] (650 Ci/mmol; 1 Ci = 37 GBq) was from Amersham, and [γ-^{32}P]ATP was from New England Nuclear. Oligonucleotide primers for sequencing and PCR were synthesized on a Biosearch 8700 DNA Synthesizer. Oligonucleotide primers were 5′-end-labeled (3×10^6 cpm/pmol) with [γ-^{32}P]ATP and T4 polynucleotide kinase (15). Single-stranded M13 DNA templates were prepared as described (16).

SEQUENCING METHODS

Annealing Reaction. Single annealing and labeling reactions were performed for each set of four sequencing reactions. The annealing mixture contained 5 μl of oligonucleotide primer (0.1 pmol/μl) in 6× *Taq* sequencing buffer (10 mM $MgCl_2$/10 mM Tris·HCl, pH 8.0, at room temperature), and 5 μl of template DNA (0.05–0.5 pmol). The mixture was heated to 90°C for 3 min, incubated at 42°C for 20 min, cooled to room temperature, and briefly spun to collect the fluid at the bottom of the tube.

Labeling Reaction. To the 10-μl annealing reaction mixture were added 2 μl of labeling mix (10 μM dGTP/5 μM dCTP/5 μM TTP in 10 mM Tris·HCl, pH 8.0), 2 μl of dATP[α-^{35}S] (5 μM in 10 mM Tris·HCl, pH 8.0), 2 μl of *Taq* DNA polymerase (5 units/μl in dilution buffer: 10 mM Tris·HCl, pH 8.0/0.5% Tween 20/0.5% Nonidet P-40), and 4 μl of H_2O. The labeling reaction mixture was incubated for 1 min at 37°C (see Fig. 3). Note: for sequencing with 5′-labeled primers, the addition of

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: c^7GTP, 7-deaza-2′-deoxyguanosine 5′-triphosphate; PCR, polymerase chain reaction.

*To whom reprint requests should be addressed.

dNTP[α-^{35}S] and the labeling reaction step were omitted, and the volume was made up with 10 mM Tris·HCl (pH 8.0).

Extension–Termination Reaction. Four separate extension–termination reactions were performed in 96-well microtiter plates (Falcon 3911) for each labeled template, using concentrated deoxy/dideoxy termination mixes: "G-mix" (30 μM each dNTP, 0.25 mM ddGTP, 0.37 mM $MgCl_2$); "A-mix" (30 μM each dNTP, 1.0 mM ddATP, 1.12 mM $MgCl_2$); "T-mix" (30 μM each dNTP, 1.5 mM ddTTP, 1.62 mM $MgCl_2$); and "C-mix" (30 μM each dNTP, 0.5 mM ddCTP, 0.62 mM $MgCl_2$). Aliquots (4 μl) from the labeling reaction mixtures were added at room temperature to wells containing 2 μl of the appropriate termination mix. Reaction mixtures were overlaid with 10 μl of mineral oil to prevent evaporation and then incubated at 70°C for 1–3 min. Reactions were stopped by the addition of 2 μl of 95% deionized formamide with 0.1% bromophenol blue, 0.1% xylene cyanol, and 10 mM EDTA (pH 7.0). Samples were heated at 80°C for 3 min before loading 1–2 μl onto a buffer gradient sequencing gel (17).

Asymmetric PCRs. The template for PCRs was single-stranded M13mp10 DNA containing a 400-base insert in the *Eco*RI site of the polylinker. Oligonucleotides (20-mers) were synthesized to flank the polylinker, immediately outside of the universal "−20" and "Reverse" sequencing primer binding sites, and these were designated RG05 (5′-AGGGTTTTCCCAGTCACGAC-3′)and RG02(5′-GTGTGG-AATTGTGAGCGGAT-3′), respectively. Each PCR contained 20 pmol of one primer and 0.2 pmol of the other, 20 μM each dNTP, 1–10 ng of DNA, 1× modified PCR buffer (10 mM Tris·HCl, pH 8.0/3.0 mM $MgCl_2$), 0.05% each of Tween 20 and Nonidet P-40, and 2.5 units of *Taq* DNA polymerase in a total vol of 100 μl. Reactions were performed in 0.5-ml microcentrifuge tubes with the Perkin–Elmer Cetus Thermal Cycler. The thermal profile involved 35 cycles of denaturation at 93°C for 30 sec, primer annealing at 50°C for 1 min, and extension at 72°C for 1 min.

Sequencing of PCR Products. Aliquots of the PCRs were directly incorporated into dideoxy chain-termination sequencing reaction mixtures. A set of four base-specific chain-termination mixes was made up, each in 1× modified PCR buffer and 20 μM each dNTP. The individual mixes contained 250 μM ddGTP, 1.28 mM ddATP, 1.92 mM ddTTP, or 640 μM ddCTP. For each PCR product to be sequenced, four wells on a 96-well microtiter plate were labeled G, A, T, and C, and each well received 2.5 μl of the appropriate termination mix. A 20-μl aliquot of each PCR mixture was mixed with 0.5 μl of fresh *Taq* DNA polymerase (48 units/μl), 1 μl of the appropriate ^{32}P-labeled M13 "forward" or "reverse" sequencing primer (5′-GTAAAACGA-CGGCCAGT-3′, 5′-AACAGCTATGACCATG-3′, respectively; 1.2 pmol/μl) and 10.5 μl of 1× modified PCR buffer. The PCR/primer preparation was immediately dispensed in 7.5-μl aliquots into the wells containing the termination mixes and mixed with the pipette. The reactions were incubated at 70°C for 2 min, and stopped by the addition of 4 μl of 91% formamide with 20 mM EDTA (pH 8.0) and 0.05% each of xylene cyanol and bromophenol blue. Aliquots (5μl) of these reaction mixtures were heated to 75°C for 5 min, and 1–2 μl was loaded on a buffer gradient sequencing gel.

RESULTS

***Taq* DNA Polymerase Is Fast and Very Processive.** The experiments shown in Fig. 1 involved extending a 5′ ^{32}P-labeled 30-mer primer hybridized to M13mp18 single-stranded DNA with an equimolar amount of *Taq* DNA polymerase at various temperatures. Aliquots were taken over time and analyzed as described. Within 2 min at 70°C the entire 7.25-kilobase template was replicated; this corresponds to an extension rate in excess of 60 nucleotides per sec. *Taq* DNA polymerase retained significant activity at lower temperatures with calculated extension rates of 24, 1.5, and 0.25 nucleotides per sec at 55°C, 37°C, and 22°C, respectively. At 70°C and at substantial substrate excess (0.1:1 molar ratio of polymerase to primer/template; data not shown) most of the initiated primers were completely extended prior to reinitiation on new primer/template substrate. These results showed *Taq* DNA polymerase to be highly processive.

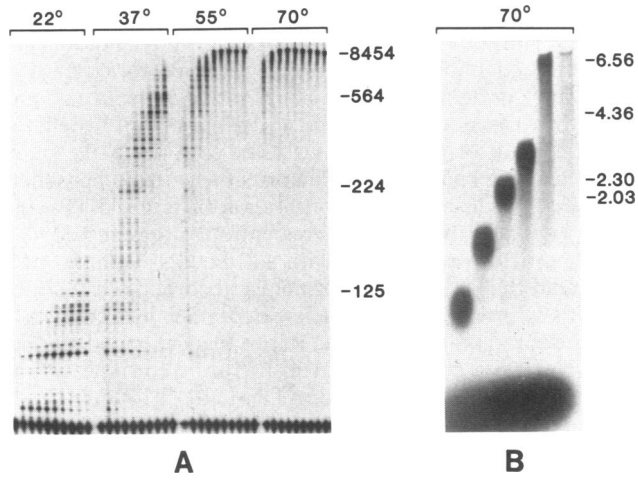

FIG. 1. Autoradiographs of a polyacrylamide/urea gel (*A*) and an alkaline agarose gel (*B*) comparing the extension rate of *Taq* DNA polymerase at different temperatures. Time points are as follows: (*A*) 0 (no enzyme), 15, 30, and 45 sec, and 1, 2, 3, 5, 7, and 10 min; (*B*) 15, 30, ahd 45 sec, and 1, 2, and 5 min. M13mp18 template DNA (2 pmol) and 5′ ^{32}P-labeled primer DG48 (4 pmol) (5′-GGGAAGGGC-GATCGGTGCGGGCCTCTTCGC-3′, calculated t_m = 78°C in 0.1 M Na^+) were annealed in 40 μl of 10 mM Tris·HCl, pH 8.0/5 mM $MgCl_2$, as described. The reaction mixtures were adjusted to 200 μM each dNTP, 0.05% each Tween 20 and Nonidet P-40, 10 mM Tris·HCl (pH 8.0), 50 mM KCl, and 2.5 mM $MgCl_2$, in a total vol of 80 μl, then brought to the desired temperature in the absence of enzyme. *Taq* DNA polymerase (2 pmol) was added to start the reactions, and 8-μl aliquots were removed and added to 8 μl of a stop solution containing 100 mM NaOH, 2 mM EDTA, 5% Ficoll, and 0.1% each bromophenol blue and xylene cyanol. The aliquots were further diluted to 40 μl with half-strength stop solution. Aliquots (5 and 20 μl) of the time points were denatured at 80°C for 3 min and loaded onto a buffer gradient sequencing gel (17) and a 0.8% alkaline agarose gel (18), respectively. Reduction in the signal of full-length product observed at the 5-min time point (*B*) is consistent with the presence of significant polymerization-dependent 5′ exonuclease activity associated with the enzyme. Markers refer to the number of bases incorporated in nucleotides (*A*) or in kilobases (*B*).

Factors Affecting the Sequencing Reactions. The buffer (14) for *Taq* DNA polymerase PCRs was modified for DNA sequencing. Each component was investigated individually by using a 5′ ^{32}P-labeled M13 forward sequencing primer (17-mer) and an M13 single-stranded DNA template. Sequencing reactions were performed as described above except that the labeling step was omitted. KCl was included at 0–300 mM. The best extensions occurred in the absence of KCl; at 50 mM KCl there was slight inhibition of enzyme activity, and at ≥75 mM KCl, the activity of *Taq* DNA polymerase was significantly inhibited. The presence of gelatin, which acts as an enzyme stabilizer in PCRs, did not affect the sequencing reactions *per se*; however, it produced distortions during electrophoresis. Addition of nonionic detergents (final concentrations, 0.05% Tween 20 and 0.05% Nonidet P-40) both stimulated the activity of the *Taq* DNA polymerase and reduced the background caused by false terminations from the enzyme (data not shown).

Taq DNA polymerase is sensitive to the free magnesium ion concentration. Accordingly, stock dNTPs and ddNTPs contained equimolar amounts of $MgCl_2$. We varied all four deoxynucleotide triphosphate concentrations between 1 and 20 μM. At concentrations of <5 μM each, or when the concentration of one dNTP was low relative to the other dNTPs, a high background of incorrect termination products was seen because of misincorporation of both dNTPs and ddNTPs. Thus, the optimum concentration for each ddNTP was empirically determined with all four dNTPs at 10 μM. We found that *Taq* DNA polymerase incorporated the four ddNTPs with varying efficiencies, and much less efficiently than the corresponding dNTPs. Ratios that generated optimal distributions of chain-termination products were [dGTP/ddGTP (1:6), dATP/ddATP (1:32), TTP/ddTTP (1:48), and dCTP/ddCTP (1:16)]. *Taq* DNA polymerase concentration was varied between 1 and 20 units per set of four reactions containing 0.2 pmol of single-stranded DNA template, 0.5 pmol of primer, and the dNTP/ddNTP concentrations just described. The signal intensity increased up to 10 units of polymerase per reaction set, representing approximately a 2.5-fold molar excess of enzyme over template/primer.

Developing a Two-Step Labeling and Extension Protocol. We then sought to develop a protocol for incorporation of labeled nucleotide during the sequencing reaction. A "Klenow-type" protocol, in which one labeled nucleotide is present at low concentration relative to the other three during the synthesis reaction, was impractical because of misincorporation of dNTPs and ddNTPs. We estimate the apparent K_m values for each of the four dNTPs to be between 10 and 20 μM. When the concentration of the labeled nucleotide was significantly below K_m (i.e., $\approx$1 μM), ddNTPs present at 80–500 μM were inappropriately incorporated at high frequency (data not shown). Concentrations higher than 1 μM for an α-^{35}S-labeled dNTP are not practical. Also, because the enzyme lacks 3'-exonuclease (proofreading) activity, misincorporated dNTPs induced chain termination. Fig. 2 shows a sequencing ladder generated in the absence of ddNTPs by forcing misincorporation of dNTPs with imbalanced dNTP concentrations. These reactions produced a doublet at most base positions, and chasing these reactions revealed that the upper band of each doublet likely represents molecules that have misincorporated a base, while the lower band represents a pause in the polymerization. Accordingly, the lower bands disappeared when the reaction was chased. Misincorporated bases appeared to be inefficiently extended by the chase.

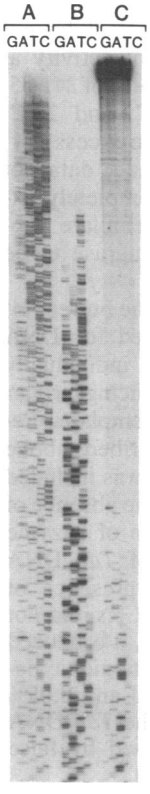

FIG. 2. Autoradiograph of a polyacrylamide/urea gel demonstrating base-specific chain termination due to misincorporation of dNTPs. The sequencing ladder generated with standard dideoxy chain terminations (*A*) is shown beside ladders generated by limiting one of the four dNTPs, before (*B*) and after (*C*) chasing with concentrated balanced dNTP mix. The standard dideoxy reactions were carried out as described for sequencing with a ^{32}P-labeled primer. In the other reactions, the primer and template were annealed in 10 μl of 10 mM Tris·HCl, pH 8.0/6 mM $MgCl_2$. Diluted *Taq* DNA polymerase (2 μl) was added and the reaction was brought to 20 μl with 10 mM Tris·HCl (pH 8.0). The sample was divided into four aliquots, identified by the nucleotide to be limited in that reaction. The "G" and "A" aliquots were brought to 0.5 μM in the limiting nucleotide and to 30 μM in the other three dNTPs; the "T" and "C" reactions were similar, with the limiting nucleotide increased to 1.5 μM. All reaction mixtures were incubated for 10 min at room temperature and then chased by addition of 0.25 vol of 10 mM Tris·HCl, pH 8.0/0.1 mM EDTA (*B*) or 250 μM (each) dNTP mix (*C*). The samples were overlaid with mineral oil and incubated at 70°C for 2 min before addition of 4 μl of formamide/EDTA stop solution. The products were denatured at 75°C for 5 min and resolved on a buffer gradient sequencing gel (17).

To circumvent these problems, we developed a two-step procedure similar in concept to one published by Tabor and Richardson for sequencing with a modified bacteriophage T7 DNA polymerase (5): an initial low-temperature labeling step using low concentrations of all four dNTPs (one of which is labeled) followed by a processive extension in the presence of higher dNTP and ddNTP concentrations. To read the sequence next to the primer, it was necessary to use both low temperature and limiting dNTP concentrations to generate an array of extension products ranging in size from a few to >100 nucleotides long. Minimum concentrations of 0.5 μM each dNTP were necessary in this step to generate signals on an overnight exposure, and increasing one of the unlabeled dNTPs to 1.0 μM made the signals very easily readable (data not shown). This effect was seen regardless of which nucleotide was increased, but increasing more than one did not provide additional benefit. The effects of temperature and incubation time on the labeling reaction are shown in Fig. 3. Termination reactions were incubated at either 55°C or 70°C using high dNTP concentrations to ensure maximum processivity and fidelity. The reactions performed at 55°C occurred at a slower rate, but there was no detectable difference in fidelity as compared with 70°C experiments. Using these conditions, we found remarkable uniformity in the band intensities, and we have not detected any idiosyncrasies in the band patterns. In addition, the same reaction conditions cover both short and long gel runs. Fig. 3 includes an autoradiograph of an extended electrophoresis, which yields DNA sequence information in excess of 1000 nucleotides from the priming site.

Using Base Analogs and High Temperature to Sequence Through G+C-Rich DNA and to Eliminate Band Compressions. Band compressions resulting from abnormal gel migration of certain sequences are frequently encountered with G+C-rich DNA templates. Substitutions of dITP (3), or the base analog c^{7}GTP (4), for dGTP have been particularly useful in resolving compression artifacts. We compared incorporation of these nucleoside triphosphates by *Taq* DNA polymerase using either an M13mp18 template or a G+C-rich insert in M13, which contains several regions of strong dyad symmetry (Fig. 4). We found that *Taq* DNA polymerase incorporated c^{7}GTP with essentially the same kinetics as dGTP and that a combination of high reaction temperature and c^{7}GTP was very efficient for resolving difficult sequences.

In contrast, inosine-containing reaction mixtures required a 4-fold higher level of dITP as compared to dGTP, the labeling reaction needed 4 min, and the ratio of ddGTP to dITP was reduced by a factor of 20 compared to dGTP. As shown in Fig. 4, dITP appears to promote frequent terminations during the extension reaction. Terminations caused by inosine result both from a higher rate of misincorporation with dITP as compared to the other dNTPs, and because *Taq* DNA polymerase lacks sufficient 3'-exonuclease activity for editing misincorporated bases. Terminations induced by dITP are reduced if the reactions are initiated at 70°C.

Coupling DNA Sequencing to the PCR. The PCRs were performed with one of the oligonucleotide primers present in a 100-fold greater concentration than the other. In this type of reaction, termed "asymmetric" PCR (13), one of the two PCR primers is depleted during the earlier thermal cycles,

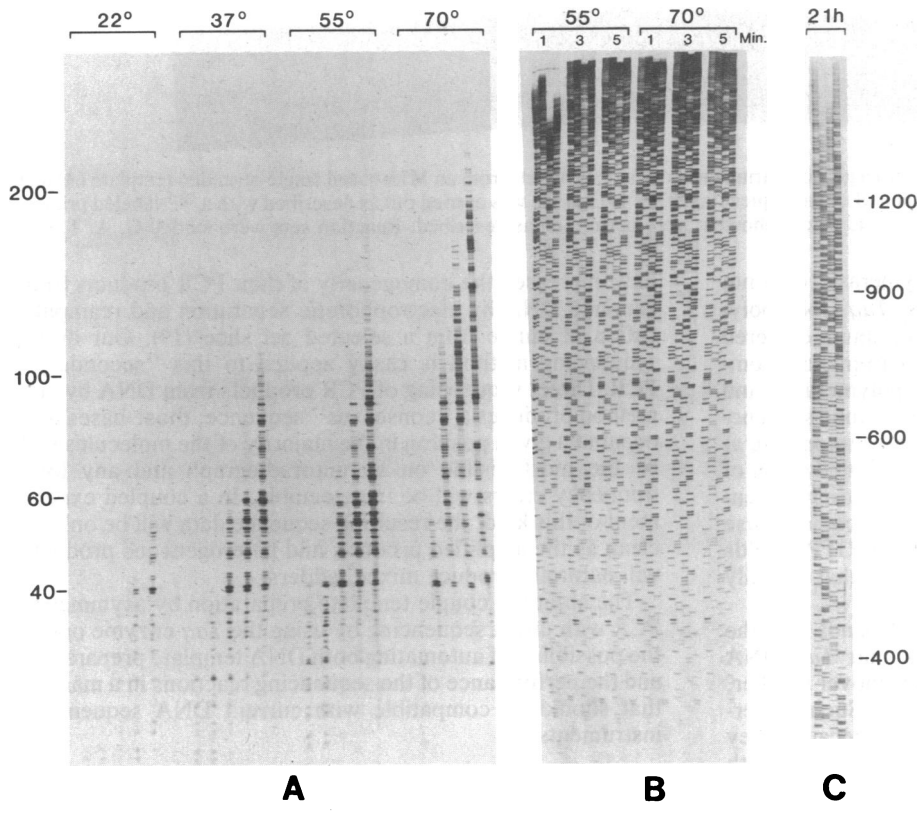

FIG. 3. Autoradiographs of polyacrylamide/urea gels showing the products of labeling reactions (*A*), extension–termination reactions performed at various temperatures (*B*), and sequencing reaction products resolved during extended electrophoresis (*C*). The labeling reactions were performed as described, except the reactions were brought up to temperature before the addition of the enzyme. Aliquots were removed at 0.5, 1, 3, 5, 7, and 10 min. The extension–termination reactions were performed as described for sequencing. Reactions were stopped and resolved on a buffer gradient sequencing gel as described in Fig. 2. Extended electrophoresis (*C*) was performed on the products of a 70°C 3-min extension–termination sequencing reaction. Samples were run at 15 W for 21 hr on a 7% acrylamide gel (18 × 50 cm × 0.4 mm) (24:1 cross-linking) with 7 M urea and 1× TBE (90 mM Tris/64.6 mM boric acid/2.5 mM EDTA, pH 8.3). Markers indicate the distance in nucleotides from the beginning of the primer. Reaction sets were loaded G, A, T, C.

and the reaction generates single-stranded product with the remaining primer.

Sequencing of asymmetric PCR-generated templates did not require purification of the product. Based on an estimated yield of 1 μg of total product, we calculate that one-third to one-half of the dNTPs initially added were used up during the PCR cycles. In addition, the stability of the dNTPs during PCR was determined to be ≈50% after 60 cycles of PCR (Corey Levenson, Cetus Corporation; personal communication). Accordingly, the termination mixes were formulated to boost the dNTPs to a final concentration of ≈10 μM in the sequencing reaction, to supply specific ddNTPs at appropriate concentrations as determined above, and to provide additional DNA polymerase. We used a ^{32}P-labeled sequencing primer to avoid purifying the PCR product and to simplify the sequencing protocol to a single extension/termination step. It is obvious that fluorescent-labeled sequencing primers could also be used, allowing the products to be analyzed on an automated DNA sequencing instrument.

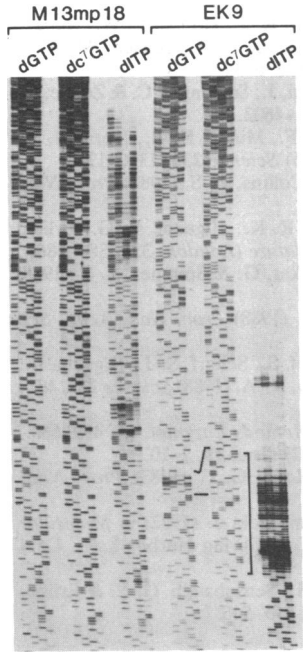

FIG. 4. Autoradiograph of a polyacrylamide/urea gel comparing extension products generated with base analogs. The effects of replacing dGTP with c^{7}GTP (dc^{7}GTP) or dITP are shown in sequencing reactions performed on M13mp18 single-stranded DNA or on a partially palindromic clone, EK9. Reaction conditions and electrophoresis were as described. Lanes are loaded G, A, T, C. Lines between the EK9 dGTP and c^{7}GTP reaction sets align the same positions upstream and downstream of the compressed region. The bracket indicates the limits of the palindrome. The correct sequence of the region is 5′-CCATGTGACCCTGCCCGACTTCGACGGGAATTCCCGTCGAAGTCGGGCAGGGTCACCATA-3′. The complementary bases are underlined and the bases compressed in the dGTP reactions are boldface.

The gel presented in Fig. 5 compares the DNA sequence obtained with *Taq* DNA polymerase using either an asymmetric PCR-generated template, or the same DNA insert cloned in M13mp18 as template. The resulting sequence ladders show the clarity and uniformity of signal characteristic of *Taq*-generated sequences. Any degradation of enzyme or dNTPs that may have occurred during the PCR thermal cycling did not seem to affect the generation of clean sequence data. Synthesis of single-stranded DNA template during 35 cycles of PCR was largely independent of the initial DNA concentration. Asymmetric PCRs performed with 0.1 to 100 ng of M13mp10 single-stranded DNA, or 10 μl of an M13 phage plaque picked directly into 100 μl of water, sequenced equivalently.

DISCUSSION

In this paper, we present convenient and efficient protocols for sequencing with *Taq* DNA polymerase. This enzyme worked equally well with either 5′-labeled primers or by incorporation of label in a two-step reaction protocol. Both approaches generated DNA sequencing ladders that were characteristically free of background bands or noticeable enzyme idiosyncrasies, were uniform in intensity, and were readable over long distances. These protocols also gave very clean results with alkali-denatured double-stranded DNA templates (data not shown).

Our results suggest that *Taq* DNA polymerase has advantages for many sequencing applications. Sequencing results obtained with the *Taq* enzyme were clearly superior to either Klenow or avian myeloblastosis virus reverse transcriptase and were often better (on G+C-rich templates) than results

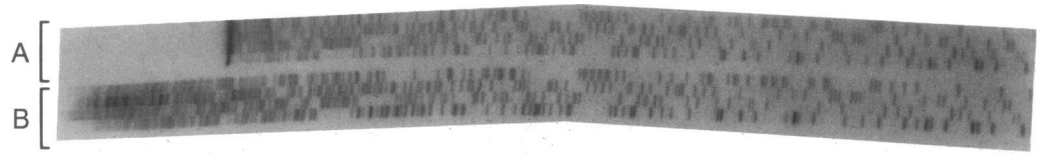

FIG. 5. Autoradiograph of a polyacrylamide/urea gel comparing the extension products from an M13-based single-stranded template (*A*) and an asymmetric PCR template of the same sequence (*B*). The sequencing of the M13 clone was carried out as described with a ^{32}P-labeled primer. The asymmetric amplification, DNA sequencing, and electrophoresis were performed as described. Reaction sets were loaded G, A, T, C.

obtained by using modified T7 DNA polymerase (data not shown). Unlike any of these polymerases, *Taq* DNA polymerase works over a broad temperature optimum centered around 75°C. Regions of DNA structure (hairpins) are commonly encountered that strongly hinder polymerases and cause premature termination bands across all four sequencing lanes. The ability of *Taq* DNA polymerase to operate at high temperature and low salt allows heat destabilization of hairpins during the sequencing reaction, permitting the enzyme to read through such structures. The concomitant use of a structure-destabilizing dGTP analog, c^{7}GTP, yields sequencing products from G+C-rich templates that are fully resolved upon electrophoresis.

We attribute the absence of background bands and the uniformity of signal to our observations that *Taq* DNA polymerase is highly processive, has a high turnover number, and has very little or no proofreading activity. Such properties of the enzyme are ideal for sequencing because they reduce pausing and premature termination at sequences with secondary structure and diminish discrimination against dideoxy nucleotide analogs (5).

Under certain circumstances, the absence of significant *Taq*-associated 3′-exonuclease activity causes chain-termination due to misincorporated bases. The misincorporation rate is enhanced when one or more of the dNTPs are well below K_m and/or when the concentration of one dNTP is very low relative to the other dNTPs. Because dITP base pairs promiscuously, we observed frequent chain termination near regions of high secondary structure with dITP and do not recommend it for sequencing with *Taq*. We do not observe misincorporation of bases if the concentration of all four dNTPs is similar and/or if they are present at $\geq$10 μM each. Sequence analysis of cloned PCR products generated with *Taq* DNA polymerase suggests that the fidelity of PCR using 50–200 μM each dNTP is quite respectable (approximately one mistake in 4000–5000 base pairs sequenced after 35 cycles of PCR and cloning of the products; unpublished results) and is comparable with that observed using other DNA polymerases for PCR. In addition, our data suggest that misincorporation errors that occur during the PCR promote chain termination (presumably because of significantly higher K_m values for mismatch extension), thus attenuating amplification of defective molecules and maintaining fidelity.

Several methods, with varying degrees of speed and reliability, have been published for sequencing PCR products (10–13). The remarkable sequencing properties demonstrated by *Taq* and its use in PCRs suggest it as the ideal enzyme for directly analyzing PCR products. Here, the protocols for sequencing with *Taq* were successfully used to sequence asymmetric PCR products without prior purification, and the results compared favorably with sequencing the same insert using M13 single-stranded DNA template.

While this approach has been developed for sequencing inserts in M13 or pUC-based vectors, it is applicable to direct sequencing of clones in λ phage and other cloning vectors. Some variability in the single-stranded DNA yield of the PCR has been observed with different primer pairs and ratios (13), and the reaction conditions for each amplification system will need to be adjusted for optimal results. Some investigators have increased the homogeneity of their PCR products from genomic DNA by electrophoretic separation and reamplification of eluate from a selected gel slice (19). Our direct sequencing method is easily applied to this "secondary" PCR. Direct sequencing of PCR products from DNA by any method produces a "consensus" sequence; those bases that occur at a given position in the majority of the molecules will be the most visible on an autoradiograph and any low-frequency errors will be undetectable. In a coupled experiment of this kind, the resulting sequence data will be only as clean as the amplified product, and heterogeneous products will naturally produce mixed ladders.

The ability to couple template preparation by asymmetric PCR with direct sequencing by using the *Taq* enzyme opens the possibility of automating both DNA template preparation and the performance of the sequencing reactions in a manner that should be compatible with current DNA sequencing instruments.

We thank Susanne Stoffel for providing *Taq* DNA polymerase; Corey Levenson, Laurie Goda, and Dragan Spasic for preparation of synthetic oligonucleotide primers; Ulf Gyllensten for sharing data prior to publication; members of the Cetus PCR Group for their continued interest in this work; and Eric Ladner and Sharon Nilson for artwork.

1. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
2. Yanisch-Perron, C., Vieira, J. & Messing, J. (1985) *Gene* **33,** 103–119.
3. Mills, D. R. & Kramer, F. R. (1979) *Proc. Natl. Acad. Sci. USA* **76,** 2232–2235.
4. Barr, P. J., Thayer, R. M., Laybourn, P., Najarian, R. C., Seela, F. & Tolan, D. R. (1986) *BioTechniques* **4,** 428–432.
5. Tabor, S. & Richardson, C. C. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 4767–4771.
6. Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., Heiner, C., Kent, S. B. H. & Hood, L. E. (1986) *Nature (London)* **321,** 674–679.
7. Prober, J. M., Trainor, G. L., Dam, R. J., Hobbs, F. W., Robertson, C. W., Zagursky, R. J., Cocuzza, A. J., Jensen, M. A. & Baumeister, K. (1987) *Science* **238,** 336–341.
8. Ansorge, W., Sproat, B., Stegemann, J., Schwager, C. & Zenke, M. (1987) *Nucleic Acids Res.* **15,** 4593–4602.
9. Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. & Arnheim, N. (1985) *Science* **230,** 1350–1354.
10. Engelke, D. R., Hoener, P. A. & Collins, F. S. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 544–548.
11. Wong, C., Dowling, C. E., Saiki, R. K., Higuchi, R. G., Erlich, H. A. & Kazazian, H. H. (1987) *Nature (London)* **330,** 384–386.
12. Stoflet, E. S., Koeberl, D. D., Sarkar, G. & Sommer, S. S. (1988) *Science* **239,** 491–494.
13. Gyllensten, U. B. & Erlich, H. A. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 7652–7656.
14. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239,** 487–491.
15. Maxam, A. & Gilbert, W. (1980) *Methods Enzymol.* **65,** 499–560.
16. Zinder, N. D. & Boeke, J. D. (1982) *Gene* **19,** 1–10.
17. Biggin, M. D., Gibson, T. J. & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 3963–3965.
18. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
19. Higuchi, R., von Beroldingen, C. H., Sensabaugh, G. F. & Erlich, H. A. (1988) *Nature (London)* **332,** 543–546.