THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 272, No. 17, Issue of April 25, pp. 11228–11235, 1997

Printed in U.S.A.

Low Fidelity Mutants in the O-Helix of *Thermus aquaticus* DNA Polymerase I*

(Received for publication, December 26, 1996, and in revised form, February 19, 1997)

Motoshi Suzuki‡, Amy K. Avicola‡, Leroy Hood§, and Lawrence A. Loeb‡¶

From ‡The Joseph Gottstein Memorial Cancer Research Laboratory, Department of Pathology, Box 357705, and the §Department of Molecular Biotechnology, University of Washington, Seattle, Washington 98195-7705

We screened 67 mutants in the O-helix of *Thermus aquaticus* (*Taq*) DNA polymerase I (pol I) for altered fidelity of DNA synthesis. These mutants were obtained (Suzuki, M., Baskin, D., Hood, L., and Loeb, L. A. (1996) *Proc. Natl. Acad. Sci. U. S. A.* 93, 9670–9675) by substituting an oligonucleotide containing random sequences for codons 659–671, and selecting for complementation of a growth defect in *Escherichia coli* caused by temperature-sensitive host pol I. Thirteen mutants decreased fidelity in a screen that employed primer extension reactions lacking one of four complementary deoxynucleoside triphosphates (dNTPs). Three mutants were purified and exhibited 29–68% of wild-type specific activity. Homogeneous polymerases A661E, A661P, and T664R extended primers further than the wild-type, synthesizing past template nucleotides for which the complementary dNTP was absent. The data indicate that both misinsertion of incorrect nucleotides and extension of mispaired primer termini were increased. In a *lacZα* forward mutation assay, A661E and T664R yielded mutation frequencies at least 7- and 25-fold greater, respectively, than that of the wild-type polymerase. These findings emphasize the importance of the O-helix in substrate recognition and are compatible with a role for pyrophosphate release in enhancing fidelity of DNA synthesis.

Thermus aquaticus (*Taq*) DNA polymerase I (pol I)[1] is used extensively in the amplification of DNA by the polymerase chain reaction. Based on its primary sequence, crystal structure, and catalytic mechanism, *Taq* pol I is classified in the same family as *Escherichia coli* pol I (1–3). The amino acid sequences of *Taq* pol I and *E. coli* pol I are 38% homologous, both enzymes have 5′-3′ exonuclease domains (4, 5), and the crystal structures of the polymerase domains are nearly identical (5, 6). The structure of the vestigial 3′-5′ exonuclease domain of *Taq* pol I is very different, however; unlike *E. coli* pol I, purified *Taq* pol I exhibits no proofreading activity (7).

The O-helix in *E. coli* pol I is constituted of 15 amino acids which form part of the DNA binding cleft (8). Crystallographic, biochemical, and mutagenesis studies indicate that four conserved residues in the O-helix which face toward the cleft interact with either the incoming dNTP or the template-primer. Considerable evidence indicates that in *E. coli* pol I, Arg-754 and Lys-758 interact with the phosphate groups in the incoming dNTP, that Phe-762 interacts with the deoxyribose moiety in the dNTP, and that Tyr-766 binds to the template-primer (9–11). The nearly identical structures of the polymerase domains of *Taq* pol I and *E. coli* pol I (6, 7) suggest that the corresponding amino acids in *Taq* pol I have analogous functions. By using random sequence mutagenesis together with genetic complementation, we recently observed that the O-helix of *Taq* pol I contains four amino acids which are immutable (or nearly immutable) and are apparently essential for catalysis *in vivo* (12). These residues, Arg-659, Lys-663, Phe-667, and Tyr-671, correspond to the four above mentioned residues in *E. coli* pol I. Importantly, Phe-667 in *Taq* pol I and the corresponding Phe-762 in *E. coli* pol I are critical for substrate discrimination (13). Substitution of Tyr for Phe results in increased incorporation of dideoxynucleoside triphosphates. Analysis of the effects of this and other mutations in the O-helix on the incorporation of non-complementary nucleotides may be important to understanding substrate recognition by DNA polymerases.

The fidelity of DNA polymerases is central to the accurate transmission of genetic information. This fidelity is conferred primarily by the accuracy of polymerization, with an additional contribution from proofreading for enzymes which possess an editing 3′-5′ exonuclease (14). In the case of *Taq* pol I, which lacks proofreading activity, the error rate during a single round of replication has been estimated as one misincorporation for every 9000 nucleotides polymerized (15). This estimate is in approximate agreement with studies on the product of polymerase chain reactions obtained after multiple rounds of DNA synthesis in some (16) but not all studies (17). The frequency of misincorporation by *Taq* pol I is greater than that of the Klenow fragment of *E. coli* pol I, even when assayed in 1 mM dNTP (conditions that substantially reduce the Klenow-associated 3′-5′ exonucleolytic activity). This comparison implies that despite the similarity in structure of the two enzymes, there are subtle differences that affect fidelity. Interestingly, the fidelity of *Taq* pol I is markedly influenced by reaction conditions; both base substitution and frameshift error rates of less than 1×10^{-5} are observed at pH 5–6 (18).

The fidelity of polymerization depends on the ability of the polymerase site to discriminate correct from incorrect substrates, *i.e.* complementary from non-complementary nucleoside triphosphates (19) and matched from mismatched primer termini (20). A molecular understanding of polymerization fidelity will require identification of the amino acids which participate in substrate discrimination and elucidation of the role(s) played by each. As recently pointed out, lack of mutants with altered accuracy, particularly for DNA polymerases whose three-dimensional structure is known, is a limiting factor in

* This work was supported by National Science Foundation Science and Technology Center for Molecular Biotechnology Grant BIR9214821 (to L. H.) and a National Cancer Institute Outstanding Investigator Grant R35-CA-39903 (to L. A. L.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¶ To whom correspondence should be addressed: The Joseph Gottstein Memorial Cancer Research Laboratory, Dept. of Pathology, Box 357705, University of Washington, Seattle, WA 98195-7705. Tel: 206-543-6015; Fax: 206-543-3967; E-mail: laloeb@u.washington.edu.

[1] The abbreviation used is: pol I, DNA polymerase I.

obtaining a molecular description of fidelity (9). Some accuracy mutants have been obtained by site-directed mutagenesis (9). However, this approach does not readily permit examination of large numbers of amino acid replacements and frequently yields mutants with markedly reduced catalytic efficiency.

In the present work, we illustrate a different approach to obtaining fidelity mutants of *Taq* pol I. We have previously used random sequence mutagenesis together with genetic complementation to create a collection of catalytically active mutants with amino acid replacements in the O-helix (12), and we have now screened some of these mutants for reduced fidelity to examine the effects of the mutations on accuracy. The importance of the O-helix in interactions with incoming dNTPs and with the template-primer suggests that amino acid substitutions within this segment that alter fidelity are likely to do so by altering these interactions. Moreover, because our mutant selection depends on complementation of a temperature-sensitive phenotype *in vivo*, *Taq* pol I derivatives with near wild type activity are obtained (21), and the effects of mutation on fidelity are therefore observed in the absence of major impairments of catalysis.

EXPERIMENTAL PROCEDURES

Polymerase Sources, Strains, and Oligonucleotides—-Wild-type *Taq* pol I and its mutant derivatives were selected by genetic complementation as previously reported (12). *E. coli* DH5α [F^-, *ϕ80dlacZΔM15, Δ(lacZYA-argF)U169, deoR, recA1, endA1, phoA, hsdR17*($r_k^-m_k^+$)*, supE44*, λ^-, *thi-1, gyrA96, relA1*] was used to express *Taq* pol I protein. *E. coli* MC1061 [*hsdR, mcrR, araD,* (139Δ^-*ara*ABC$^-$leu)*, Δ$^-$7679lacX74, galU, galK, rspL, thi*$^-$] was used for transfection in the forward mutation assay, and transfectants were plated on the indicator strain *E. coli* CSH50 [Δ(*pro*-Blac)/F′+raD36,*thi*$^-$*, ara*$^-$*,proAB, lacI*^{q}*Z*$^-$Δ*M15*]. Oligonucleotides were synthesized and purified by Operon Technologies Inc. (Alameda, CA).

Preparation of Extracts—A single colony of *E. coli* DH5α carrying wild-type or mutant *Taq* pol I was inoculated into 40 ml of 2xYT (16 g/liter tryptone, 10 g/liter yeast extract, 5 g/liter NaCl, pH 7.3) containing 30 mg/liter chloramphenicol. After incubation at 37 °C overnight with vigorous shaking, an equal amount of fresh medium with 0.5 mM isopropyl-β-D-thiogalactoside was added, and incubation was continued for 4 h. Cells were harvested, washed once with TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA), and suspended in 100 μl of buffer A (50 mM Tris-HCl, pH 8.0, 2.4 mM phenylmethylsulfonyl fluoride, 1 mM dithiothreitol, 0.5 mg/liter leupeptin, 1 mM EDTA, 250 mM KCl). Bacteria were lysed by incubating with lysozyme (0.2 mg/ml) at 0 °C for 2 h. The lysate was centrifuged at 15,000 rpm (Sorvall, SA-600 rotor) for 15 min, and the supernatant solution was incubated at 72 °C for 20 min. Insoluble material was removed by centrifugation.

Purification of Polymerases—Purification was carried out according to Lawyer *et al.* (22) with modifications. A single colony of *E. coli* DH5α carrying wild-type or mutant *Taq* pol I was inoculated into 10 ml of 2xYT. Two ml of the inoculum was immediately added to each of 5 bottles containing 1 liter of 2xYT with 30 mg/liter chloramphenicol. After overnight incubation at 37 °C with vigorous shaking, 1 liter of 2xYT containing 30 mg/liter chloramphenicol and 0.5 mM isopropyl-β-D-thiogalactoside was added, and incubation was continued for 4 h. Cells were harvested, washed once with TE buffer, and suspended in 100 ml of buffer A. Bacteria were lysed by incubating with lysozyme (0.2 mg/ml) at 0 °C for 2 h and then sonicating on ice for 45 s by using a micro-tip probe (Sonifier, Branson Sonic Power, Danbury, CT).

The lysate was centrifuged at 15,000 rpm (Sorvall, SA-600 rotor) for 15 min, and the supernatant solution was incubated at 72 °C for 20 min. Insoluble material was removed by centrifugation. Ammonium sulfate (0.2 M) and Polymin P (0.6%) were added and the suspension was held on ice for 1 h. After removal of the precipitate by centrifugation and filtration through a Costar 8310 filter, the filtrate was applied to a 3 × 8-cm phenyl-Sepharose HP (Pharmacia Biotech Inc.) column equilibrated with buffer A containing 0.2 M ammonium sulfate and 0.01% Triton X-100. The column was washed with the same buffer (300 ml) and activity was eluted with buffer B (TE buffer containing 0.01% Triton X-100 and 50 mM KCl). The eluate (100 ml) was dialyzed overnight against 4 liters of buffer B and loaded onto a 0.8 × 8-cm heparin-Sepharose CL-6B (Pharmacia Biotech Inc.) column equilibrated with buffer B. After washing with buffer B (50 ml), activity was eluted in a 30-ml linear gradient of 50–500 mM KCl in TE buffer containing 0.01% Triton X-100. Active fractions were collected, dialyzed against 50 mM Tris-HCl (pH 8.0) containing 50 mM KCl and 50% glycerol, and stored at −80 °C.

Polymerase Assay—Enzyme was incubated at 72 °C for 5 min in 50 mM Tris-HCl (pH 8.0), 2 mM $MgCl_2$, 100 μM each dATP, dGTP, dCTP, and dTTP, 0.2 μCi of [^{3}H]dATP, and 200 μg/ml activated calf thymus DNA. Incorporation of radioactivity into an acid-insoluble product was measured according to Battula and Loeb (23). One unit represents incorporation of 10 nmol of dNMP in 1 h, corresponding to 0.1 unit as defined by Perkin-Elmer.

Primer Extension Assay—The 14-mer primer 5′-CGCGCCGAATTCCC was ^{32}P-labeled at the 5′ end by incubation with [γ-^{32}P]ATP and T4 polynucleotide kinase and annealed to an equimolar amount of the template 46-mer 5′-GCGCGGAAGCTTGGCTGCAGAATATTGCTAGCGGGAATTCGGCGCG. Crude extracts containing 0.3–1 unit of mutant or wild-type *Taq* pol I were incubated at 45 °C for 60 min in 50 mM Tris-HCl (pH 8.0), 2 mM $MgCl_2$, 50 mM KCl, 20 μM each dATP, dGTP, dCTP, and TTP, and 1.4 ng of the annealed template-primer. A set of four additional reactions, each lacking a different dNTP, was carried out for each polymerase. Purified enzyme (1 unit) was incubated for the times indicated under the same conditions as for crude extracts. After electrophoresis in a 14% polyacrylamide gel containing 8 M urea, reaction products were analyzed by autoradiography. Extension was quantified by using an NIH imaging program.[2]

Forward Mutation Assay and DNA Sequence Analysis—The noncoding strand of the *lacZ*α gene contained in 200 ng of gapped M13mp2 DNA (24) was copied by using 5 units of wild-type or mutant *Taq* pol I in a reaction mixture containing 50 mM Tris-HCl (pH 8.0), 2 mM $MgCl_2$, 20 μM each dNTP, and 50 mM KCl. After incubation at 72 °C for 5 min, the DNA was transfected and the plaques were scored (15). The nucleotide sequence of mutants was determined by using a Thermo Sequenase™ cycle sequencing kit (Amersham LIFE SCIENCE, Cleveland, OH).

RESULTS

Screening for Fidelity Mutants—We have previously generated a large collection of functional mutants of *Taq* pol I (12). Extracts of *E. coli* expressing these mutated *Taq* pol I were used to screen for mutations that alter fidelity. The mutants were created by substituting random nucleotides for the 13 codons specifying amino acids 659–671 in the O-helix of plasmid-borne *Taq* pol I. Functional mutations were then selected based on their ability to complement the growth defect of an *E. coli* strain carrying a temperature-sensitive host pol I (12). Sequence analysis of the selected, functional mutants revealed that 4 of the 13 mutagenized amino acids, Arg-659, Lys-663, Phe-667, and Tyr-671, tolerated no substitutions or only a few conservative substitutions, in accord with biochemical and crystallographic evidence that Arg-659 and Lys-663 participate in the formation of a metal-dNTP binding pocket (10, 11). The nine remaining residues were tolerant of several or many replacements. To explore the contribution of individual amino acids in the O-helix to replicational fidelity, we screened functional mutants in a primer elongation assay to identify substitutions that affect ability to utilize non-complementary dNTPs.

We screened 67 of the 75 active mutants previously sequenced, including all 38 with single amino acid substitutions (described in Fig. 3*A* of Ref. 12). Plasmids encoding the mutant polymerases were cloned, purified, and grown in *E. coli*, and host cells were analyzed for expression of *Taq* pol I by measuring the activity of crude extracts (see "Experimental Procedures"). *E. coli* DNA polymerases and nucleases were inactivated by heating at 72 °C for 20 min. The ability of heat-treated extracts to elongate primers in the absence of a complete complement of four dNTPs was then determined using a set of five reactions. One reaction contained all four complementary triphosphates while each of the others lacked a different dNTP ("minus conditions"). Elongation in the minus reactions is lim-

[2] See http//www.nih.gov/.

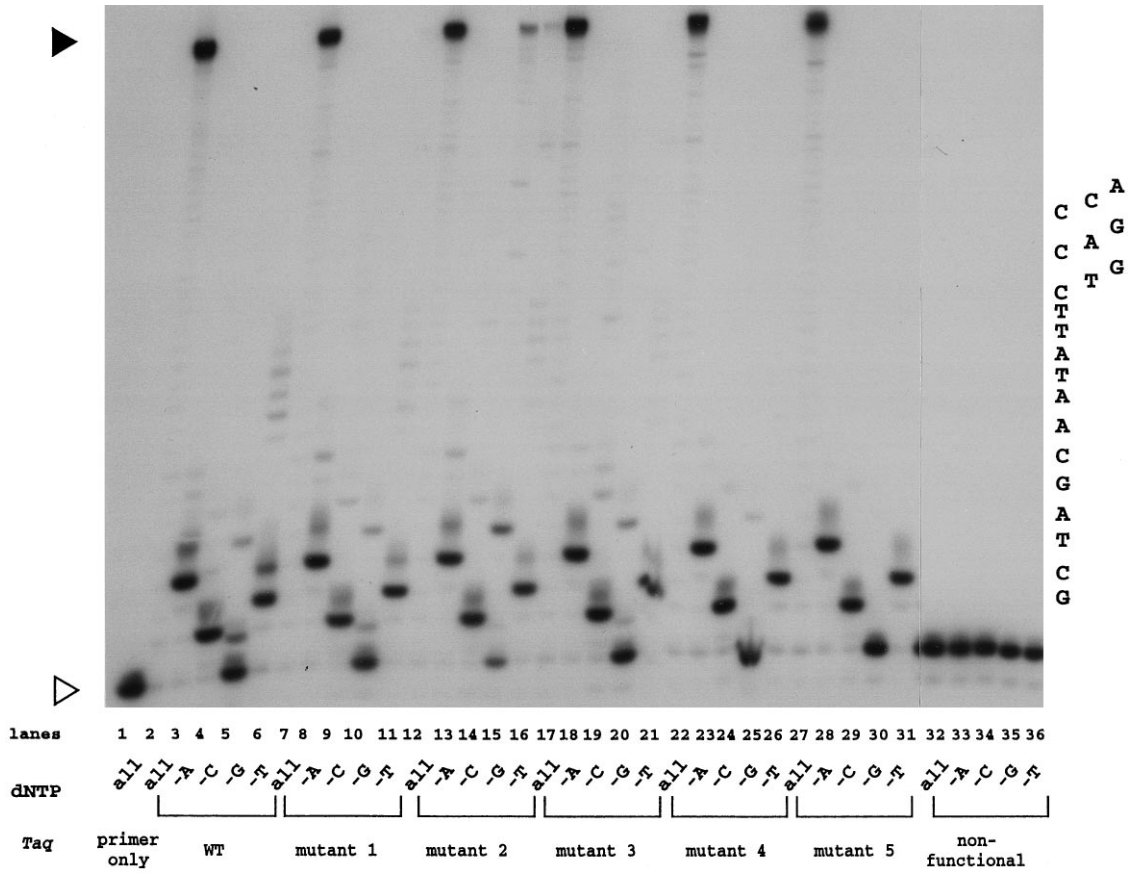

FIG. 1. **Screening for altered fidelity of *Taq* pol I mutants in a primer extension assay.** Heat-treated extracts of *E. coli* expressing wild-type *Taq* pol I or one of its mutant derivatives were used to catalyze primer elongation in the absence of a full set of complementary dNTPs. For each pol I, ^{32}P-labeled primer was extended in a control reaction with all four dNTP and in a set of four additional reactions in which one of the dNTPs was omitted, as indicated in the figure. Elongation patterns for seven extracts are shown to exemplify results of the screen, including wild-type *Taq* pol I (*lanes 2–6*) and a catalytically non-functional derivative (*lanes 32–36*) in which the O-helix was replaced with another sequence (12). M1, A661E, I665T; M2, A661P, N666I; M3, T664N, L670I; M4, F667Y; M5, R660K. The non-functional mutant was displayed in a separate portion of the gel and spliced with the others. *Open* and *filled triangles* represent the positions of the primer and the fully extended product, respectively. Positions of nucleotides complementary to the template are shown at the *right*.

ited by the rate of misincorporation at template positions complementary to the missing dNTP.

Fig. 1 illustrates representative results obtained for wild-type *Taq* pol I and for mutants with elongation patterns that differed from wild-type. In the presence of all four dNTPs, every extract examined extended more than 90% of the hybridized primer to a product of length similar to that of the template. In the minus reactions, wild-type *Taq* pol I (*lanes 2–6*) extended 48–60% of the primer up to, but not opposite, the first template position complementary to the missing dNTP. The remaining primer was terminated opposite the missing dNTP, presumably by incorporation of a single non-complementary nucleotide, or was terminated further downstream, presumably by extension of the mispaired primer terminus. A variety of elongation patterns was observed for the 67 mutants. Thirteen mutants extended more of the primer and/or synthesized a greater proportion of longer products than the wild-type enzyme in three or four of the minus reactions (*e.g.* mutant 2 (*lanes 12–16*) which formed full-length products in reactions lacking dGTP (*lane 15*) or TTP (*lane 16*)). This increased extension presumably reflects increased incorporation and/or extension of non-complementary nucleotides. Other mutants (*e.g.* mutant 5, *lanes 27–31*) extended less of the primer or synthesized shorter products than the wild-type enzyme. In several cases, different amino acid substitutions at the same position either increased or decreased extension in comparable minus reactions.

Fig. 2 contains a compilation of amino acid replacements in

	659				663				667			671
WT :	R	R	A	A	K	T	I	N	F	G	V	L	Y
29 :			E										
36 :						P					I		
40 :			P										
45 :						P							
53 :						N							
130 :		P										T	
156 :			S	G		S		I					
175 :		W				K							
206 :						R							
240 :				G		N							
247 :			G					I					
248 :								V					
306 :			P					I					

FIG. 2. **Low fidelity mutants of *Taq* pol I identified in the primer extension screen.** Amino acid replacements are listed for mutants that elongated primers to a greater extent than wild-type *Taq* pol I in at least three of the "minus" reactions in the primer extension screen (see Fig. 1). Amino acid substitutions at positions Arg-659 through Tyr-671 are shown below the wild-type (*WT*) sequence. Mutant designations at the left are the same as described previously (12). Note that Arg-659, Lys-663, Phe-667, and Tyr-671, for which no replacements are listed, were previously found to be immutable or nearly immutable (12).

the 13 mutants that displayed increased extension in at least three of the minus reactions. With the exception of Gly-668, one or more substitutions that putatively reduce the accuracy of DNA synthesis were observed for each of the 9 non-conserved amino acids. Eleven mutants harbored substitutions at either Ala-661 or Thr-664, including several single mutants. This initial screen with crude extracts suggested that a large number of changes are permitted in the O-helix that do not reduce

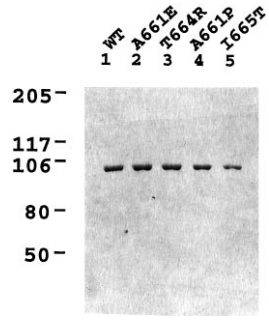

FIG. 3. **SDS-polyacrylamide gel electrophoresis of purified *Taq* DNA polymerases.** Purified wild-type or mutant pol I (0.5 μg) was heated at 72 °C for 10 min in buffer containing 1% SDS, 50 mM Tris-HCl (pH 6.8), 10% glycerol, 1% 2-mercaptoethanol. Samples were subjected to electrophoresis in a 7.5% SDS-polyacrylamide gel (53) and the gel was stained with Coomassie Brilliant Blue R. The positions of size markers are indicated at the *left*.

the ability of *Taq* pol I to complement the growth defect of *recA718 pol A12* (12). Many of the substitutions in the O-helix that do not reduce the ability of *Taq* pol I to carry out functional complementation reduce the fidelity of DNA synthesis *in vitro*.

Purification of Mutant Polymerases—To demonstrate that the reduction in fidelity exhibited by crude extracts is due to the mutant *Taq* pol I, we purified the wild-type enzyme as well as the polymerase from the three single mutants Ala-661 → Glu (A661E), Ala-661 → Pro (A661P), and Thr-664 → Arg (T664R). As a control, we also purified enzyme from the single mutant Ile-665 → Thr (I665T), which the screen suggested would have no alteration in fidelity. As shown in Fig. 3, each polymerase was purified to apparent homogeneity and had mobility corresponding to the expected molecular weight of 99,000. The specific activities of the purified enzymes, determined in assays with activated calf thymus DNA, were 66,000, 45,000, 19,000, 23,000, and 30,000 units/mg protein for the wild-type, A661E, A661P, T664R, and I665T, respectively. That the mutated enzymes retain at least 29% of wild-type activity *in vitro* is in accord with their ability to complement the growth defect caused in *E. coli* by temperature-sensitive host pol I, and ensures that analysis of fidelity will not be complicated by major impairments of catalytic efficiency.

Primer Elongation by Purified Polymerases—Primer extension assays were carried out with the homogenous mutant polymerases. Fig. 4*A* shows reactions incubated for 5 min at 42 °C. Wild-type *Taq* pol I extended most of the primer to one nucleotide before the template position opposite the missing complementary dNTP. Only about 30% of the primers were elongated further. In reactions containing equivalent activity, the mutant polymerases A661E, T664R, and A661P extended a larger proportion of the primers past the sites where the wild-type polymerase ceased synthesis. The control enzyme I665T yielded an elongation pattern similar to that of the wild-type enzyme. Elongation reactions with the three polymerases were also carried out for 60 min (Fig. 4*B*). Again, A661E and T664R synthesized a greater proportion of longer products than obtained with the wild-type and I665T polymerases. Notably, A661E, T664R, and A661P synthesized longer products in 5 min (Fig. 4*A*) than the wild-type did in 60 min (Fig. 4*B*).

To further document the reduced fidelity illustrated in Fig. 4, *A* and *B*, we carried out a time course of primer elongation. Fig. 4*C* shows results for the minus dCTP reaction catalyzed by the wild-type and T664R polymerases. Wild-type *Taq* pol I extended 9% of the primers past the first deoxyguanosine template residue within the 60-min incubation period, but elongation past the second deoxyguanosine was not detected (*lane 6*). In the same interval, T664R extended 93% of the primer past the first template deoxyguanosine, and elongation proceeded past as many as five template deoxyguanosines (*lane 12*). Importantly, a comparable proportion of primers was extended at all time points, despite the striking difference in the length of the products. These time course data indicate that greater elongation reflects increased ability to utilize non-complementary substrates and primer termini, rather than a putative difference in the amount of activity present. Such a difference might be hypothesized, because activity of the individual polymerases was based on assays utilizing a gapped DNA template-primer (*i.e.* activated calf thymus DNA).

Interestingly, very little primer was extended to a length greater than 21 nucleotides in any of the minus dATP reactions (Fig. 4). Cessation of elongation may be due to the run of two Ts in the template at positions corresponding to a 21- and 22-mer. Similarly, extension may have ceased at the template Gs corresponding to 33- and 34-mer products in minus dCTP reactions, and at template As corresponding to 25- and 26-mer products in minus dTTP reactions. Apparently, extension of primers terminating with two non-complementary nucleotides is much less efficient than extension of primers terminating with a single mismatched nucleotide. These results suggest that extension of mismatched primer termini is catalyzed by a template-dependent mechanism characteristic of DNA polymerases, rather than by a template-independent mechanism such as that of terminal deoxynucleotidyl transferase.

Forward Mutation Assay—To quantitate the fidelity of DNA synthesis by the purified polymerases, we measured the frequency of mutations produced by copying a biologically active template *in vitro* (25). The target sequence was the *lacZ*α gene located within a single-stranded region in gapped circular double-stranded M13mp2 DNA (26). The gapped segment was filled by synthesis with the wild-type or mutant enzymes. The double-stranded circular product was transfected into *E. coli*, and the mutation frequency was determined by scoring white and pale blue mutant plaques. A comparison of the specific activities and mutation frequencies of the purified enzymes is presented in Table I. After synthesis by wild-type *Taq* pol I, the mutation frequency was not greater than that of the uncopied control. Synthesis by A661E and T664R gave rise to mutation frequencies more than 7- and 25-fold greater, respectively, than that of the wild-type polymerase.

A sample of independent, randomly chosen mutants produced by T664R was characterized by DNA sequence analysis. Both base substitutions and frameshifts were found throughout the targeted *lacZ*α gene and its regulatory sequence. Of the 64 independent plaques, 57 had mutations in the target. Other mutations presumably occurred outside the target region. Some had more than one base substitution and a total of 66 mutations were observed (Fig. 5). Among them, 61 were base substitutions. Transitions (38/61) were more frequent than transversions (23/61). T → C transitions accounted for 31 of 61 base substitutions, while T → A (9/61), A → T (8/61), and G → A (5/61) substitutions were less frequent. This base substitution spectrum is essentially the same as that reported for wild-type *Taq* pol I (27). From these data, the base substitution fidelity of T664R can be calculated according to Tindall and Kunkel (27) as 8.6×10^{-4} or 1 error per 1200 nucleotides. On the basis of the five frameshift mutants detected, the frameshift error can be calculated as 4.9×10^{-5} or 1 error per 20,000 nucleotides.

DISCUSSION

We have utilized random sequence mutagenesis and a genetic complementation system to obtain a large series of catalytically hardy mutants in the O-helix of *Taq* pol I (12). In the present work, we screened preparations of 67 of the mutant enzymes for alterations in the fidelity of DNA synthesis by

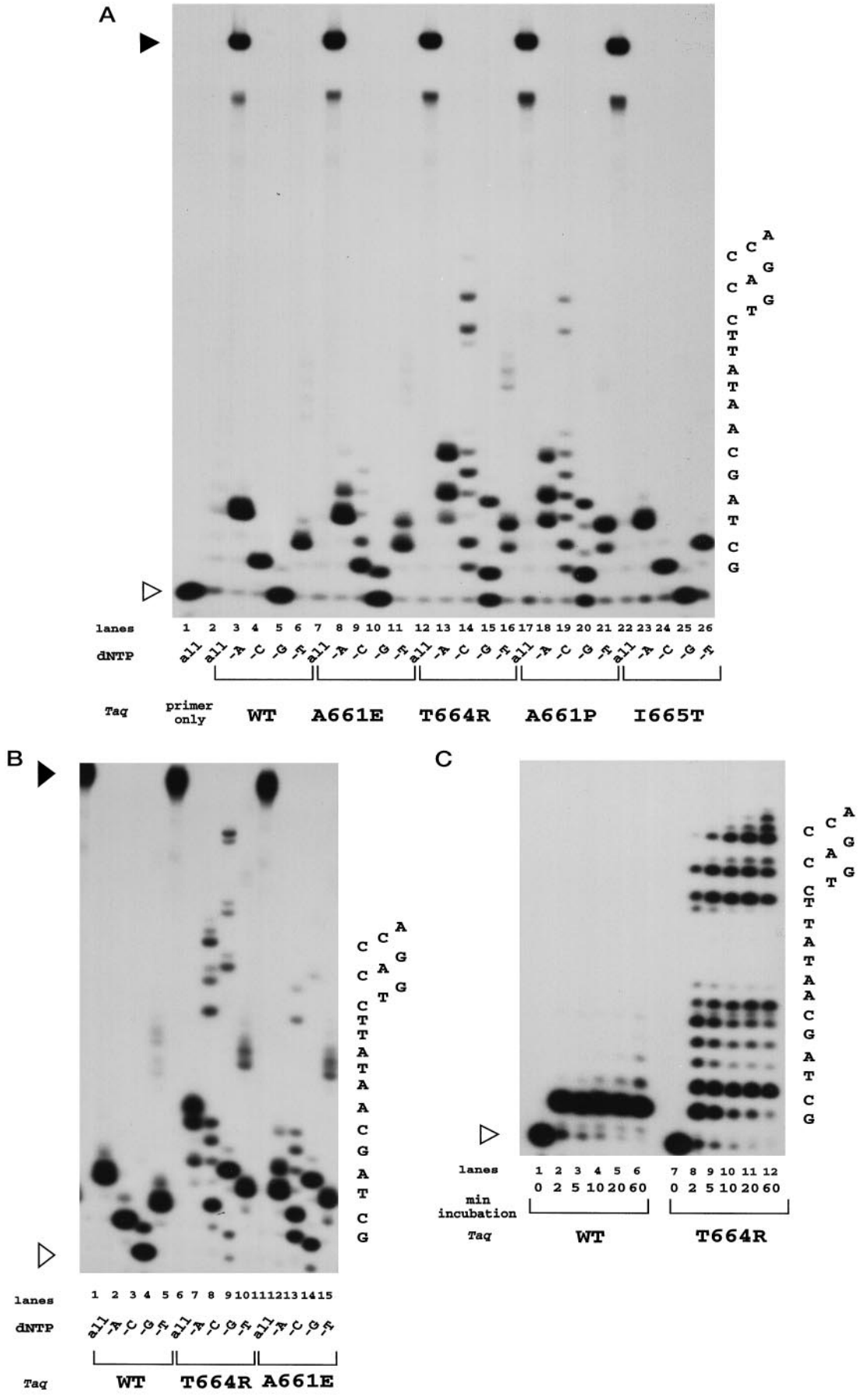

FIG. 4. **Primer extension by purified *Taq* DNA polymerases.** Primer elongation was catalyzed by purified wild-type *Taq* pol I or one of the mutant derivatives A661E, A661P, T664R, or I665T. *A,* incubation was for 5 min at 45 °C in reactions containing all four complementary dNTPs

TABLE I

Mutation frequency in the lacZα forward mutation assay

Mutation frequencies were measured after *in vitro* copying of M13mp2 DNA by purified mutant and wild type *Taq* pol Is. Both white and pale blue plaques were scored as mutants. Each mutant plaque was verified by replating with an equal number of wild-type M13mp2. In this experiment, the mutation frequency observed for uncopied DNA (no enzyme control) was 3.5×10^{-3}, greater than that of DNA copied with the wild-type, and therefore not subtracted from any of the mutation frequencies.

Taq pol I	Specific activity	Plaques scored		Mutation frequency
		Total	Mutant	
	units/mg			$\times 10^{-3}$
WT	66,000	8,637	22	2.5
A661E	45,000	6,782	116	17.1
T664R	23,000	5,148	324	62.9

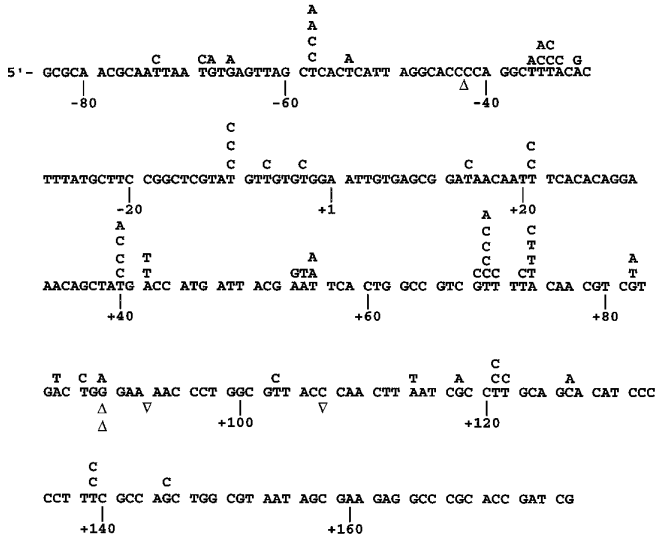

FIG. 5. **Spectrum of single base changes generated in the *lacZα* forward mutation assay by the *Taq* pol I mutant T664R.** The DNA sequence of the viral strand of wild-type M13mp2 is shown, with base substitutions catalyzed by T664R indicated above the wild-type sequence and frameshifts indicated below. For frameshifts, deletion of a base is indicated by a Δ and addition of a base is indicated by a ▽. Since frameshifts occurred in a run of three or four of the same nucleotide, the symbol is centered under the run. The gap in M13mp2 starts at the deoxyguanosine residue at position +145.

measuring elongation of DNA primers in reactions lacking one of four complementary dNTPs. Thirteen of the mutants displayed reduced fidelity compared with that of the wild-type polymerase in this primer extension assay. Three low fidelity enzymes were purified to homogeneity, found to have specific activities that were at least 29% of wild-type, and re-examined in the primer extension assay. Relative to the wild-type enzyme, the purified mutant enzymes extended a greater proportion of primers to a position opposite that of a template residue for which there was no complementary dNTP, indicating that they discriminate less strictly between a correct and an incorrect incoming nucleotide (19, 28). In addition, the mutant enzymes synthesized longer products than the wild-type polymerase by utilizing nascent primers with non-complementary 3′-termini, indicating that they also discriminate less stringently between matched and mismatched primer termini. Concomitant increases in misincorporation and extension of mispaired termini has also been observed in human immunodeficiency virus type-1 reverse transcriptase (29).

We also assessed the fidelity of the purified mutant enzymes in the M13mp2 forward mutation assay (30). This assay has several important advantages for the measurement of *in vitro* misincorporation. First, one can quantitate mutation frequency and identify the types of mutations that are produced (Table I and Fig. 5). Second, synthesis can be carried out with equal concentrations of the four dNTPs, whereas misincorporation in the primer extension assay occurs in the absence of one of the nucleotide substrates (Figs. 1 and 4). Third, one can examine the sequence dependence of mutations. DNA polymerases such as *E. coli* pol I (31), pol α (32, 33), pol β (30), pols δ and ϵ (35), and human immunodeficiency virus type 1 reverse transcriptase (36) produce mutations non-randomly when copying DNA templates, producing a disproportionate number of mutations at "hot spots" that are unique to each polymerase. We observed at least a 7- and 25-fold increase for mutants A661E and T664R, respectively, compared with the wild-type enzyme. Although the relative misincorporation, T664R > A661E > wild-type, was the same in the forward mutation and primer extension experiments, the base substitution specificity appeared to differ. The primer extension experiments indicated that approximately equal amounts of primer were extended past the target nucleotides by T664R in each of the four reactions lacking a different dNTP. Thus, the A, C, G, and T target nucleotides may have about equal probability of substitution (Fig. 4*A, lanes 13–16*). In contrast, half of the base substitutions produced by T664R in the forward mutation assay were T → C transitions (Fig. 5). This difference may be attributable to differences in the sequence context of the target nucleotides in the two assays.

Most work on the fidelity of DNA polymerases has focused on the contribution of proofreading by the 3′-5′ exonuclease. Fewer studies have aimed at understanding the basis of discrimination between complementary and non-complementary nucleotide substrates. Such studies are most readily interpretable for polymerases that lack a proofreading exonuclease, because mutation in the polymerase domain can cause differences in partitioning between states active for polymerization or exonucleolytic proofreading (37). Studies on the fidelity of mutant DNA polymerases that lack editing exonuclease activity have been limited. D1002N and T1003S in the highly conserved region I of human DNA polymerase α have been shown to display 70- and 40-fold higher insertional fidelity in Mn^{2+}-catalyzed reactions, respectively, than that of the wild-type enzyme, as well as reduced Mn^{2+}-catalyzed extension of mismatched termini (38). G262A and W266A in helix H of the thumb domain of human immunodeficiency virus type-1 reverse transcriptase displayed reduced processivity and increased template-primer slippage errors (39). In addition, Y766S in the O-helix of exonuclease-deficient *E. coli* pol I Klenow fragment displayed reduced insertional fidelity (40); since this mutation also decreased DNA binding affinity (10) it would appear that Tyr-766 may influence dNTP discrimination via an interaction with the template. Because the fidelity of DNA synthesis has been examined in just a few mutant enzymes, it is possible that a wide variety of other mutations within the polymerase site may also alter fidelity.

Most of the substitutions that reduced fidelity in the present study of the O-helix in *Taq* pol I were found in the N-terminal part of the helix, and many of these substitutions clustered on

and in reactions lacking one dNTP, as indicated. *B*, incubation was for 60 min at 45 °C in reactions containing all four complementary dNTPs and in reactions lacking one dNTPs as indicated. *C*, incubation was at 45 °C for the times indicated in the absence of dCTP. *Open* and *filled triangles* represent the positions of the primer and the fully extended product, respectively. The sequence of the incorporated complementary nucleotides is given at the *right*. Only the primer was loaded in *lanes 1* and 7.

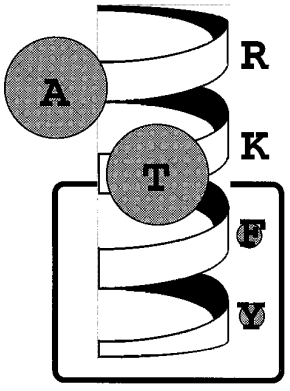

FIG. 6. **A schematic representation of amino acids Arg-659 through Tyr-671 in the O-helix of *Taq* pol I.** The diameter of circles around individual amino acids is proportional to their mutability, *i.e.* to their tolerance for single amino acid substitutions. Arg-659, Lys-663, Phe-667, and Tyr-671 may face the incoming dNTP (8) and are immutable or nearly immutable; Arg-659 and Lys-663 may interact with phosphate groups in the dNTP (8–11) and Phe-667 may interact with the sugar (13). Ala-661 and Thr-664 may face away from the incoming dNTP (8), are highly mutable, and are shown in the present study to contribute to the maintenance of wild-type fidelity. The hydrophobic residues Ile-665 through Tyr-671 are enclosed in a *square*.

either side of the essential (10, 11, 41, 42) and non-mutable Lys-663 (12) (Figs. 2 and 6). The most frequent substitutions among our low fidelity mutants occurred at Ala-661 and Thr-664 (Fig. 2). These amino acids are not on the face of the helix that interacts with the incoming dNTP, as judged from the crystal structure of *E. coli* pol I (8), nor are they essential for catalytic activity *in vivo*, based on the ability of other amino acids to replace them (12). In fact, we have reported that six different amino acids could replace either Ala-661 or Thr-664 and yield active DNA polymerases, in the absence of mutations at other positions that could compensate (12). A variety of amino acids are also found in the corresponding positions in other DNA polymerases within the pol I family. Ser, Asn, Lys, or Glu is found at the position corresponding to Ala-661, and Ala, Val, Lys, or Ile is found at the position corresponding to Thr- 664 (3). Thus, many substitutions are tolerated adjacent to the essential Lys-663 and Arg-659 in *Taq* pol I and homologous residues in other DNA polymerases.

Among the 13 low fidelity mutants we identified (Fig. 2), only two contain substitutions in the five C-terminal amino acids of the O-helix (Phe-667 through Tyr-671). This segment interacts with the terminal region of the template-primer duplex as judged from the crystal structure of *Taq* pol I in the presence of a blunt-ended oligonucleotide (43). The double-stranded oligonucleotide could represent the product of the polymerase reaction. Moreover the two substitutions we observed occurred in the presence of other replacements that could alter fidelity.

We offer two explanations for how the diverse substitutions we identified in the O-helix could reduce the fidelity of DNA synthesis. First, these substitutions might alter the local conformation around the catalytic pocket and thereby reposition critical residues that interact with either the template-primer or the incoming dNTPs. Among the substitutions in the 13 low fidelity *Taq* pol I mutants were five prolines, which disrupt helical structure, and three glycines, which could relax the conformation around critical residues at the active site.

A second explanation for the reduced fidelity of our mutants involves a pyrophosphate error prevention mechanism (44, 45). It might have been predicted that amino acid substitutions adjacent to the Phe and Tyr (Fig. 6) would alter the fidelity of DNA synthesis since they interact with the template-primer binding site (38). However, most of the substitutions that alter fidelity are located upstream from these residues in proximity to Lys-663. The assignment of individual amino acid residues that interact with specific atoms in the incoming dNTPs is tentative, at best (46). It is based primarily on inferences from kinetic studies (45), binding studies with binary complexes (44), and cross-linking studies with wild-type and mutant DNA polymerases (11). Nevertheless, evidence suggests that both Lys-663 and Arg-659, and the homologous residues in other DNA polymerases, interact with the β- and γ-phosphates in the incoming dNTPs and could thus facilitate removal of pyrophosphate during catalysis (44). The incoming dNTP may initially interact with the carboxylate region of the palm domain (46, 47). The carboxylates could anchor a pair of divalent metal ions to facilitate deprotonation of the 3′-hydroxyl terminus of the primer strand (46). As judged from the structure of *Taq* pol I with a double-stranded oligonucleotide, Lys-663 and Arg-659 in the O-helix in the fingers domain are distant from the carboxylate residues in the palm region and thus a conformational change may be required to facilitate their interaction with the β- and γ- phosphates in the incoming dNTP. Support for the involvement of Lys-663 and/or Arg-659 in pyrophosphate release is largely indirect. The corresponding Arg-754 and Lys-758 in *E. coli* pol I have been shown to interact with phosphate residues in the incoming dNTP (8, 9). Substitution of Ala at Lys-758 in *E. coli* pol I reduced k_{cat} 330–3,500-fold with poly(dA)·$(dT)_{15}$ as template-primer, while substitution of Ala at Arg-754 reduced k_{cat} 19–30-fold (10, 48). These data suggest that Arg-754 and Lys-758 in *E. coli* pol I are involved in catalysis but they do not establish a direct interaction with the dNTP or involvement in pyrophosphate release. More direct evidence stems from crystallographic data that showed that the Klenow fragment is complexed with pyrophosphate near the conserved Lys and Arg residues (42). Thus, our low fidelity mutants containing substitutions near Lys-663 or Arg-659 may exhibit attenuated pyrophosphate release. A two-step chemical mechanism for pyrophosphate release that effects the fidelity of DNA synthesis has been suggested. The first step involves binding of the substrate via the γ-phosphoryl group and the second step involves verification of complementary base pair interaction by coordination with the β-phosphoryl group (44). This error preventing step follows binding of the substrate and precedes primer elongation. Enhancement of misincorporation by the addition of pyrophosphate has been demonstrated for DNA polymerases (49) and has been attributed to a kinetic proofreading step in a multi-step proofreading mechanism (45).

The essentiality of individual amino acids in DNA polymerases has been inferred from sequence comparisons (1) that identify evolutionarily conserved and presumably critical residues. Chemical modification, site-specific mutagenesis, and kinetic studies follow, and provide quantitative data on which to assess detailed catalytic function. We have taken a different approach to evaluating the role of amino acids in the O-helix of *Taq* pol I. We constructed a library of active mutants by using random mutagenesis and genetic selection and identified immutable/essential residues (12). We then screened these active mutants for alterations in the fidelity of DNA synthesis. Our results suggest that many substitutions within the polymerase site can alter fidelity while preserving near wild-type activity. The mutants we have identified may be of value in ascertaining the mechanisms of substrate discrimination by DNA polymerases. In addition, they may have utility in applied molecular evolution (50) by increasing mutagenesis during amplification of genes by the polymerase chain reaction (51, 52).

Acknowledgments—We thank Bella S. Charurat and Karl Rose for their excellent technical assistance. We are grateful to Ann Blank and Premal Patel for helpful discussions.

REFERENCES

1. Blanco, L., Bernad, A., Blasco, M. A., and Salas, M. (1991) *Gene* (*Amst.*) **100,** 27–38
2. Delarue, M., Poch, O., Tordo, N., Moras, D., and Argos, P. (1990) *Protein Eng.* **3,** 461–467
3. Braithwaite, D. K., and Ito, J. (1993) *Nucleic Acids Res.* **21,** 787–802
4. Joyce, C. M., Kelley, W. S., and Grindley, N. D. F. (1982) *J. Biol. Chem.* **257,** 1958–1964
5. Lawyer, F. C., Stoffel, S., Saiki, R. K., Myambo, K., Drummond, R., and Gelfand, D. H. (1989) *J. Biol. Chem.* **264,** 6427–6437
6. Ollis, D. L., Brick, P., Hamlin, R., Xuong, N. G., and Steitz, T. A. (1985) *Nature* **313,** 762–766
7. Kim, Y., Eom, S. H., Wang, J., Lee, D.-S., Suh, S. W., and Steitz, T. A. (1995) *Nature* **376,** 612–616
8. Beese, L. S., Derbyshire, V., and Steitz, T. A. (1993) *Science* **260,** 352–355
9. Joyce, C. M., and Steitz, T. A. (1994) *Annu. Rev. Biochem.* **63,** 777–822
10. Astatke, M., Grindley, N. D. F., and Joyce, C. M. (1995) *J. Biol. Chem.* **270,** 1945–1954
11. Kaushik, N., Pandey, V. N., and Modak, M. J. (1996) *Biochemistry* **35,** 7256–7266
12. Suzuki, M., Baskin, D., Hood, L., and Loeb, L. A. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 9670–9675
13. Tabor, S., and Richardson, C. C. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 6339–6343
14. Kornberg, A., and Baker, T. A. (1992) *DNA Replication,* 2d Ed., pp. 130, W. H. Freeman and Co., New York
15. Tindall, K. R., Stein, J., and Hutchinson, F. (1988) *Genetics* **118,** 551–560
16. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) *Gene* (*Amst.*) **77,** 51–59
17. Fucharoen, S., Fucharoen, G., Fucharoen, P., and Fukumaki, Y. (1989) *J. Biol. Chem.* **264,** 7780–7783
18. Eckert, K. A., and Kunkel, T. A. (1990) *Nucleic Acids Res.* **18,** 3739–3744
19. Boosalis, M. S., Petruska, J., and Goodman, M. F. (1987) *J. Biol. Chem.* **262,** 14689–14696
20. Perrino, F. W., and Loeb, L. A. (1989) *J. Biol. Chem.* **264,** 2898–2905
21. Sweasy, J. B., and Loeb, L. A. (1992) *J. Biol. Chem.* **267,** 1407–1410
22. Lawyer, F. C., Stoffel, S., Saiki, R. K., Chang, S. Y., Landre, P. A., Abramson, R. D., and Gelfand, D. H. (1993) *PCR Methods Application* **2,** 275–287
23. Battula, N., and Loeb, L. A. (1974) *J. Biol. Chem.* **249,** 4086–4093
24. Feig, D. I., Sowers, L. C., and Loeb, L. A. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 6609–6613
25. Kunkel, T. A., and Loeb, L. A. (1979) *J. Biol. Chem.* **254,** 5718–5725
26. Feig, D. I., and Loeb, L. A. (1993) *Biochemistry* **32,** 4466–4473
27. Tindall, K. R., and Kunkel, T. A. (1988) *Biochemistry* **27,** 6008–6013
28. Kuchino, Y., Mori, F., Kasai, H., Inoue, H., Iwai, S., Miura, K., Ohtsuka, E., and Nishimura, S. (1987) *Nature* **327,** 77–79
29. Preston, B. D., Poiesz, B. J., and Loeb, L. A. (1988) *Science* **242,** 1168–1171
30. Kunkel, T. A. (1985) *J. Biol. Chem.* **260,** 5787–5796
31. Hillebrand, G. G., McCluskey, A. H., Abbott, K. A., Revich, G. G., and Beattie, K. L. (1984) *Nucleic Acids Res.* **12,** 3155–3171
32. Feig, D. I., and Loeb, L. A. (1994) *J. Mol. Biol.* **235,** 33–41
33. Kunkel, T. A. (1985) *J. Biol. Chem.* **260,** 12866–12874
34. Deleted in proof
35. Thomas, D. C., Roberts, J. D., Sabatino, R. D., Myers, T. W., Tan, C.-K., Downey, K. M., So, A. G., Bambara, R. A., and Kunkel, T. A. (1991) *Biochemistry* **30,** 11751–11759
36. Bebenek, K., Abbotts, J., Roberts, J. D., Wilson, S. H., and Kunkel, T. A. (1989) *J. Biol. Chem.* **264,** 16948–16956
37. Reha-Krantz, L. J., and Nonay, R. L. (1994) *J. Biol. Chem.* **269,** 5635–5643
38. Copeland, W. C., Lam, N. K., and Wang, T. S. M. (1993) *J. Biol. Chem.* **268,** 11041–11049
39. Bebenek, K., Beard, W. A., Casas-Finet, J. R., Kim, H.-R., Darden, T. A., Wilson, S. H., and Kunkel, T. A. (1995) *J. Biol. Chem.* **270,** 19516–19523
40. Carroll, S. S., Cowart, M., and Benkovic, S. J. (1991) *Biochemistry* **30,** 804–813
41. Osumi-Davis, P. A., de-Aguilera, M. C., Woody, R. W., and Woody, A. Y. M. (1992) *J. Mol. Biol.* **226,** 37–45
42. Osumi-Davis, P. A., Sreerama, N., Volkin, D. B., Middaugh, C. R., Woody, R. W., and Woody, A. Y. M. (1994) *J. Mol. Biol.* **237,** 5–19
43. Eom, S. H., Wang, J., and Steitz, T. A. (1996) *Nature* **382,** 278–281
44. Ferrin, L. J., and Mildvan, A. S. (1986) *Biochemistry* **25,** 5131–5145
45. Beckman, R. A., and Loeb, L. A. (1993) *Q. Rev. Biophys.* **26,** 225–331
46. Joyce, C. M., and Steitz, T. A. (1995) *J. Bacteriol.* **177,** 6321–6329
47. Polesky, A. H., Dahlberg, M. E., Benkovic, S. J., Grindley, N. D. F., and Joyce, C. M. (1992) *J. Biol. Chem.* **267,** 8417–8428
48. Pandey, V. N., Williams, K. R., Stone, K. L., and Modak, M. J. (1987) *Biochemistry* **26,** 7744–7748
49. Loeb, L. A., Dube, D. K., Beckman, R. A., Koplitz, M., and Gopinathan, K. P. (1981) *J Biol. Chem.* **256,** 3978–3987
50. Loeb, L. A. (1996) *Adv. Pharmacol.* **35,** 321–347
51. Cadwell, R. C., and Joyce, G. F. (1992) *PCR Methods Application* **2,** 28–33
52. Stemmer, W. P. C. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 10747–10751
53. Laemmli, U. K. (1970) *Nature* **227,** 680–685