Interaction of DNA Polymerase I (Klenow Fragment) with the Single-Stranded Template beyond the Site of Synthesis†

Robert M. Turner, Jr., Nigel D. F. Grindley, and Catherine M. Joyce*

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520

Received August 2, 2002; Revised Manuscript Received December 18, 2002

ABSTRACT: Cocrystal structures of DNA polymerases from the Pol I (or A) family have provided only limited information about the location of the single-stranded template beyond the site of nucleotide incorporation, revealing contacts with the templating position and its immediate 5′ neighbor. No structural information exists for template residues more remote from the polymerase active site. Using a competition binding assay, we have established that Klenow fragment contacts at least the first four unpaired template nucleotides, though the quantitative contribution of any single contact is relatively small. Photochemical cross-linking indicated that the first unpaired template base beyond the primer terminus is close to Y766, as expected, and the two following template bases are close to F771 on the surface of the fingers subdomain. We have constructed point mutations in the region of the fingers subdomain implicated by these experiments. Cocrystal structures of family A DNA polymerases predict contacts between the template strand and S769, F771, and R841, and our DNA binding assays provide evidence for the functional importance of these contacts. Overall, the data are most consistent with the template strand following a path over the fingers subdomain, close to the side chain of R836 and a neighboring cluster of positively charged residues.

Klenow fragment of DNA polymerase I is a 68 kDa protein which has the polymerase and 3′−5′ (proofreading) exonuclease activities of the parent molecule (*1*). The extensive genetic, biochemical, and structural studies that have been carried out on Klenow fragment make it an ideal model system for investigating the molecular mechanism of template-directed DNA synthesis and the way in which the polymerase interacts with its DNA and nucleotide substrates.

The large number of nucleic acid polymerase crystal structures now available show a common polymerase domain structure, the overall shape of which has been likened to that of a half-open right hand, with auxiliary activites, such as proofreading, on separate domains (*2*−*4*). The subdomains of the polymerase domain, named fingers, palm, and thumb, have analogous functions even in distantly related polymerases. The palm subdomain forms the base of the polymerase cleft; its highly conserved "polymerase fold" is the scaffold for many of the residues necessary for catalysis of phosphoryl transfer. The "family X" polymerases, exemplified by mammalian DNA polymerase β and terminal transferase, have a distinct palm domain fold, more closely related to nucleotidyl transferases (*5*, *6*). The mostly α-helical thumb subdomain is an important part of the binding site for the duplex product of DNA synthesis, while the fingers subdomain contributes much of the binding site for the incoming dNTP.

For several DNA polymerases, binary complex (polymerase−DNA) and ternary complex (polymerase−DNA−dNTP) cocrystal structures have been determined (*7*−*16*). In the Pol I (family A), Pol α (family B), and reverse transcriptase families, comparison of binary and ternary complex structures indicates that dNTP binding is accompanied by substantial movement of the fingers subdomain, causing the polymerase domain to change from an open to a closed conformation, and forming a tight binding pocket for the nascent base pair (*9*, *10*, *13*, *15*). The binary and ternary complex cocrystal structures give a detailed picture of the binding of the primer−template duplex to side chains of the palm and thumb subdomains, but provide relatively limited information about the binding of the single-stranded template beyond the site of nucleotide addition. The cocrystal structures of family A DNA polymerases show single-stranded template residues corresponding only to the position opposite the incoming nucleotide (hereafter called the 0 position) and its immediate 5′ neighbor (hereafter called the +1 position); the HIV-1 reverse transcriptase and bacteriophage RB69 DNA polymerase (family B) ternary complexes also show the +2 template nucleotide. When additional single-stranded template residues were present on the cocrystallization substrates, these were disordered in the crystals. On the basis of the position of these few nucleotides close to the templating position, the consensus is that the single-stranded template probably exits the polymerase complex over the surface of the fingers subdomain.

Although no structures of Klenow fragment with DNA at the polymerase active site have been determined, there are

† This work was supported by National Institutes of Health Grant GM-28550 and by NIH Minority Predoctoral Fellowship GM-18508 (to R.M.T.).

* To whom correspondence should be addressed: Bass Center for Molecular and Structural Biology, Department of Molecular Biophysics and Biochemistry, Yale University, 266 Whitney Ave., P.O. Box 208114, New Haven, CT 06520-8114. Phone: (203) 432-8992. Fax: (203) 432-3104. E-mail: catherine.joyce@yale.edu.

10.1021/bi026566c CCC: $25.00
Published on Web 02/07/2003

several informative cocrystal structures of other family A homologues, specifically, binary and ternary complexes of the Klenow fragment analogue from *Thermus aquaticus* (Klentaq) (*10*, *12*), a ternary complex of the DNA polymerase from bacteriophage T7 (T7 DNA pol) (*9*), and several binary complexes of the Klenow fragment analogue from *Bacillus stearothermophilus* DNA polymerase (*Bst* pol) (*11*). The ternary complexes of T7 DNA pol and Klentaq (four independent structures) all show the templating (0 position) base positioned, as expected, within the binding pocket for the nascent base pair. Although the DNA oligonucleotides all have three single-stranded template bases, no DNA is visible beyond the +1 position. The +1 base is dislocated by ~180° from the stacking arrangement of the template−primer duplex and occupies a similar, but by no means identical, location in the five complexes. In contrast, the binary complexes of Klentaq and *Bst* pol indicate quite different locations for the 0 position base. In the *Bst* pol binary complexes, which have been demonstrated to carry out catalysis in the crystal (*11*), the 0 position base occupies a binding pocket between the O and O1 helices. The Klentaq binary complex, formed by allowing the bound dNTP to diffuse out of a ternary complex crystal, shows the 0 position base positioned above its 3′ neighbor in the template−primer duplex (*10*). The +1 base is visible only in one of the *Bst* pol complexes (Protein Data Bank entry 4BDP) and is stacked with an aromatic side chain from the O1 helix.

In view of the limited and sometimes contradictory structural data described above, we have undertaken a study of the binding of the single-stranded template to Klenow fragment, determining first how much of the template overhang contacts the enzyme. We then used photochemical cross-linking to identify probable protein−DNA contacts and investigated the functional consequences of mutating side chains likely to be within the contact region.

EXPERIMENTAL PROCEDURES

Materials

Oligonucleotides. The DNA oligonucleotides shown in Figure 1 were synthesized by the Keck Biotechnology Resource Laboratory at Yale Medical School (New Haven, CT). The hairpin oligonucleotides (Figure 1A,C) were purified by gel electrophoresis. Oligonucleotides containing halopyrimidine bases for photochemical cross-linking (Figure 1B) were ethanol-precipitated in minimal light. The concentrations of all oligonucleotides were determined spectrophotometrically using calculated extinction coefficients (*17*). The halopyrimidine-containing templates were 5′ end-labeled with [γ-^{32}P]ATP using T4 polynucleotide kinase and then annealed to the primer (≥3-fold molar excess) in 10 mM Tris-HCl (pH 7.5) and 5 mM $MgCl_2$.

Klenow Fragment and Mutant Derivatives. Mutant derivatives of Klenow fragment were constructed, expressed, and purified as described previously (*18*, *19*), with one modification. The final gel filtration column of the published purification method was omitted; instead, peak fractions from the phenyl-Superose column were precipitated by addition of ammonium sulfate to 85% saturation, redissolved to give a concentration of ~10 mg/mL protein, and dialyzed into 50 mM Tris-HCl (pH 7.5), 0.5 mM DTT, and 50% (v/v) glycerol for storage at −20 °C. Protein concentrations were measured by the Bradford colorimetric assay (*20*). All polymerases in this study carried the D424A mutation, which eliminates the 3′−5′ exonuclease activity (*21*). Each protein is described by the mutation(s) in the polymerase domain; thus, the D424A control is termed wild type, and S769A is the D424A/S769A double mutant.

A. Hairpin substrates

```
                                  H1 n=0
                                  H2 n=1
                                  H3 n=2
    A A
    A GGCTACCTGCATGA-3′           H4 n=3
    A CCGATGGACGTACTAC n-5′       H5 n=4
    G                             H6 n=5
                                  H8 n=7
```

B. Photo cross-linking substrates (X=BrdU or IdU)

Primer (P)	5′-ACGCCTGGCGGATACGACGAC$_{dd}$-3′
Template (T_{0-4}X)	3′-TGCGGACCGCCTATGCTGCTGXXXXXGAC-5′
Template (T_0X)	3′-TGCGGACCGCCTATGCTGCTGXAGAAGAC-5′
Template (T_1X)	3′-TGCGGACCGCCTATGCTGCTGCXGAAGAC-5′
Template (T_2X)	3′-TGCGGACCGCCTATGCTGCTGCAXAAGAC-5′
Template (T_3X)	3′-TGCGGACCGCCTATGCTGCTGCAGXAGAC-5′
Template (T_4X)	3′-TGCGGACCGCCTATGCTGCTGCAGAXGAC-5′
Template (TCNTL)	3′-TGCGGACCGCCTATGCTGCTGCAGAAGAC-5′

C. Hairpin substrate for DNase I footprinting assay

```
    A A
    A GACCAGTTCATCGGCTACCTGCATGA-3′
    A CTGGTCAAGTAGCCGATGGACGTACTAGTATGCATGTG-5′
    G
```

FIGURE 1: DNA oligonucleotides used in this study. Hairpin structures are closed using a stable tetraloop (*33*). (A) A set of 14 bp hairpin oligonucleotides, with 5′ template overhangs of different lengths, used in gel mobility shift and competition assays. (B) Oligonucleotides used in photochemical cross-linking experiments. The common primer (P) is complementary to a series of template oligonucleotides containing bromodeoxyuridine or iododeoxyuridine at the position marked with an X. The control template sequence, TCNTL, has no photoactivatable bases (including thymine) in the template overhang. (C) Hairpin 68mer oligonucleotide used in DNase I footprinting. This oligonucleotide has the same sequence, over the duplex region bound by the polymerase, as the oligonucleotides in panel A. The longer duplex was necessary to obtain a DNase I footprint.

Methods

Photochemical Cross-Linking (Analytical). Small-scale cross-linking reactions (20 μL) with duplex oligonucleotides containing BrdU at various positions in the 5′ template overhang (see Figure 1B) were performed in Falcon U-bottom microtiter dishes on ice. The reaction mix contained 13 nM wild-type (D424A) Klenow fragment, 13 nM duplex oligonucleotide, 5′-labeled on the template strand, 50 mM Tris-HCl (pH 7.5), and 10 mM $MgCl_2$. Gel mobility shift analysis showed that this concentration was sufficient to promote efficient (≥80%) formation of the polymerase−DNA complex. When ternary complexes were studied, the dNTP complementary to the first template position was present at a concentration of 20 μM. The reaction mixtures were covered and equilibrated for 5 min, and then irradiated at a distance of 1 cm with a 115 V, 302 nm, UV transilluminator (UV Products). After appropriate times of irradiation, 5 μL of 250 mM Tris-HCl (pH 7.5), 20% (v/v) glycerol, 3.4% (w/v) SDS, 0.01 M DTT, and 0.01% (w/v) bromophenol blue was added. Samples were fractionated by electrophoresis on a 10% SDS−polyacrylamide gel, and the cross-linking yield was quantitated by phosphorimaging. A series of control experiments were carried out by making adjustments to the reaction conditions as described in Table 2.

Preparative Cross-Linking Reactions and Trypsin Digestion. Large-scale cross-linking reaction mixtures (1 mL) for peptide analysis contained 1 μM 5′-labeled duplex oligonucleotide with a single photoactivatable base, 1 μM Klenow fragment, 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, and, when studying ternary complexes, the next complementary dNTP (20 μM). Usually, the photoagent was IdU, and the mixture was irradiated in a QS 1 cm quartz cuvette using a Kimmon IK Series HeCd laser at 325 nm for 1 h at 4 °C. In a few cases, BrdU was used and the transilluminator procedure described above was scaled up appropriately. The efficiency of cross-linking was quantitated as described above. For subsequent analysis, three 1 mL cross-linking reaction mixtures were pooled and concentrated using a Microcon YM-10 centrifugal concentrator (Millipore), and then washed twice with 10 mM Tris-HCl (pH 7.5) and 0.1 mM EDTA and once with 2 mM ammonium bicarbonate. The concentrated sample (~8 μL) was digested at 37 °C for 12 h in an 80 μL reaction mixture containing 0.4 mM ammonium bicarbonate, 8 M urea, and 2.5 mg/mL trypsin. The extent of digestion was monitored by electrophoresis on 10% SDS−polyacrylamide and 10% polyacrylamide−urea gels.

Isolation of the Cross-Linked Complex and Peptide Sequencing. The tryptic digest was mixed with 2 volumes of 10 mM Tris-HCl (pH 7.5), 0.1 mM EDTA, and 1 M NaCl, concentrated on a Microcon YM-10 concentrator, washed two additional times with the same buffer, and finally desalted with 10 mM Tris-HCl (pH 7.5) and 0.1 mM EDTA. The eluate, concentrated to 8 μL, was fractionated on a 10% polyacrylamide−urea gel and transferred electrophoretically to a PVDF membrane (ProBlott) in 0.01 M CAPS buffer (pH 11) at 70 V for 45 min using a TransBlot apparatus (Bio-Rad). Successful transfer to the membrane was verified by checking with a Geiger counter for the removal of radioactivity from the gel. The membrane was stained with Coomassie blue, extensively destained with 50% (v/v) methanol and 10% (v/v) acetic acid, and rinsed multiple times with HPLC-grade H_2O. After the membrane had been air-dried overnight, the cross-linked peptide was excised using an autoradiogram as a guide. Cross-linked peptides (typically, 5−20 pmol) were sequenced by automated Edman degradation at the Keck Biotechnology Resource Laboratory at Yale Medical School.

Competition Assay. Competition of a mixture of hairpin oligonucleotides for a limiting amount of Klenow fragment was assessed in reaction mixtures (30 μL) containing each of four oligonucleotides (H1−H4, or H4−H6, and H8; ~1 μM) and 2 nM wild-type Klenow fragment in 10 mM Tris-HCl (pH 7.5) and 5 mM $MgCl_2$. After equilibration for 5 min at 22 °C, the reaction was initiated by addition of 3 μL of 10 nM [α-^{32}P]dTTP (3000 Ci/mmol). Samples (5 μL) were removed at appropriate intervals and the reactions quenched by addition of 2 μL of quench buffer [20 mM EDTA, 80% (v/v) deionized formamide, 0.2% (w/v) SDS, 0.1% (w/v) bromophenol blue, and 0.1% (w/v) xylene cyanol]. Samples were fractionated on a 10% polyacrylamide−urea gel, and the incorporation of [α-^{32}P]dTTP into each DNA substrate was quantitated on a phosphorimager. The exact ratios of each DNA oligonucleotide present in the competition assay were measured in a 30 μL reaction mixture, as above, except that $^1/_{10}$ the amount of DNA was used, together with 0.3 μM [α-^{32}P]dTTP and 7 μM unlabeled dTTP, and the mixture was incubated for 20 min to allow the reaction to proceed to completion.

Mutant derivatives of Klenow fragment were compared with the wild type by performing a competition assay using the H1−H4 oligonucleotide set. The reaction was performed as described above except that the oligonucleotide concentrations were 3.48 μM H1, 0.36 μM H2, 0.12 μM H3, and 0.04 μM H4; under these conditions, Klenow fragment gave approximately equal levels of [α-^{32}P]dTTP incorporation into all four substrates. The incubation times were increased as needed for mutant Klenow fragment derivatives with low catalytic activity.

Measurement of K_D by DNase I Protection. DNase I footprinting was performed essentially as described previously (*18*). The DNA substrate was a 68mer hairpin oligonucleotide (Figure 1C) which was labeled at the 5′ end, using T4 polynucleotide kinase and [γ-^{32}P]ATP, and purified by gel electrophoresis. The binding reaction mixture contained 0.1 nM labeled DNA, 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 50 mM NaCl, and variable concentrations of wild-type or mutant Klenow fragment.

Gel Mobility Shift Assay. The procedure was essentially as described previously (*22*, *23*). Hairpin oligonucleotides from the series shown in Figure 1A were labeled at the 3′ end by incorporation of [α-^{32}P]dTTP and were purified by gel electrophoresis. In some instances, the 5′ end was phosphorylated with unlabeled ATP.

Steady-State Kinetic Measurements of the Polymerase Reaction. The steady-state k_{cat}, for the polymerase reaction catalyzed by wild-type and mutant Klenow fragment derivatives, was determined as described previously (*18*), using a homopolymeric poly(dA)−oligo(dT) substrate. A dNTP concentration of 40 μM was used to provide saturating conditions.

RESULTS

Binding of Wild-Type Klenow Fragment to the Single-Stranded Template Overhang. The location of important binding contacts in the single-stranded template overhang was determined by a competition assay in which an equimolar mixture of oligonucleotides with different lengths of single-stranded 5′ overhang (Figure 1A) competed (in groups of four) for a limiting amount of wild-type Klenow fragment. The relative amounts of each DNA that became enzyme-bound were measured by incorporation of [α-^{32}P]dTTP and subsequent gel fractionation (Figure 2A). The radioactivity incorporated into each DNA should reflect its binding affinity, provided that the polymerase and DNA substrates remain in equilibrium as the reaction proceeds. By using extremely low dTTP concentrations (<10 nM), we slowed the polymerase reaction sufficiently to allow the polymerase to re-equilibrate with the mixture of DNAs before each round of nucleotide addition. Under these conditions, the relative yields of the radioactive products were the same as when the reaction was carried out in a Rapid Quench-Flow apparatus so as to monitor only the first turnover, reflecting enzyme−DNA complexes that were in equilibrium before the start of the reaction (data not shown). Despite the low dTTP concentration, reversal of the polymerase reaction by pyrophosphorolysis did not appear to be a problem, as shown

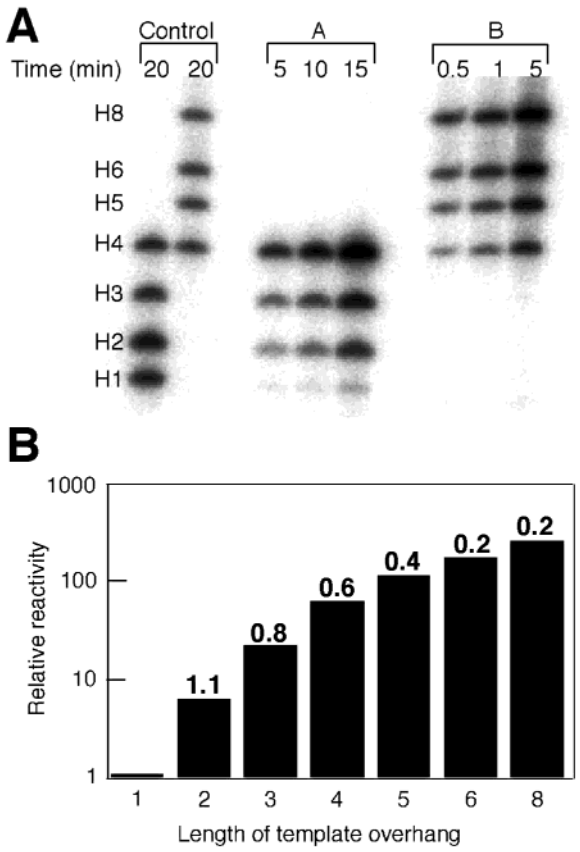

FIGURE 2: Preference of wild-type Klenow fragment for DNA substrates with differing template overhang lengths. (A) Competition assays using the hairpin oligonucleotides (H1−H8, Figure 1A), which were tested as two sets of four (H1−H4, in the lanes marked A, H4−H6 and H8, in the lanes marked B) in an approximately equimolar mixture. Samples were removed at the indicated times, and the amount of radioactivity in each oligonucleotide reflects its binding affinity for wild-type Klenow fragment. The control lanes correspond to reactions which were allowed to proceed to completion, so the measured radioactivity indicates the actual relative molarities of the oligonucleotides in each competition assay mixture. (B) Graphical representation of the results of the competition assay. The radioactivity incorporated into each hairpin substrate was corrected for differences in the molarities of the oligonucleotides, using the values from the control lanes, and then normalized to the value for the worst substrate (H1). The relative reactivity of each substrate is plotted on a logarithmic axis vs template overhang length. The number above each bar indicates the increased free energy of binding (in kilocalories per mole) relative to the DNA molecule that has one fewer template base.

by the lack of an effect of adding inorganic pyrophosphatase to the reaction mixtures.

Quantitation of the competition assay, corrected for differences in the concentrations of the input oligonucleotides, is shown graphically in Figure 2B. Wild-type Klenow fragment showed the greatest degree of discrimination among substrates having short template overhangs, as demonstrated by the 60-fold preference for the four-nucleotide template over the one-nucleotide template, compared with the modest 4-fold preference for the eight-nucleotide template over the four-nucleotide template. The stepwise increase in binding affinity with each additional template base from position 0 to +3 suggests that Klenow fragment contacts each of the first four nucleotides in the single-stranded region beyond the primer terminus.

Table 1: Gel Shift Analysis of Binding by Wild-Type Klenow Fragment to the 5′ Single-Stranded Template

template overhang	K_d(DNA) (nM)	relative affinity	no. of determinations
5′ (OH)C	8.0	1	2
5′ pC	1.3	6	2
5′ (OH)CC	0.8	10	2
5′ pCC	0.1	80	3
5′ pCCC	0.03	270	1

Binding contacts on the single-stranded template overhang were investigated further by gel mobility shifts. The hairpin oligonucleotides H1−H4 were labeled at the 3′ end by incorporation of [α-^{32}P]dTTP, giving a series of substrates having from zero to three bases as the 5′ template overhang. To evaluate the contribution of the phosphate groups, a portion of each oligonucleotide was phosphorylated with unlabeled ATP before 3′ labeling, generating pairs of DNA molecules that differed only with respect to the presence of a 5′ phosphate. The binding of each DNA was evaluated by gel mobility shift (Table 1). Oligonucleotides with no template overhang, or with a single nonphosphorylated template base, gave unstable complexes that were hard to quantitate with confidence. The addition of a 5′ phosphate contributed substantially to the binding affinity of single-nucleotide and two-nucleotide template extensions.

Protein−DNA Cross-Linking. We used photochemical cross-linking to identify the region of the polymerase that is close to the single-stranded template. A series of duplex oligonucleotides was designed in which the template strand contained a single photoactivatable base within the 5′ template overhang (Figure 1B). By varying the position of the photoactivatable base and sequencing the resulting cross-linked peptides, we planned to identify pairwise contacts between particular nucleotides and corresponding regions of the polymerase. As photoagents we used halopyrimidines, 5-bromodeoxyuridine (BrdU), and 5-iododeoxyuridine (IdU), which have two important advantages; the bromo or iodo substituent is similar in size to the 5-methyl group of thymidine and therefore unlikely to be disruptive, and the bases are photoactivated at wavelengths above 300 nm, which decreases the likelihood of nonspecific photodamage to DNA or protein (*24*). The primer strand had a dideoxynucleotide at the 3′ end, allowing formation of the polymerase−DNA−dNTP ternary complex but preventing further reaction. To test the feasibility of the cross-linking approach and to optimize reaction conditions, we conducted small-scale pilot experiments with the P−T_{0-4}Br duplex DNA. By using a template oligonucleotide with five potentially cross-linkable positions, we increased the chances that at least one would be close to a suitably reactive amino acid side chain. Irradiation of a ternary complex of Klenow fragment with the P−T_{0-4}Br duplex and dATP gave a good yield of a putative cross-linked product, reaching 30−40% after irradiation for 1 h (Figure 3A,B). Irradiation of Klenow fragment for 1 h under the same conditions but in the absence of DNA had no effect on polymerase activity (data not shown), so it is unlikely that significant protein damage occurs under these irradiation conditions.

A series of control experiments using the P−T_{0-4}Br and P−T_{0-4}I duplexes confirmed that the new species observed on gel electrophoresis (Figure 3A) did indeed correspond to

Table 2: Effect of Reaction Conditions on Cross-Linking Yield

reaction conditions	cross-linking (%)[a] $T_{0-4}Br$	$T_{0-4}I$
complete reaction	100	100
no Klenow fragment	0	0
no photoagent[b]	0	0
no UV light	0	0
no Mg^{2+}	80	80
no dATP	88	88
dsDNA[c]		
4 nM	60	55
200 nM	0	0
2 μM	0	0
ssDNA[d]		
4 nM	100	110
200 nM	80	81
2 μM	55	70
125 mM NaCl, no dATP[e]	0	0
500 mM NaCl, no dATP[e]	0	0
125 mM NaCl, dATP[e]	93	88
500 mM NaCl, dATP[e]	0	0

[a] Cross-linking to ^{32}P-labeled $T_{0-4}Br$ and ^{32}P-labeled $T_{0-4}I$ was assessed in 20 μL reaction mixtures irradiated for 30 min as described in Methods. Reagents were added or omitted as indicated. [b] The DNA template used was TCNTL. See Figure 1. [c] Unlabeled $T_{0-4}Br$ or $T_{0-4}I$ was annealed to the primer as described in Methods. It was added to the labeled DNA at the stated concentrations. Additional components of the cross-linking reaction were then added and, after a 5 min equilibration, 10 μL reaction mixtures were irradiated for 30 min. [d] Unlabeled TCNTL (not annealed to primer) was added to the labeled DNA, and irradiations were as described above. [e] NaCl was added at the stated concentrations to the labeled DNA and irradiations were as described above.

cross-linking of a bona fide polymerase−DNA complex (Table 2). As would be expected, the new species was not observed in the absence of UV irradiation, or if the DNA did not contain a halopyrimidine base. The putative cross-linked complex was not formed in the absence of protein, ruling out the possibility of a DNA−DNA cross-link. When unlabeled duplex DNA was added to the reaction mixture at 4 nM (equimolar with the labeled DNA), the amount of radioactivity in the cross-linked species was reduced by approximately half, whereas single-stranded DNA was a much poorer competitor. Omission of $MgCl_2$ caused a slight decrease in the yield of the cross-linked product, as did formation of the binary (as opposed to ternary) complex by omitting dATP. In the presence of 500 mM NaCl, no cross-linking was observed with either binary or ternary (with dATP) complex conditions, while 125 mM NaCl abolished cross-linking from the binary but not the ternary complex. The persistence of cross-linking in the presence of 125 mM NaCl indicates that 20 μM dATP was sufficient to promote ternary complex formation, even though other authors have used higher dNTP concentrations for this purpose (*25*). We have also demonstrated by gel mobility shifts that 20 μM complementary dNTP caused a decrease in $K_{d(DNA)}$ of ~10-fold, which is indicative of ternary complex formation (*25*) (data not shown).

Once suitable reaction conditions had been established, cross-linking reactions were carried out using oligonucleotides with a single halopyrimidine located at different positions along the 5′ template overhang. Small-scale reactions were used to determine which halopyrimidine positions were sufficiently close to cross-linkable amino acids (Table 3). When the halopyrimidine was located at the 0 position

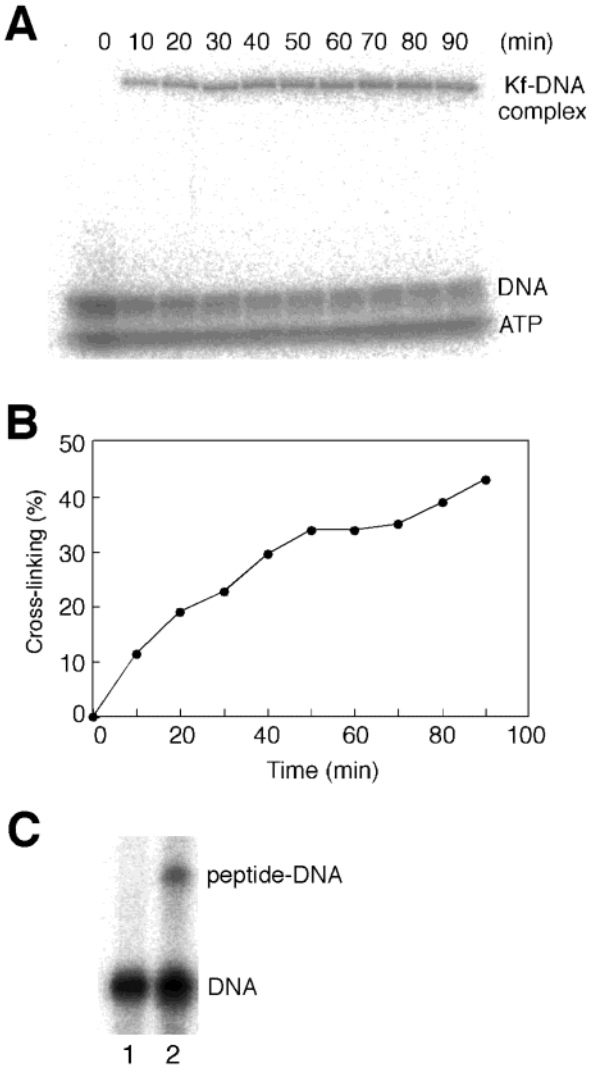

FIGURE 3: Photochemical cross-linking of DNA to Klenow fragment. (A) Time course of cross-linking of the P−$T_{0-4}Br$ oligonucleotide duplex. The DNA substrate, ^{32}P-labeled at the 5′ end of the template strand, was mixed with Klenow fragment and dATP to form a ternary complex, and irradiated for the indicated times as described in Methods. Electrophoresis on a 10% SDS−polyacrylamide gel separated the cross-linked protein−DNA complex from free DNA and from [γ-^{32}P]ATP remaining from the labeling reaction. (B) Quantitation of the experiment shown in panel A. The percentage of DNA cross-linked to Klenow fragment is plotted. (C) Isolation of the cross-linked peptide. Klenow fragment, cross-linked to the P−T_2I DNA substrate, was digested with trypsin. The labeled DNA-linked peptide was separated from unlinked DNA and from other peptides on a 10% polyacrylamide−urea gel.

(opposite the incoming dNTP), the cross-linking yield was higher in the absence of the complementary dNTP (binary complex) than in its presence (ternary complex). Good yields of cross-linked protein were obtained from both binary and ternary complexes with oligonucleotides substituted at the +1, +2, and +4 positions. No cross-linking was observed at the +3 position, presumably due to the lack of a suitably reactive amino acid side chain close to the +3 template base.

Reactions that gave a good level of cross-linking [0 (binary complex), and +1, +2, and +4 as ternary complexes] were scaled up to allow isolation and sequencing of the cross-linked peptide. Most of the preparative cross-linking reactions were carried out using IdU as the photoagent and irradiating with a HeCd laser. Because the laser produces a narrower

Table 3: Effect of the Location of the Halopyrimidine Base on Cross-Linking Efficiency

position of halopyrimidine base[a]	cross-linking (%)		
	BrdU[b]		IdU[c]
	without dNTP	with dNTP	with dNTP
0	2.1	1.1	2.9
+1	7.9	7.7	3.8
+2	6.5	8.6	6.4
+3	0	0	
+4	10.5	8.5	3.4

[a] Oligonucleotides used in the assay contained a single halopyrimidine template base as described in Figure 1B. Template positions are designated as follows: 0 indicates the template base opposite the incoming dNTP, +1 is the next base on the 5′ side, and so on. [b] Cross-linking reactions were performed as described in Methods in the absence (without dNTP) or presence (with dNTP) of the nucleotide complementary to the templating base, corresponding to a binary or a ternary complex, respectively. The results for BrdU cross-linking are the average from three analytical-scale experiments in which all the cross-linking positions were examined in parallel. Although the overall yield of cross-linked protein varied from experiment to experiment, the trends were consistent. [c] The results for IdU cross-linking were the average of two or more preparative-scale cross-linking reactions.

spectrum of UV light than the transilluminator used in our initial exploratory experiments, it decreases the likelihood of photodamage to the protein. Trypsin digestion of the cross-linked protein yielded a novel labeled species that migrated more slowly than the single-stranded template DNA on a TBE−urea gel (Figure 3C). This was transferred to a PVDF membrane and sequenced by automated Edman degradation. A single tryptic peptide, A759−R775, was identified unambiguously from experiments in which the photoactivatable base was present at the 0, +1, or +2 position. Cross-linking from the +4 position gave a mixture of peptides, resulting in several amino acids being identified at each Edman degradation cycle. The A759−R775 peptide could be identified as a major component of the mixture; the remaining amino acids did not correspond in a straightforward way to any of the predicted tryptic peptides.

The yields of phenylthiohydantoin (PTH) amino acid derivatives at each Edman degradation cycle indicated the probable sites of cross-linking (Figure 4). When the photoactivatable base was at the template 0 position in a binary complex (Figure 4A), the PTH amino acid yield showed a pronounced decrease at Y766 relative to the preceding amino acid. This implies that the 0 position cross-linked to Y766; moreover, continuation of the sequence beyond these residues suggests that the cross-link was to the amino acid side chain, not to the main chain, which would have blocked further degradation of the peptide. The low PTH amino acid yields at M768 and S769 in Figure 4A probably indicate some cross-linking of IdU at the 0 position to these side chains, although the low signal from these late degradation cycles makes it difficult to tell whether these decreases are significant. As can be seen from the examples in Figure 4, the signals at Gly positions were consistently lower than from neighboring amino acids; this could be attributed to the high background for glycine in the chromatograms and does not indicate cross-linking at these positions. When IdU was at the template +2 position in a ternary complex, decreases in the PTH amino acid yield were diagnostic of cross-linking to F771 and perhaps S769 (Figure 4C). With the photoactivatable base at the +1 position in a ternary complex, there

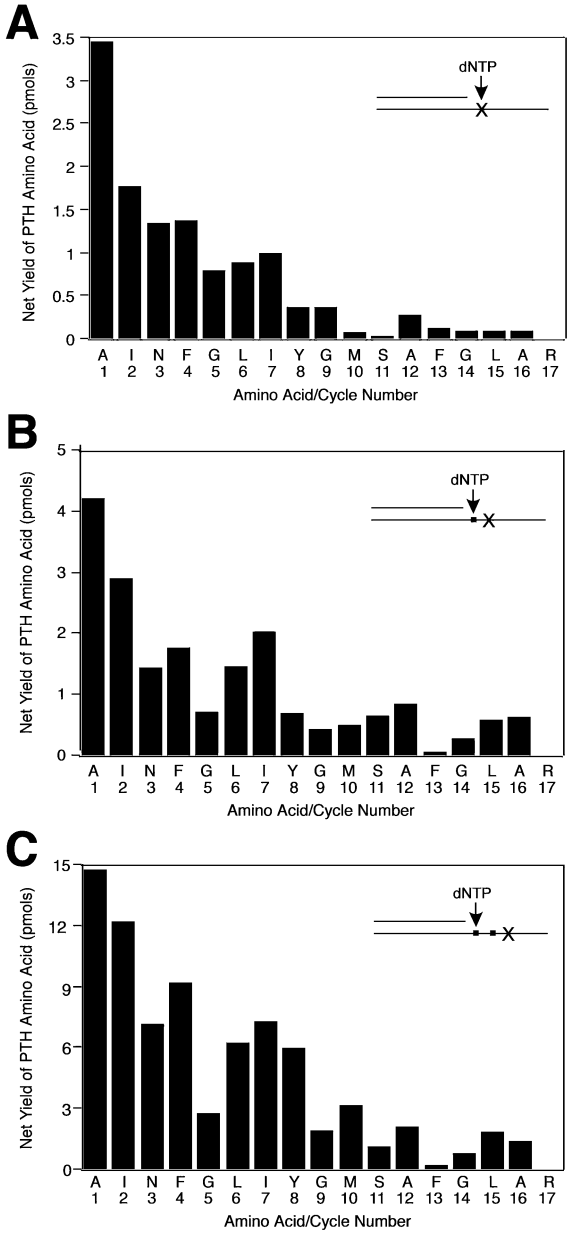

FIGURE 4: Sequencing of cross-linked peptides. Representative examples of data are shown for cross-linking from BrdU or IdU, represented with an X in the schematic diagrams, at the template 0 (A), +1 (B), and +2 (C) positions. The data in panel A were from a binary complex; the data in panels B and C were from ternary complexes. In each case, the peptide sequence that was obtained corresponded to A759−R775. The PTH amino acid yield at each cycle (except the first) was corrected for background by subtracting the yield for the same PTH amino acid at the previous cycle.

were decreases in the PTH amino acid yield at both Y766 and F771, suggesting that the isolated peptide was a mixture of species, cross-linked at either Y766 or F771 (Figure 4B).

Mutations in the Fingers Subdomain. We used site-directed mutagenesis to study amino acids within the fingers subdomain that were possible candidates for interaction with the template overhang. Residues were chosen (Figure 5a) on the basis of their proximity to the cross-linked amino acid F771, their ability to interact with the template strand, and information obtained from binary and ternary complex crystal structures of Klentaq, *Bst* pol, and T7 DNA pol (*9*−*11*). The

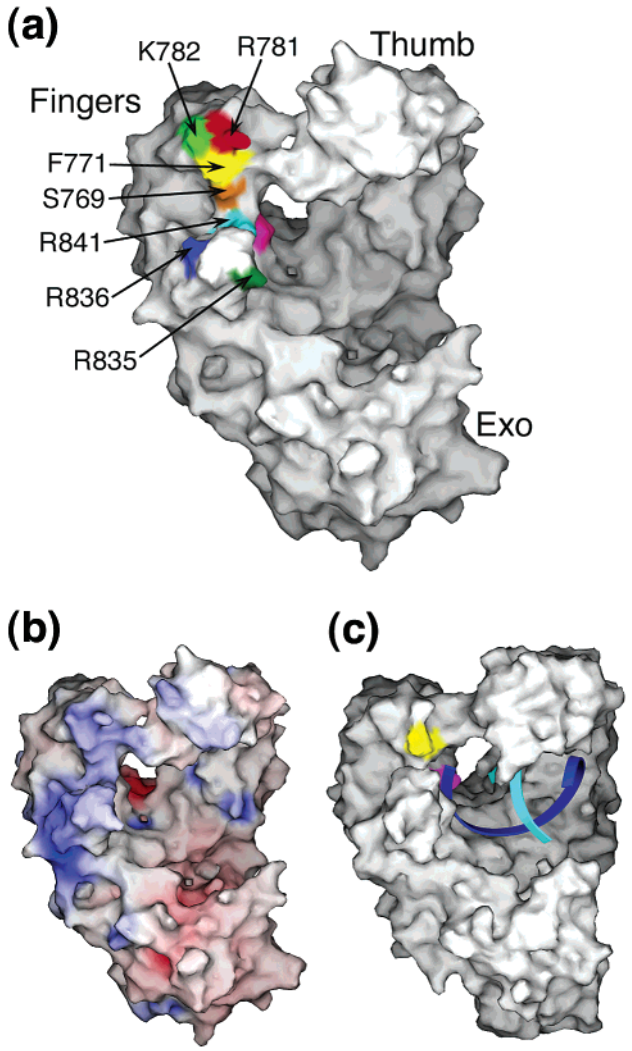

FIGURE 5: Location of relevant amino acid residues on a surface representation of the Klenow fragment structure. (a) Residues mutated in this study are colored and labeled. As a reference point, Y766 in the nascent base pair binding pocket is shown in magenta. (b) Surface charge distribution of Klenow fragment. (c) The duplex DNA binding site in family A DNA polymerases is illustrated using the binary complex of *Bst* DNA polymerase with DNA. The DNA backbone is represented as a ribbon with the template strand in blue and the primer strand in cyan. Residues Y714 and Y719 (equivalent to Y766 and F771, respectively, of Klenow fragment) are colored magenta and yellow, respectively. These figures were generated from the protein coordinates in PDB entries 1KFS (*34*) and 4BDP (*11*), using the program SPOCK (*35*).

S769 and R841 residues were chosen because both are predicted to interact with the single-stranded template based on the Klentaq ternary complex structure; the side chain equivalent to S769 of Klenow fragment interacts with the phosphate 5′ to the templating base (i.e., between the 0 and +1 bases), and the side chain equivalent to R841 is predicted to interact with the sugar of the +1 base and the phosphate 3′ to the templating base. The F771 side chain that became cross-linked in our experiments is highly conserved as F, Y, or H in family A polymerases, suggesting that it might make stacking interactions with the template overhang. Examination of the surface of the fingers subdomain of Klenow fragment showed two patches of positive charge that might be involved in anchoring the template strand (Figure 5b). One, in a cleft extending beyond R841 and down the far side of the fingers subdomain, contains the highly conserved R836; the other is just beyond F771 and involves R781 and K782. The side chain of R835, although not well conserved in Pol I homologues (including *Taq* and T7 DNA polymerase), is also interesting because it appears to extend toward the DNA template strand. With the exception of R835 and R841, these side chains had not been mutated previously in our lab. We therefore constructed the alanine replacements of S769, F771, and R836, and the double replacement, R781A/K782A.

Table 4: Kinetic Parameters of Klenow Fragment and Mutant Derivatives

enzyme[a]	k_{cat} (s^{-1})[b]	$K_{d(DNA)}$ (nM)[b,c]	relative DNA affinity [$K_{d(wild\text{-}type)}/K_{d(mutant)}$]
wild type	2.9	2.1	1
S769A	1.8	10.5	0.2
F771A	3.2	5.6	0.4
R781A/K782A	5.2	1.4	1.5
R835L	3.6	1.7	1.2
R836A	2.9	0.72	3
R841A	0.5	19	0.1

[a] All enzymes in this study carry the D424A (exonuclease-deficient) mutation. [b] Data from at least two determinations were combined to calculate k_{cat} and $K_{d(DNA)}$. Differences of ≤2-fold are not considered significant. [c] The substrate and reaction conditions for the DNA binding measurements differed from those in our previous studies (*18*, *22*), and therefore the K_d values reported for wild-type Klenow fragment are different. Our original study (*18*) used a labeled oligonucleotide annealed to a single-stranded circular M13 and reported a K_d of 8 nM, higher than values obtained subsequently, probably due to interference from the large amount of single-stranded DNA present in the reaction mixture. More recent work, using a 68mer hairpin oligonucleotide (Figure 1C), gave a K_d of 0.2 nM (*22*). Because it is difficult to measure such a low K_d accurately, we included 50 mM NaCl in the measurements reported here; this increased the K_d for wild-type Klenow fragment by ~10-fold. Despite the modifications to the protocol, the relative DNA affinities of wild-type and R841A Klenow fragment reported here agree well with the earlier study (*18*). A typical DNase I footprinting titration is shown in Figure 9 (Supporting Information).

The purified mutant proteins were characterized by measuring the steady-state k_{cat} on a homopolymeric substrate in the standard DNA polymerase assay, and by measuring DNA binding affinity by DNase I protection (Table 4). The k_{cat} values determined for wild-type (WT) and R841A Klenow fragment were in good agreement with those previously obtained (*18*). Aside from R841A, none of the other mutations affected k_{cat} significantly, consistent with their location away from the catalytic site. The R841A mutation caused a 9-fold decrease in DNA binding affinity, while mutations at S769 and F771 resulted in smaller decreases. The R836A mutation resulted in a 3-fold increase in binding affinity. The R781A/K782A double mutation and the R835L mutation did not have a significant effect on DNA binding.

Competition Assay with Mutant Proteins. The competition assay described earlier was modified to determine whether any of the mutations being studied affected the interaction of Klenow fragment with the single-stranded template. Competition among oligonucleotides H1−H4 was assayed because it appeared that the most significant binding contacts were within the first four unpaired template bases. To improve the signal from the shortest oligonucleotides, the concentrations of the four oligonucleotides were adjusted so that wild-type Klenow fragment gave approximately equal

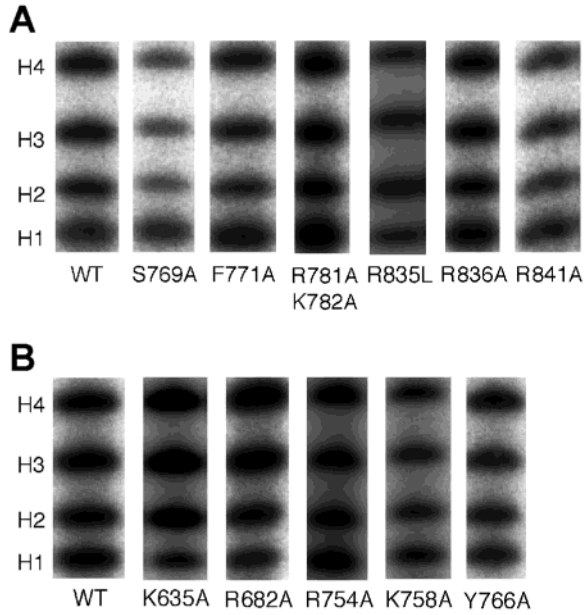

FIGURE 6: Competition assay with wild-type and mutant Klenow fragment derivatives. Competition was assessed with the concentrations of oligonucleotides H1−H4 balanced such that wild-type Klenow fragment gave approximately equal incorporation of [α-^{32}P]-dTTP into all four substrates. Samples were fractionated on a 10% polyacrylamide−urea gel, and the data were calculated as described in Table 5. Panel A shows data for mutations in the proposed binding site for the template overhang. Panel B contains data for mutations elsewhere in the Klenow fragment structure, specifically in the duplex DNA binding site (K635A and R682A) and in the nascent base pair binding pocket (R754A, K758A, and Y766A).

incorporation of [α-^{32}P]dTTP into all four. Specifically, since the relative reactivities of H1:H2:H3:H4 were ≈1:10:30:90, the concentrations were adjusted to reflect the inverse ratio. Figure 6 shows that wild-type Klenow fragment, as intended, gave approximately equal incorporation of radioactivity into all four bands, whereas some of the mutant proteins gave substantial deviations from this. Most notable was S769A, where the extent of labeling of H2 was visibly lower than that of H1. This is easily explained because S769 is expected to interact with the phosphate 5′ to the templating base, which is present in H2 and longer oligonucleotides but absent in H1; as a result, the mutation should be detrimental to the binding of H2 but not to the binding of H1. The dTTP incorporation into adjacent pairs of bands was compared for each mutant protein, and normalized by dividing by the same ratio for wild-type Klenow fragment (Table 5). Aside from R835L, all of the mutations in the putative template-binding region resulted in a decrease in the preference for H2 over H1, but had very little effect on the relative affinities of H2−H4. No other category of mutant proteins that we tested showed the same relative bias toward H1. Mutations in side chains that form the binding site for the nascent base pair (R754A, K758A, and Y766A) had very little effect on the outcome of the competition assay, while mutations in residues that contact the DNA duplex upstream of the primer terminus (K635A and R682A) resulted in a bias in favor of longer substrates so that, in every case, the ratio of H_{n+1}/H_n was greater for the mutant protein than for the wild type. The R835L mutation had an atypical phenotype; compared with the wild type, R835L exhibited an enhanced preference

Table 5: Effect of Template Overhang Length on Substrate Preference by Mutant Derivatives of Klenow Fragment[a]

	ratio of dTTP incorporation[b]		
enzyme[a]	H2:H1	H3:H2	H4:H3
wild type	1	1	1
Mutations in the Putative Template Overhang Binding Region			
S769A	0.35	0.80	1.1
F771A	0.54	0.97	1.2
R781A/K782A	0.60	0.83	0.92
R835L	2.1	0.82	0.67
R836A	0.68	0.85	0.96
R841A	0.63	0.81	0.85
Mutations within the Nascent Base Pair Binding Pocket			
R754A	0.84	1.2	1.1
K758A	1.2	1.0	1.1
Y766A	1.2	1.2	1.3
Mutations in the Duplex DNA Binding Region			
K635A	1.9	1.9	1.7
R682A	1.3	1.7	1.7

[a] All proteins in this study carry the D424A (exonuclease-deficient) mutation. [b] Calculated as $(H_{n+1}/H_n)_{mutant}$ divided by $(H_{n+1}/H_n)_{WT}$ for [α-^{32}P]dTTP incorporation into each pair of hairpin oligonucleotides (see Figure 1). The values given are averages of three determinations, which were in close agreement.

for H2 over H1, but a weakened preference for H3 over H2, or H4 over H3.

Cross-Linking with Mutant Klenow Fragment Derivatives. Small-scale cross-linking experiments were carried out to measure the effect of mutations in the proposed template-binding region on cross-linking to BrdU at the 0, +1, or +2 position in both binary and ternary complexes (Figure 7). With the photoagent at the +1 and +2 positions, the cross-linking yields were very similar for all but one of the mutant proteins in both binary and ternary complexes. The exception was F771A, which gave a much lower cross-linking yield from both ternary complexes. With BrdU at the 0 position, wild-type Klenow fragment and most of the mutant proteins gave ∼5-fold less cross-linking from the ternary complex than from the binary complex. In contrast, Y766A, S769A, and R841A gave similar yields from binary and ternary complexes. Figure 7 shows results obtained using our standard conditions of 20 μM complementary dNTP, but this behavior was seen even at dNTP concentrations of 400 μM, showing that it could not be due merely to a failure to bind dNTP and form a ternary complex (data not shown).

Analysis of Sequence Conservation. To aid in understanding which residues in the fingers subdomain of Klenow fragment may have important functions, we analyzed sequence conservation in the fingers subdomains of 50 bacterial DNA polymerase I enzymes (Table 6 of the Supporting Information). Approximately half of the residues in the fingers subdomain show at least 80% conservation. Of these conserved residues, almost half play an obvious role in maintaining the protein structure, with the majority, mostly hydrophobic, contributing to the buried core of the subdomain. A small number of highly conserved glycine or proline side chains are seen at the boundaries of secondary structure elements, and an invariant buried polar arginine (R821) interacts with two conserved carboxylates so as to link the fingers, palm, and exonuclease subdomains. The remaining half of the conserved amino acids contribute to the surface of the fingers subdomain.

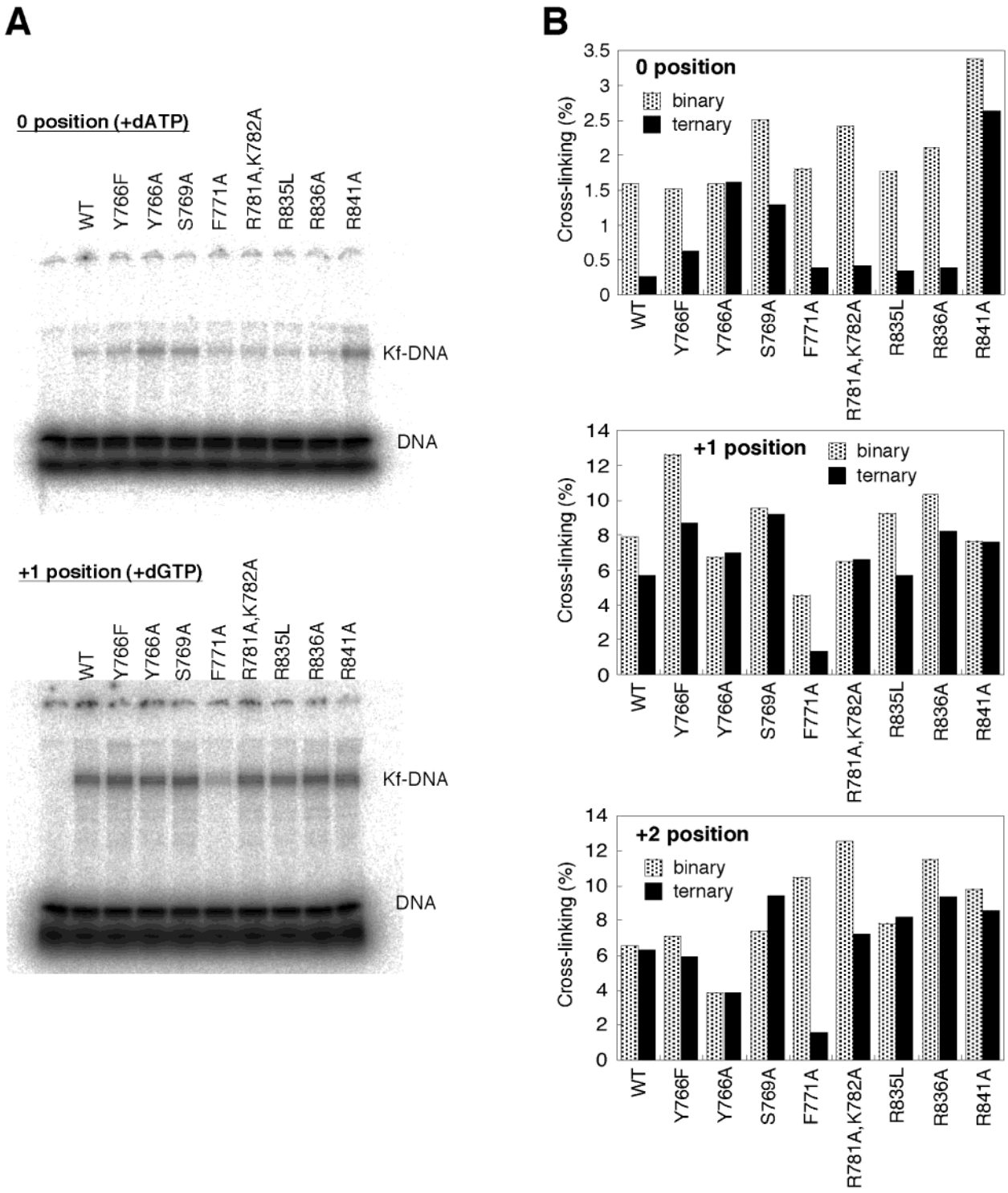

FIGURE 7: Photochemical cross-linking of labeled DNA to mutant derivatives of Klenow fragment. Panel A shows two SDS−polyacrylamide gels, fractionating cross-linked protein from free DNA. The first gel shows the results of cross-linking ternary complexes of polymerase with the ^{32}P-labeled P−T_0Br duplex and dATP; the second is from cross-linking ternary complexes with the P−T_1Br duplex and dGTP. Panel B shows the quantitation of a series of gels analyzing cross-linking from the P−T_0Br, P−T_1Br, and P−T_2Br duplexes in binary and ternary complexes. The cross-linking conditions were those described in Methods for analytical-scale reactions. Similar results were obtained with a 10-fold higher concentration of Klenow fragment, implying that the variations in cross-linking yield were not caused by differences in DNA binding affinity of the mutant proteins.

DISCUSSION

Although they did not define the path of the single-stranded template beyond the polymerase active site, structural studies indicated the likelihood that some part of this strand would contact the polymerase, probably on the surface of the fingers subdomain. The competition binding assay with wild-type Klenow fragment showed the most important contacts to be at the first four template bases, with a 6-fold increase in binding affinity as the template overhang is increased from one to two nucleotides, and 3-fold increases for addition of the next two template bases. In contrast, the next four nucleotides (increasing from a four- to an eight-

nucleotide 5′ extension) contribute only a 4-fold increase in binding affinity. Gel shift measurements with 5′-phosphorylated and nonphosphorylated DNAs suggested that the phosphates on either side of the +1 template position make the most significant contribution to binding, subject to the caveat that a terminal phosphate is not a perfect model for a backbone phosphodiester. In the competition assay, the largest increment in binding affinity for addition of a single nucleotide to the template strand was 6-fold, corresponding to a binding energy of 1.1 kcal/mol; thus, binding of the template strand is the aggregate of several relatively weak interactions.

Cross-linking with photoprobes at the templating base (0 position) and the two bases immediately 5′ to the templating base (+1 and +2 positions) places this part of the template strand on the surface of the fingers subdomain between Y766 and F771. Cross-linking experiments with mutant proteins confirm the assignment of F771 as a major cross-linking target in ternary complexes with BrdU at the +1 or +2 template position. The evidence supporting this conclusion is the much lower cross-linking yield from ternary complexes of the F771A polymerase (Figure 7). Interestingly, the F771A mutation had little or no effect on cross-linking yields in the binary complex, implying that the template strand in a binary complex is more free to move around and can cross-link to a variety of sites on the polymerase. Meisenheimer and Koch have documented the tendency of halopyrimidine photoprobes to react preferentially with aromatic side chains and have noted that, where cocrystal structures are available, efficient cross-linking may be correlated with a stacking interaction between the halopyrimidine and an aromatic side chain (*24*). In the binary, but catalytically active, complex of *Bst* DNA polymerase, the Y719 side chain (equivalent to F771 of Klenow fragment) is stacked against the +1 template base (*11*); however, the ternary complex of Klentaq does not show evidence of a stacking interaction between the +1 base and the corresponding H676 side chain.

When the photoactivatable base is at the 0 position, the wild type and most of the mutant Klenow fragment derivatives gave low cross-linking yields from the ternary complex, consistent with the expectation that the templating base would be buried in the binding pocket, with the Br substituent on the major groove side at the top of the pocket, while Y766, a likely target, would be positioned low down in the binding pocket (*9, 10*). The higher cross-linking yield from the binary complex, and the failure of any of the mutants we have tested to diminish this yield, suggest that the binary complex is more flexible than the ternary complex, allowing the 0 position photoagent to cross-link to more than one site on the protein. Both of the observed locations for the 0 position template base, in binary complexes of Klentaq (folded back over its 3′ neighbor) and *Bst* pol (inserted into a pocket between the O and O1 helices), place it within range of the target residues inferred from peptide sequencing (Y766, M768, and S769) (*10*−*12*). Interestingly, mutation of Y766, S769, or R841 resulted in cross-linking yields from a ternary complex with a 0 position photoagent that were abnormally high and more similar to the yields seen with binary complexes. All three mutations are predicted to affect contacts involved in positioning the templating base, and the cross-linking yields suggest that the absence of these contacts prevents the templating base from becoming properly buried within the binding pocket, even in the presence of dNTP concentrations high enough to ensure ternary complex formation.

In contrast to the cross-linking from the 0, +1, and +2 positions, the +3 position is presumably out of range of any suitably reactive side chain and the +4 position may not occupy a unique well-defined position on the protein surface.

Because our experimental design connects individual positions on the DNA with particular protein side chains, our results provide positional information that was largely absent from two previous photo-cross-linking investigations on Klenow fragment−DNA complexes which identified the same cross-linked tryptic peptide, A759−R775, as we report here. In the first study (*26*), Catalano et al. observed cross-linking to Y766 using an aryl azide attached via a flexible tether to the primer terminal base. Because the azide group could be located up to 10 Å from the primer 3′-OH group, the data are compatible with our conclusion that Y766 is close to the 0 position templating base. In the second study (*27*), Pandey et al. identified Y766, F771, I765, and S769 as likely sites of cross-linking; however, their data provided no positional information for the DNA because the cross-linking relied on the naturally occurring DNA bases, and it was therefore not possible to deduce which part of the DNA substrate was involved in the cross-link. Our data, using single photoprobes at unique locations in each DNA substrate, indicate that the DNA sites involved in the cross-links observed by Pandey et al. were most probably the templating base and its two 5′ neighbors. In contrast, a third cross-linking study is at odds with all the data described above. This study used a nucleotide derivatized with an aryl azide photoprobe, intended to target the template binding site, which became cross-linked to a region of the protein now known to be involved in binding to the incoming dNTP (*28*).

Mutational analysis of side chains within the likely contact region provides evidence for the functional importance of interactions inferred from cocrystal structures of family A DNA polymerases. Of the side chains that have been mutated in this study, R841 is the closest to the active site region, and R841A was the only mutation that had a significant effect on k_{cat}. In the Klentaq and *Bst* pol cocrystal structures, the side chains equivalent to R841 interact with the phosphodiester 3′ to the templating (0 position) base and with the sugar of the +1 base. Our DNA binding data for the R841A Klenow fragment indicate that the loss of these interactions results in a 9-fold decrease in DNA binding affinity, and the additional effect on k_{cat} suggests that the R841 interaction plays a role in the correct positioning of substrates in the transition state. The smaller decreases in DNA binding affinity that resulted from mutation of S769 and F771 can also be correlated with interactions seen in cocrystal structures of family A polymerases. In the Klentaq ternary complex, the homologue of S769 interacts with the phosphodiester 5′ to the templating base (*10*), and one of the *Bst* pol binary complex structures suggests a stacking interaction between F771 and one of the template bases (*11*). Studies on the partitioning of a DNA primer terminus between polymerase and 3′−5′ exonuclease sites of Klenow fragment suggested that R835 and R836 destabilize binding at the polymerase site, perhaps contributing to strain that promotes discrimination against mismatched primer termini (*29*).

Consistent with this idea, the R836A mutation caused a small increase in DNA binding affinity; however, the R835L mutation had little or no effect.

By carrying out the competition assay with Klenow fragment derivatives having mutations in the proposed template contact region, we hoped to determine which positions on the DNA substrate interact with particular side chains. The clearest example was provided by S769A, where the loss of the predicted interaction between S769 and the phosphate 5′ to the templating base accounts for the shift in preference (relative to the wild-type enzyme) toward the shortest substrate, which has a 5′ overhang of only a single nucleotide and therefore lacks this phosphate. Quantitatively, S769A caused the largest change in the competition assay, showing that the S769−DNA contact predicted from the family A polymerase crystal structures is energetically significant and probably accounts for most of the 6-fold increase in binding affinity seen when comparing the H2 and H1 substrates. Similar, though smaller, shifts in preference toward the shortest substrate resulted from the other mutations in the putative template binding region, with the exception of R835L (discussed below). As with S769A, the effects of R841A and F771A can be explained by interactions with the +1 position inferred from cocrystal structures, R841 with the sugar and F771 with the base. However, the changes were similar in magnitude to those resulting from the R836A and R781A/K782A mutations, where the altered side chains are too far away to interact with the +1 position. We should therefore consider the possibility that some mutations in the putative template binding region may indirectly affect the +1 position and diminish the degree of discrimination against the substrate with only a single nucleotide 5′ overhang. The tendency of polymerases with mutations in the proposed template binding region to favor the shortest substrate appears to be a characteristic of this class of mutations. In contrast, mutations within the binding site for the nascent base pair had no effect on the competition assay, whereas mutations in side chains that bind the primer−template duplex exaggerated the preference for DNA molecules with longer template overhangs. In the latter case, the loss of an important DNA binding contact in the duplex region may increase the relative importance of contacts to the single-stranded template.

The phenotype of the R835L mutant in the competition assay was unlike that of any of the other proteins with mutations in the proposed template binding region. Relative to wild-type Klenow fragment, R835L showed a stronger preference for the H2 over the H1 substrate, contrasting with the stronger preference for H1 showed by the other mutant proteins. As mentioned above, Thompson et al. concluded that the R835 side chain destabilizes or strains a DNA substrate bound at the polymerase site (*29*), and this may provide an explanation for the behavior of the R835L mutant protein. The location of the R835 side chain is such that any interaction with DNA bound at the polymerase site must involve the part of the template that is in duplex at or close to the primer terminus. We suggest that the putative destabilizing effect of the R835 side chain could result from nonoptimal positioning of a second DNA contact in the single-stranded template. If this contact is present in the H2 oligonucleotide but not in H1 (the phosphate that interacts with S769 would be an obvious example), then the effect of the R835L mutation on the competition between the H2 and H1 oligonucleotides can be explained. The absence of the interaction with R835 would improve binding to, for example, S769; this would enhance the binding of H2, but have no effect on the binding of H1 which lacks the proposed DNA contact. In the competition assay, the effect of the R835L mutation on the H4:H3 ratio was different from that of the other mutant proteins in our study, suggesting that R835L might cause the single-stranded template to follow a different path on the fingers subdomain.

The mutational studies confirm the expectation that the template overhang is bound on the surface of the fingers subdomain and provide direct evidence for the contributions of S769, R841, F771, and R835, inferred from structural and other studies. Quantitatively, the effects of each mutation are small, as expected from our analysis of the binding of wild-type Klenow fragment to the template overhang. This is consistent with the idea of a number of protein−DNA contacts, each making a relatively small energetic contribution. It is puzzling that the competition assay suggested that both R836 and the R781-K782 region contribute to DNA binding, since they are separated by almost 20 Å. We believe R836 is more likely to play an important role for the following reasons. First, the R836A mutation affected overall DNA binding (this work and ref *29*), whereas the R781A/K782A mutation, although it removed two positively charged side chains, had no detectable effect on $K_{d(DNA)}$. Second, the R836 side chain is highly conserved in bacterial DNA polymerase I sequences (96% have R or K in this position), whereas R781 and K782 are very poorly conserved (48 and 34% basic residues, respectively). Finally, the positively charged surface region around the position of R836 or its homologues is present in all four family A DNA polymerase structures, whereas the second patch of positive charge around R781 and K782 in Klenow fragment has no equivalent in the other structures. If R781 and K782 are not themselves involved in binding the single-stranded template, then the effect on the competition assay of removing these side chains may be indirect. It is certainly possible that the substitution of Ala for these two large side chains may affect the position of the neighboring F771 side chain, which interacts with the template overhang.

As implied above, sequence conservation provides additional circumstantial evidence concerning the importance of particular side chains in the fingers subdomain. Most of the side chains described as making contacts with the single-stranded template (S769, F771, R836, and R841) are well-conserved in an alignment of 50 bacterial DNA polymerase I sequences (Table 6 of the Supporting Information). In the more distantly related T7 DNA polymerase, the level of conservation of these side chains is poorer, but analogous interactions with the single-stranded template are often present. For example, the phosphodiester 5′ to the templating base interacts with a main chain amide nitrogen instead of the Ser side chain seen in the bacterial DNA polymerase I enzymes (*9*). In Klenow fragment, the highly conserved R836 side chain, together with the less conserved R806, K810, and K829 residues, forms a patch of positive charge the approximate position of which is maintained in all of the four Pol I family crystal structures (data not shown), even though the individual side chains contributing to the positive charge may not always be highly conserved. Curiously,

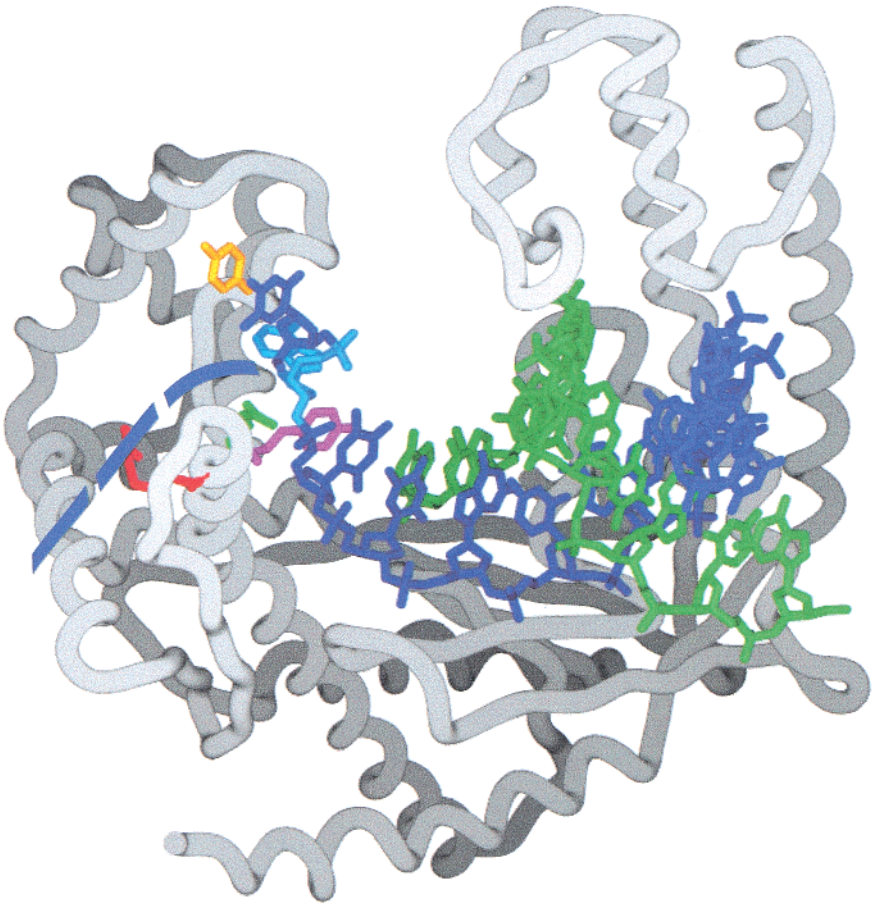

FIGURE 8: Binding site for the single-stranded template inferred from this study, illustrated using the coordinates of a complex of *Bst* DNA polymerase with duplex DNA (PDB entry 4BDP) (*11*). The protein backbone of the polymerase domain is shown, and four side chains, whose proposed interactions with DNA are discussed in the text, are illustrated. The side chains (with the equivalent Klenow fragment side chains in parentheses) are Y714 (Y766) in magenta, Y719 (F771) in gold, R784 (R836) in red, and R789 (R841) in green. The DNA primer strand is in green and the template strand in dark blue, with the exception of the 0 position which is illustrated in cyan for clarity. The 0 position base is inserted into a pocket in the protein between the O and O1 helices, and the +1 position base is stacked with Y719. The proposed path of the template strand beyond the +1 position is illustrated by the segmented ribbon, where each segment is approximately equivalent to the distance spanned by one nucleotide. Thus, this model predicts that the polymerase interacts with approximately five residues of the single-stranded template.

mutation of the nonconserved R835 gave interesting phenotypes in both the competition assay described here and the binding studies of Thompson et al. (*29*). It is possible either that different side chains perform the equivalent function in other Pol I enzymes or that replacement of the polar Arg with Leu causes local structural adjustments that reposition another side chain that interacts with the template, giving rise to the observed phenotypes.

In conclusion, the experiments described here suggest that the single-stranded template follows the approximate path indicated in Figure 8 and is held in place by a number of relatively weak interactions with the protein. The weakness of these interactions may account for the failure of cocrystal structures to reveal this binding site, and for the substantial differences in the position of the +1 base observed in ternary complexes of Klentaq with different DNA sequences (*12*). It is also possible, as suggested in a study of HIV-1 reverse transcriptase (*30*), that there may not be a unique location for the single-stranded template. For an enzyme such as bacterial DNA polymerase I, the role of which in vivo is to fill gaps and remove the RNA primers of Okazaki fragments, different binding contacts might be used in different phases of gap filling. Thus, at the start of gap filling, there could be a substantial stretch of single-stranded DNA beyond the point of synthesis, but as gap filling nears completion (or during nick translation), the polymerase will encounter double-stranded DNA close to the site of synthesis. Interestingly, the location we have suggested for the single-stranded template is very similar to the position of the downstream duplex DNA in an elongating complex of T7 RNA polymerase (*31*). Given the strong structural similarity between the polymerase domains of T7 RNA polymerase and the family A DNA polymerases (*32*), this suggests that our proposed path (Figure 8) may also be able to accommodate duplex DNA.

ACKNOWLEDGMENT

We are grateful to Xiaojun Chen Sun, Lixing Liu, and Olga Potapova for excellent technical support, to William Konigsberg, Kenneth Williams, and Myron Crawford for advice on cross-linking and peptide sequencing, to Fred Richards and Gerry Olack for the use of the HeCd laser, and to Whitney Yin and Tom Steitz for helpful discussions and for communicating results prior to publication.

SUPPORTING INFORMATION AVAILABLE

Analysis of sequence conservation on the fingers subdomain of 50 bacterial DNA polymerase I homologues (Table 6) and DNA binding affinity of S769A Klenow fragment (Figure 9). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Kornberg, A., and Baker, T. (1992) *DNA Replication*, 2nd ed., W. H. Freeman and Co., San Francisco.
2. Brautigam, C. A., and Steitz, T. A. (1998) Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes, *Curr. Opin. Struct. Biol. 8*, 54−63.
3. Jäger, J., and Pata, J. D. (1999) Getting a grip: polymerases and their substrate complexes, *Curr. Opin. Struct. Biol. 9*, 21−28.
4. Steitz, T. A. (1999) DNA polymerases: structural diversity and common mechanisms, *J. Biol. Chem. 274*, 17395−17398.
5. Steitz, T. A., Smerdon, S. J., Jäger, J., and Joyce, C. M. (1994) A unified polymerase mechanism for nonhomologous DNA and RNA polymerases, *Science 266*, 2022−2025.
6. Holm, L., and Sander, C. (1995) DNA polymerase β belongs to an ancient nucleotidyltransferase superfamily, *Trends Biochem. Sci. 20*, 345−347.
7. Beese, L. S., Derbyshire, V., and Steitz, T. A. (1993) Structure of DNA polymerase I Klenow fragment bound to duplex DNA, *Science 260*, 352−355.
8. Eom, S. H., Wang, J., and Steitz, T. A. (1996) Structure of *Taq* polymerase with DNA at the polymerase active site, *Nature 382*, 278−281.
9. Doublié, S., Tabor, S., Long, A. M., Richardson, C. C., and Ellenberger, T. (1998) Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution, *Nature 391*, 251−258.
10. Li, Y., Korolev, S., and Waksman, G. (1998) Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of *Thermus aquaticus* DNA polymerase I: structural basis for nucleotide incorporation, *EMBO J. 17*, 7514−7525.
11. Kiefer, J. R., Mao, C., Braman, J. C., and Beese, L. S. (1998) Visualizing DNA replication in a catalytically active *Bacillus* DNA polymerase crystal, *Nature 391*, 304−307.
12. Li, Y., and Waksman, G. (2001) Crystal structures of a ddATP-, ddTTP-, ddCTP, and ddGTP-trapped ternary complex of Klentaq1: insights into nucleotide incorporation and selectivity, *Protein Sci. 10*, 1225−1233.
13. Franklin, M. C., Wang, J., and Steitz, T. A. (2001) Structure of the replicating complex of a pol α family DNA polymerase, *Cell 105*, 657−667.
14. Jacobo-Molina, A., Ding, J., Nanni, R. G., Clark, A. D., Jr., Lu, X., Tantillo, C., Williams, R. L., Kamer, G., Ferris, A. L., Clark, P., Hizi, A., Hughes, S. H., and Arnold, E. (1993) Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA, *Proc. Natl. Acad. Sci. U.S.A. 90*, 6320−6324.
15. Huang, H., Chopra, R., Verdine, G. L., and Harrison, S. C. (1998) Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance, *Science 282*, 1669−1675.
16. Ling, H., Boudsocq, F., Woodgate, R., and Yang, W. (2001) Crystal structure of a Y-family DNA polymerase in action: a mechanism for error-prone and lesion-bypass replication, *Cell 107*, 91−102.
17. Puglisi, J. D., and Tinoco, I., Jr. (1989) Absorbance melting curves of RNA, *Methods Enzymol. 180*, 304−325.
18. Polesky, A. H., Steitz, T. A., Grindley, N. D. F., and Joyce, C. M. (1990) Identification of residues critical for the polymerase activity of the Klenow fragment of DNA polymerase I from *Escherichia coli*, *J. Biol. Chem. 265*, 14579−14591.
19. Joyce, C. M., and Derbyshire, V. (1995) Purification of *Escherichia coli* DNA polymerase I and Klenow fragment, *Methods Enzymol. 262*, 3−13.
20. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, *Anal. Biochem. 72*, 248−254.
21. Derbyshire, V., Freemont, P. S., Sanderson, M. R., Beese, L., Friedman, J. M., Joyce, C. M., and Steitz, T. A. (1988) Genetic and crystallographic studies of the 3′,5′-exonucleolytic site of DNA polymerase I, *Science 240*, 199−201.
22. Astatke, M., Grindley, N. D. F., and Joyce, C. M. (1995) Deoxynucleoside triphosphate and pyrophosphate binding sites in the catalytically competent ternary complex for the polymerase reaction catalyzed by DNA polymerase I (Klenow fragment), *J. Biol. Chem. 270*, 1945−1954.
23. Minnick, D. T., Astatke, M., Joyce, C. M., and Kunkel, T. A. (1996) A thumb subdomain mutant of the large fragment of *Escherichia coli* DNA polymerase I with reduced DNA binding affinity, processivity, and frameshift fidelity, *J. Biol. Chem. 271*, 24954−24961.
24. Meisenheimer, K. M., and Koch, T. H. (1997) Photocross-linking of nucleic acids to associated proteins, *Crit. Rev. Biochem. Mol. Biol. 32*, 101−140.
25. Alekseyev, Y. O., Dzantiev, L., and Romano, L. J. (2001) Effects of benzo[*a*]pyrene DNA adducts on *Escherichia coli* DNA polymerase I (Klenow fragment) primer-template interactions: evidence for inhibition of the catalytically active ternary complex formation, *Biochemistry 40*, 2282−2290.
26. Catalano, C. E., Allen, D. J., and Benkovic, S. J. (1990) Interaction of *Escherichia coli* DNA polymerase I with azidoDNA and fluorescent DNA probes: identification of protein-DNA contacts, *Biochemistry 29*, 3612−3621.
27. Pandey, V. N., Kaushik, N., and Modak, M. J. (1994) Photoaffinity labeling of DNA template-primer binding site in *Escherichia coli* DNA polymerase I. Identification of involved amino acids, *J. Biol. Chem. 269*, 21828−21834.
28. Moore, B. M., II, Jalluri, R. K., and Doughty, M. B. (1996) DNA polymerase photoprobe 2-[(4-azidophenacyl)thio]-2′-deoxyadenosine 5′-triphosphate labels an *Escherichia coli* DNA polymerase I Klenow fragment substrate binding site, *Biochemistry 35*, 11642−11651.
29. Thompson, E. H. Z., Bailey, M. F., van der Schans, E. J. C., Joyce, C. M., and Millar, D. P. (2002) Determinants of DNA mismatch recognition within the polymerase domain of the Klenow fragment, *Biochemistry 41*, 713−722.
30. Peletskaya, E. N., Boyer, P. L., Kogon, A. A., Clark, P., Kroth, H., Sayer, J. M., Jerina, D. M., and Hughes, S. H. (2001) Cross-linking of the fingers subdomain of human immunodeficiency virus type 1 reverse transcriptase to template-primer, *J. Virol. 75*, 9435−9445.
31. Yin, Y. W., and Steitz, T. A. (2002) Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase, *Science 298*, 1387−1395.
32. Sousa, R., Chung, Y. J., Rose, J. P., and Wang, B. C. (1993) Crystal structure of bacteriophage T7 RNA polymerase at 3.3 Å resolution, *Nature 364*, 593−599.
33. Antao, V. P., and Tinoco, I. (1992) Thermodynamic parameters for loop formation in RNA and DNA hairpin tetraloops, *Nucleic Acids Res. 20*, 819−824.
34. Brautigam, C. A., and Steitz, T. A. (1998) Structural principles for the inhibition of the 3′-5′ exonuclease activity of *Escherichia coli* DNA polymerase I by phosphorothioates, *J. Mol. Biol. 277*, 363−377.
35. Christopher, J. A. (1998) *SPOCK: the structural properties observation and calculation kit (program manual)*, The Center for Macromolecular Design, Texas A&M University, College Station, TX.

BI026566C