Cocrystal structure of an editing complex of Klenow fragment with DNA

(3'-5' exonuclease/DNA polymerase/protein-DNA interaction/x-ray crystallography/metal ion catalysis)

P. S. Freemont*, J. M. Friedman, L. S. Beece, M. R. Sanderson†, and T. A. Steitz

Department of Molecular Biophysics and Biochemistry and the Howard Hughes Medical Institute at Yale University, New Haven, CT 06511

Communicated by Julian M. Sturtevant, August 26, 1988 (received for review July 21, 1988)

ABSTRACT High-resolution crystal structures of editing complexes of both duplex and single-stranded DNA bound to Escherichia coli DNA polymerase I large fragment (Klenow fragment) show four nucleotides of single-stranded DNA bound to the 3'-5' exonuclease active site and extending toward it. Two divalent metal ions interacting with the phosphodiesters are hydrolyzed to catalyze exonuclease reaction by a mechanism that may be related to mechanisms of other enzymes that catalyze phospho-group transfer including RNA enzymes. We suggest that the editing active site competes with the polymerase active site some 30 Å away for the newly formed 3' terminus. Since a 3' terminal mismatched base pair favors the melting of duplex DNA, its binding and excision at the editing exonuclease site that binds single-stranded DNA is enhanced.

The large proteolytic fragment (Klenow fragment) of Escherichia coli DNA polymerase I utilizes an editing 3'-5' exonuclease activity (1) to reduce the misincorporation of erroneous nucleotides by about 10-fold (2) at an active site that is some 30 Å away from the polymerase site of misincorporation (3). How might this be accomplished? The crystal structure of the Klenow fragment shows that it is folded into two domains (3). Various experiments (reviewed in ref. 4) establish that the dNMP binds in the crystal catalyzes the 3'-5' exonuclease activity, whereas the larger C-terminal domain contains the active site for the polymerase reaction. Mutant proteins that contain amino acid changes in the dNMP binding site have been made by directed mutagenesis; they are devoid of exonuclease activity but retain full polymerase activity (5). Furthermore, the DNA encoding the C-terminal domain has been cloned, and the product has been expressed, isolated, and shown to possess significant DNA polymerase activity with no measurable 3'-5' exonuclease activity (6). The observation (3) that these two active sites are ~25-30 Å apart poses the interesting question of how they work together to achieve high-fidelity synthesis of DNA.

The C-terminal domain contains a cleft that is large enough to accommodate the double-stranded B-DNA product of DNA synthesis (3). The approximate position of the 3' terminus of the primer strand has been derived from the cross-linking of 8-azido-dATP to Tyr-776, footprinting of Klenow fragment on DNA (7), and the position of site-directed mutants that alter polymerase activity but not exonuclease activity (A. Polesky and C. Joyce, personal communication). This model of DNA at the polymerase active site places about 8 base pairs (bp) of duplex product DNA in the cleft.

A more detailed understanding of the structural basis of the polymerase and exonuclease activities requires the separate determination of the crystal structures of suitable DNAs complexed with each of these two active sites on the Klenow fragment. We report here the high-resolution crystal structure analyses of three different complexes between DNA and the 3'-5' exonuclease active site. The structures of these complexes suggest a possible mechanism for the exonuclease reaction and provide some insight into how these two separate active sites might work together to enhance the accuracy of DNA synthesis.

EXPERIMENTAL METHODS

Structural Analysis. The first crystals of polymerase-DNA complex examined were obtained by cocrystallizing the Klenow fragment in the presence of a 2-fold molar excess of an 8-bp duplex DNA containing a 3-base single-stranded 5' overhang (5' AGACGGCCCGG
GCGGGGCC/)

These crystals were grown from solutions containing 38% (wt/vol) saturated ammonium sulfate and 1 mM EDTA, which inhibits the 3'-5' exonuclease activity. In spite of the crystallization at high ionic strength, HPLC analysis of the crystals suggested a one-to-one stoichiometry of duplex DNA to protein as well as the presence of equimolar amounts of each strand (data not shown). The cocrystals are tetragonal space group P4₁ (a = b = 104.7 Å, c = 86.0 Å) and are isomorphous with the crystals of a native Klenow fragment (3) except for a 1-Å increase in the length of the a/b axes.

Two other complexes of the Klenow fragment with DNA were prepared by diffusing single-stranded deoxytetranucleotides into crystals of two different mutant proteins (5) that are devoid of 3'-5' exonuclease activity. Crystals of D424A (Asp-424 to Ala) and crystals of D355A, E357A (Asp-355 to Ala; Glu-357 to Ala) were soaked in solutions containing 1 mM p[dT]₄ for a period of 1 week. X-ray diffraction data were measured from crystals of all three of these DNA complexes to high resolution by using a Mark II two-dimensional position-sensitive area detector (8) (Table 1).

RESULTS

Editing Complex with Duplex DNA. Although the duplex DNA substrate was expected to form a complex at the polymerase active site, it binds to the 3'-5' exonuclease active site under the high ionic strength conditions of the crystallization. The 3.5-Å resolution difference electron density map between the DNA complex and the metal-free

Table 1. Data collection and soaking

<table>
<thead>
<tr>
<th>Crystals*</th>
<th>Maximum resolution, Å</th>
<th>Unique reflections, no.</th>
<th>Rmerge, %</th>
<th>Isomorphous differences, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Klenow (citrate)</td>
<td>2.8</td>
<td>18,900</td>
<td>4.1</td>
<td>11.4</td>
</tr>
<tr>
<td>Klenow (SAS)</td>
<td>3.3</td>
<td>11,100</td>
<td>3.3</td>
<td>—</td>
</tr>
<tr>
<td>Klenow cocrystals</td>
<td>3.8</td>
<td>6,500</td>
<td>4.1</td>
<td>19.6 (19.4)</td>
</tr>
<tr>
<td>D242A/p(DT)</td>
<td>2.9</td>
<td>16,200</td>
<td>5.1</td>
<td>15.1</td>
</tr>
<tr>
<td>D355A,E357A/p(DT)</td>
<td>3.1</td>
<td>13,700</td>
<td>5.6</td>
<td>14.5</td>
</tr>
</tbody>
</table>

*Klenow DNA cocrystals and Klenow (citrate) crystals were soaked in 60% saturated ammonium sulfate/1 mM EDTA/200 mM citrate, pH 5.6. Klenow (SAS) crystals were soaked in 70% saturated ammonium sulfate/1 mM ZnSO4/20 mM MgSO4/66 mM Pipes, pH 7.0 (SAS buffer). The D242A and D355A,E357A mutant crystals were soaked for 1 week in SAS buffer containing 1 mM p(dT)₄.

Maximum resolution was chosen as the resolution where (I)ₘᵦₑᵦ = 2(σ(I) + σ(I)ₘᵦₑᵦ)/2.

Rmerge = Σₖ(Σᵢ[Iᵢ] - [Σᵢ[Iᵢ]])/Σᵢ[Iᵢ], where Iᵢ is intensity of the kth observation of reflection h.

The percent isomorphous differences were calculated against Klenow (SAS) between 20- and 3.3-Å resolution (20-3.8 Å for the Klenow cocrystal). ∆SAD = SAD - SAD + SAD, where SAD and SAD are the observed diffraction intensities of the complexed and uncomplexed protein, respectively.

Calculated between Klenow cocrystal and Klenow (citrate), 20- to 3.8-Å resolution.

Klenow fragment shows electron density arising from three to four bases of ordered single-stranded DNA bound at the 3'-5' exonuclease active site and extending toward the cleft believed to bind duplex DNA (Fig. 1 Upper). The protein appears to have at least partially denatured the DNA and bound to the 3' terminal sequence since the difference electron density can be fitted by the 3' terminal sequence but not the sequence of the 5' single-stranded overhang with its associated 5' phosphate. Numerous conformational changes are seen in the protein, prominent among these being a movement of two helices (helices I and H) (3) that make up one side of the DNA binding cleft.

It appears, therefore, that we have cocrystallized and determined the structure of an editing complex with DNA. It is not certain why the rest of the DNA does not appear in this difference electron density map. It may be partially disordered. Furthermore, the difference map has a high noise level due to the significant differences in structure between the complexed and uncomplexed enzyme.

Single-Stranded DNA Complex. Because it was clear from the complex with duplex DNA that the exonuclease active site contains a binding site for single-stranded DNA, p(dT)₄ was diffused into preformed crystals of two different mutant proteins deficient in 3'-5' exonuclease activity. High-resolution crystallographic analyses of these mutant proteins

Fig. 1. (Upper) Portion of the 3.8-Å resolution difference electron density map between the cocrystallized DNA complex and the metal-free native protein superimposed on the structure of the Klenow fragment. The structure factor amplitudes were measured from crystals soaked in 60% saturated ammonium sulfate/1 mM EDTA/200 mM citrate, pH 6.5. Electron density for 3 or 4 nucleotides is seen emanating from 3'-5' exonuclease active site and extending towards the large cleft in the protein. (Lower) Model of the single-strand DNA sequence (dT₄) fitted into a 2.8-Å resolution difference electron density map between the (dT₄) complex with D242A mutant protein and free native wild-type protein containing Zn²⁺ calculated by using F_dna complex − F_native as coefficients. The structure factor amplitudes F_dna complex were measured from crystals soaked in 1 mM p(dT)₄/70% saturated ammonium sulfate/1 mM ZnSO₄/20 mM MgSO₄/66 mM Pipes, pH 7.0. The phases used in the calculation were experimentally derived isomorphous replacement phases between 20- and 8.0-Å resolution and calculated phases from 8.0- to 2.9-Å resolution. The positive electron density is contoured at a level of 4σ where σ is the root-mean-square density throughout the unit cell. The α-carbon backbone of the protein is superimposed. Only three nucleotides are shown. Although electron density for the fourth nucleotide is weak in this difference map, it is significantly stronger in a difference map in which data from the dTMP complex with the D242A mutant Klenow fragment are used as the parent. The two spheres are at the positions of the two divalent metal ions found in the dTMP complex. The high-resolution difference electron density maps of the Klenow–DNA cocrystal and the tetranucleotide complex with the D355A, E357A mutant Klenow fragment indicate a similar position for the tetranucleotide.
complexed with dNMP and divalent metals showed no major structural changes (5). The loss of exonuclease activity seems to be a direct consequence of the loss of the metal ion at site B in the D424A mutant and both metal ions in the D355A, E357A mutant protein. The high-resolution difference electron density maps (Fig. 1 Lower) show a substantially greater degree of structural isomorphism with the uncomplexed protein and essentially the same mode of binding single-stranded DNA as seen in the co-crystallized complex. Low-resolution studies (7 Å) of crystals soaked in longer single-stranded deoxyribonucleotides did not show additional difference electron density suggesting that single-stranded DNA beyond four residues from the 3' terminus may be partially or totally disordered in these high ionic strength soaking conditions.

**DNA-Protein Interactions.** In all three complexes, the single-stranded DNA is bound to the 3'-5' exonuclease active site through several interactions with the protein (Fig. 2). Three protein side chains are making hydrophobic interactions with the 3' terminal three bases. Leu-361 is wedged between (and interacting with) both the 3' terminal and penultimate bases. Phe-474 interacts with the other side of the 3' terminal base. The third base from the 3' terminus is interacting with His-660, which moves from its position in the native Klenow structure. The 3' terminal residue binds nearly identically to dTMP although its 5' phosphate may be in a slightly different orientation. The 3'-OH of the sugar moiety is buried and hydrogen-bonded to the backbone amide of Thr-358, as previously observed in the binding of dTMP (3).

The sugar-phosphate backbone of the single-stranded DNA interacts with three amino acid side chains and presumably also with the two divalent metal ions (Fig. 2). Although refinement will be required to confirm the existence of hydrogen bonds, it appears that the third phosphate from the 3' terminus is hydrogen-bonded to the guanidinium group of Arg-455; the penultimate phosphate is hydrogen-bonded to Gln-419 and a main chain amide of Met-458. In the complex with D424A, the phosphate of the bond to be cleaved is hydrogen-bonded to Tyr-497 and bound to a zinc ion at site A. If we assume that metal ion B binds as observed in the complex with dNMP, then it would also interact with the same phosphate in the DNA complex.

The conformation of the single-stranded DNA is extended. The binding site will only accommodate single-stranded DNA; it appears impossible to build into the model a complementary single strand without creating numerous close contacts. This is consistent with the results of Brutlag and Kornberg (9) who concluded from the temperature dependence of 3'-5' exonuclease reaction that the 3' terminus must be single stranded in the exonuclease active site. However, our observation that four 3' terminal nucleotides of a duplex become single stranded indicates a larger degree of melting of the duplex than previously anticipated. The bound single-stranded DNA in all three complexes has a very similar, but probably not identical, structure and orientation on the protein.

**DISCUSSION**

**Mechanism of 3'-5' Exonuclease.** The structural studies reported here as well as previous studies place constraints on the possible catalytic mechanisms of the 3'-5' exonuclease reaction. Assuming that the complex observed with the deoxytetranucleotide substrate approximates a Michaelis complex, it appears that only the two divalent metal ions assumed to be bound at sites A and B and the γ-OH of Tyr-497 are in close enough contact with the phosphate of the bond to be cleaved to directly affect hydrolysis. Other protein side chains appear too far away to be directly involved with catalysis. These metal ion sites in the crystalline complex with dTMP can be substituted by Mn$^{2+}$, Zn$^{2+}$, or Mg$^{2+}$ although the metal ion in vivo is unknown. The fact that the divalent metal ion at site B is essential for 3'-5' exonuclease activity was demonstrated by the failure of the D424A mutant protein lacking this metal site B to catalyze this hydrolysis (5). Gupta and Benkovic (10) showed that the hydrolysis results in inversion of configuration at phosphorus. Assuming that the reaction is an associative in-line mechanism and proceeds through a transient pentacovalent species, then the water molecule must attack from an apical position opposite the leaving 3'-OH.

Although we do not see a water molecule bound in the appropriate attacking position in this inactive complex, the observed position of the leaving group plus the requirement that the reaction proceed with inversion of configuration (10) necessitates that this water molecule attack from a position near the γ-OH of Tyr-497 and metal ion A (Figs. 2 and 3). There are no other protein side chains in the vicinity that could function as a general base.

It appears plausible that metal ion B may function to stabilize the transient pentacovalent species and/or to facilitate the leaving of the 3' oxyanion from an axial position; the metal ion A may facilitate the formation of an attacking hydroxide ion. Herschlag and Jencks (11) have concluded from a series of nucleophilic displacement reactions at phosphorus that divalent metal ions can serve to stabilize an oxyanion leaving group. Thus for phosphomonooesters, divalent metal ions render phospho-group transfer less sensitive to the nucleophility of the attacking group while increasing the overall reaction rate. We suggest that the metal ion in site
B could be playing an additional role in stabilizing the transient pentacovalent species. The hydrolysis of methyl-ethylene phosphate proceeds some 10^6 times faster than the trimethyl analog (12), presumably because the ethylene bridge strains the O—P—O bond angle to 99°, 10° less than the tetrahedral angle (13). Perhaps metal ion B performs a function analogous to that played by the ethylene group in methylethylene phosphate by reducing the energy required to form the 90° O—P—O bond angle between apical and equatorial oxygens in the transition state (Fig. 3). A similar hypothesis has been proposed to explain the enhancement of hydrolysis of methyl phosphate when bound to the cobalt(III) ion (14). The metal ion A might be stabilizing a hydroxide ion to facilitate the attack on phosphate by water. Both the structure of the catalytic site and the proposed mechanism are reminiscent of the active site and mechanism for one step of the two-step reaction proposed for alkaline phosphatase (15), which has two bound zinc atoms 4 Å apart that interact with phosphate.

Whether Tyr-497 serves primarily to orient the substrate and/or a water molecule or plays a more vital role in catalysis can be tested by site-directed mutagenesis.

The phospho-group transfer reaction catalyzed by the 3′-5′ exonuclease of DNA polymerase I is similar to (or in some cases the same as) the phospho-group transfer reactions catalyzed by those enzymes in which the catalytic function is provided by an RNA molecule alone (21, 22). This raises the possibility that the catalytic mechanisms of these RNA enzymes may also be related to the catalytic mechanism of the DNA polymerase I 3′-5′ exonuclease. What makes this type of mechanism a particularly attractive possibility for RNA enzymes is that it may require only the involvement of two appropriately positioned metal ions and perhaps a tyrosyl hydroxyl and does not appear to necessitate the involvement of any side chains with the specialized chemical properties found in protein-based enzymes. The RNA enzymes would be required to (i) bind two magnesium ions 4.3 Å apart and (ii) bind and orient the RNA substrate(s) in appropriate juxtaposition to these two metal ions.

**Relationship Between the Exonuclease and Polymerase Active Sites.** How then do the polymerase and 3′-5′ exonuclease active sites work together to assure that mismatched base pairs incorporated at the polymerase active site are edited out at the 3′-5′ exonuclease active site some 25 to 30 Å away? An earlier hypothesis proposed that the exonuclease activity was a manifestation of a different conformational state of the polymerase active site and that the two active sites were in the same position (1). It has now been established, however, that the polymerase and exonuclease activities reside on different subunits in the case of DNA polymerase III (16, 17) and on different domains of the same polypeptide chain in the case of DNA polymerase I (5, 6). We have demonstrated here that the exonuclease active site is extensive and contains a binding site for four single-stranded nucleotides spanning a distance of some 15 Å from the 3′ terminus.

Comparison of the crystal structure of the Klenow fragment complexed with single-stranded DNA at the exonuclease active site and the previously proposed model (3) with duplex DNA model built at the polymerase active site suggests that these two active sites might be in “communication” by virtue of the DNA sliding between them (Fig. 4).

The 5′ end of experimentally observed single-stranded DNA comes close to the 3′ end of model-built duplex DNA forming a continuous pathway between the observed exonuclease and putative polymerase active sites. The path that the 3′ terminus must follow to proceed from the presumed polymerase active site to the exonuclease active site involves ≈4 bp of duplex DNA plus four bases of single-stranded frayed end.

**Structural Basis of Editing.** We suppose that after each polymerization step the 3′ terminus of the DNA may dissociate from the polymerase active site but not necessarily from the protein; rather, the protein may diffuse along the DNA. DNA polymerases I, like other polymerases, is processive (18) in that it incorporates about 20–25 nucleotides between each dissociation from the DNA. Thus, access of both active sites to the 3′ terminus after each nucleotide incorporation may require the 3′ terminus to slide between the sites some
portion of the time. Since the 3′→5′ exonuclease activity excises ~10% of all correctly incorporated nucleotides (19), it appears that the polymerase and exonuclease activities are in a delicately poised competition. In order for the two active sites to compete for the newly formed 3′ terminus, the rate of sliding between these two active sites would have to be faster than either of the two enzymatic rates. This would require that the DNA must slide the distance of 25 Å in less than 100 msec, a rate that is not unreasonably large (18).

For the 3′→5′ exonuclease to improve the fidelity of DNA synthesis, the enzyme must exert a significant bias in favor of excising mismatched base pairs over correctly matched base pairs (20). This requires the enzyme to exhibit an increased rate of exonuclease activity and/or a decreased rate of polymerase activity on 3′ terminal mismatched base pairs. How is this achieved? Since the 8-bp duplex DNA used in the recocrystallization experiment is found to bind with at least four nucleotides becoming single stranded, the melting of 4 bp at the 3′ terminus appears necessary for the 3′ terminus to bind to the exonuclease active site. If formation of single-stranded DNA is rate limiting, an increased exonuclease activity on DNAs containing mismatched base pairs could result from an increased rate of melting reflected in its lower thermal stability. Since a mismatched base pair in any of the four 3′ terminal nucleotides reduces free energy cost of forming a four nucleotide single-stranded DNA, binding of a mismatched 3′ terminus to the exonuclease active site would be enhanced relative to its binding to the polymerase active site. In fact, this model predicts that any factor favoring the formation of single-stranded DNA at the 3′ terminus and its binding to the exonuclease active site would enhance the efficiency of editing.

A reduced polymerase activity on DNAs containing mismatched base pairs would require their detection by the polymerase active site. Since many mismatched base pairs can assume the sugar-phosphate backbone structure of normal B-DNA, some other aspect of the mismatched base pair may be recognized by the enzyme. It has been suggested (23) that the enzyme may have a "reading head" in the minor groove that can detect mismatched base pairs and suppress either translocation or formation of a catalytically competent active site.

To provide a structural basis for understanding the role of the polymerase active site in the fidelity of this enzyme will require determination of the structure of a ternary complex of DNA and dNTP with the polymerase active site. Large crystals of D424A Klenow fragment have been grown at low ionic strength and neutral pH in the presence of DNA, α, β-methylene dTTP, and Mg²⁺ that may prove appropriate for these issues.

We thank Catherine Joyce, Nigel Grindley, Victoria Derbyshire, and Andrea Polesky for their continued help and suggestions; Harold Wyckoff and Olke Uhlenbeck for discussions; and Sherin Abdel-Meguid and colleagues at Monsanto for use of their Mark II area detector. This work was supported by American Cancer Society Grant NP-421 and National Institutes of Health Grant GM-22778 to T.A.S. and a Jane Coffin Childs Memorial Fund Fellowship to J.M.F.