Contacts between the 5' Nuclease of DNA Polymerase I and Its DNA Substrate*

Yang Xu, Olga Potapova, Andres E. Leschziner, Nigel D. F. Grindley, and Catherine M. Joyce†

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

The 5' nuclease of DNA polymerase I (Pol I) of Escherichia coli is a member of an important class of prokaryotic and eukaryotic nuclease, involved in DNA replication and repair, with specificity for the junction between single-stranded and duplex DNA. We have investigated the interaction of the 5' nuclease domain with DNA substrates from the standpoint of both the protein and the DNA. Phosphatase ethylation interference showed that the nuclease binds to the nucleotides immediately surrounding the cleavage site and also contacts the complementary strand one-half turn away, indicating that contacts are made to one face only of the duplex portion of the DNA substrate. Phosphodiester contacts were investigated further using DNA substrates carrying unique methylphosphonate substitutions, together with mutations in the 5' nuclease. These experiments suggested that two highly conserved basic residues, Lys78 and Arg81, are close to the phosphodiester 3' to the cleavage site. Our results provide strong support for a DNA binding model proposed for the related exonuclease from bacteriophage T5, in which the conserved basic residues mentioned above define the two ends of a helical arch that forms part of the single-stranded DNA-binding region. The nine highly conserved carboxylates in the active site region appear to play a relatively minor role in substrate binding, although they are crucial for catalysis. In addition to binding the DNA backbone around the cleavage point, the 5' nuclease also has a binding site for one or two frayed bases at the 3' end of an upstream primer strand. In agreement with work in related systems, 5' nuclease cleavage is blocked by duplex DNA in the 5' tail, but the enzyme is quite tolerant of abasic DNA or polarity reversal within the 5' tail.

DNA polymerase I (Pol I) of Escherichia coli has an intrinsic 5' nuclease activity that is important for the removal of RNA primers from Okazaki fragments during lagging strand replication and for the removal of damaged nucleotides in DNA excision repair (1). Homologous 5' nuclease domains are found in most other bacterial Pol I enzymes and in some bacteriophages, where the polymerase and 5' nuclease exist as separate polypeptides (2, 3). The bacterial 5' nuclease family also shows substantial sequence similarity to the FEN-1 eukaryotic nucleases, which are involved in various DNA transactions (4–8). Although the 5' nuclease were originally called 5'-3' exonucleases, these enzymes are more accurately described as structure-specific nuclease, with specificity for the junction between a base paired region and a single-stranded 5' overhang or "flap" (9–11). Cleavage by these nuclease usually takes place between the first two paired bases at the junction between the duplex and the single-stranded 5' tail (see Fig. 1), although some variability has been observed (10–13). Recent work suggests that the preferred substrate of the bacterial nuclease is a single unpaired base at the 3' end of the primer upstream of the 5' nuclease cleavage site (thus facilitating formation of a ligatable nick), and it seems likely that any variability in the observed cleavage position may reflect rearrangement via branch migration of the 5' and 3' single-stranded flaps (14, 15).

Alignment of the sequences of the bacterial and bacteriophage 5' nuclease revealed nine invariant carboxylates (2, 3), the majority of which are also conserved in the FEN-1 family (6, 16). The large number of conserved carboxylates raises the possibility that the 5' nuclease reaction, like the polymerase and 3'-5' exonuclease reactions, may be catalyzed by divalent metal ions coordinated to the enzyme via carboxylate ligands (17). Moreover, mutagenesis studies in several systems have demonstrated the importance of the carboxylates in the 5' nuclease reaction (3, 18–21).

Crystal structures have been determined for three of the prokaryotic 5' nuclease, the intrinsic 5' nuclease of Thermus aquaticus (Taq) DNA polymerase I (22), bacteriophage T4 RNase H (6), and bacteriophage T5 exonuclease (23), and for two archaeabacterial FEN-1 analogs (24, 25). The core structures of all five enzymes are very similar with the invariant carboxylates clustered in the central portion of the molecule, some coordinated to divalent metal ions, although the precise locations of the metal ions and the ligand geometries are somewhat variable (3, 26).

In the absence of any crystalline complexes with bound DNA, little information exists on how the 5' nuclease interact with their substrate. A plausible candidate for a DNA-binding region in the prokaryotic 5' nuclease is a cluster of highly conserved and mostly basic residues. The relevant region of the protein structure is not well resolved in the Taq DNA polymerase and T4 RNase H structures, but in T5 exonuclease it is located on one side of an unusual helical arch structure (23). The two FEN-1 structures have a flexible loop in place of the...
helical arch (24, 25). In all three structures the loop or arch is large enough to accommodate single-stranded DNA, suggesting that the single-stranded 5′ end of the DNA substrate may be threaded through this part of the protein (Refs. 23, 24, and 26, see Fig. 2B), with the duplex DNA portion held close to the base of the loop or arch by a helix-loop-helix motif present in all five of the available structures (25). The threading model was originally proposed based on observations that 5′ nuclease cleavage requires a free 5′ end and is blocked by significant obstructions (e.g. annealed primers) in the single-stranded tail (3, 10, 27). However, recent work has shown that FEN-1 can tolerate a variety of modifications, including an 11-nucleotide branch, within the 5′ flap DNA, prompting re-examination of the threading model (28). In this work, we have investigated the interaction between the 5′ nuclease and its DNA substrate from the perspective of both the protein and the DNA, and have attempted to identify contacts between particular protein side chains and individual phosphodiester groups on the DNA.

EXPERIMENTAL PROCEDURES

Materials—Oligonucleotides for site-directed mutagenesis and for 5′ nuclease substrates were synthesized by the Rockefeller University Resource Laboratory at Yale Medical School. Those used as reaction substrates were purified by gel electrophoresis. Oligonucleotide concentrations were determined spectrophotometrically using calculated extinction coefficients (29). DNA oligonucleotides containing abasic spacers or regions with reversed polarity were kindly provided by Dr. Daniel Kaplan, and have been described elsewhere (30). Radiolabeled nucleotides were synthesized with Nα-ethyl-N′-nitrosourea was purchased from Sigma. DNase I was from Cooper Biomedical. T4 polynucleotide kinase, DNA ligase, and restriction enzymes were from New England Biolabs or Roche Molecular Biochemicals and were used according to the accompanying instructions.

Mutagenesis and Protein Purification—Following our published methods (3), mutations in the 5′ nuclease were constructed and subcloned into an expression plasmid for the 323-amino acid 5′ nuclease domain, and the mutant proteins were purified by fast protein liquid chromatography (Amersham Pharmacia Biotech). To remove low levels of contamination by intact Pol I, the purification method was modified in either of two ways. When preparing 5′ nuclease mutants on a small scale, the peak fractions from the MonoQ HR16/20 column were assayed for polymerase activity (31), and only those fractions that had no measurable polymerase activity were combined and applied to the phenyl-Superose column. For large scale purification, a final gel filtration step was included. The 5′ nuclease pool from the phenyl-Superose column was concentrated by ammonium sulfate precipitation, applied to a HiLoad 16/20 Superdex 200 prep grade column (120-ml bed volume) and eluted with 50 mM Tri-HCl, pH 7.5, 1 mM dithiothreitol, 100 mM NaCl. This purification method was necessary for the present study because low-level contamination by full-length Pol I interfered with DNase I footprinting experiments, particularly with weak binding 5′ nuclease mutants. Because the DNA binding affinity of the polymerase domain is much greater than that of the 5′ nuclease, competition between the two was a problem even though the amount of Pol I was not enough to affect enzyme activity assays and was not detectable on a Coomassie Blue-stained gel. We believe that the contaminating Pol I in our preparations was derived from the chromosomal polA locus rather than from readthrough of the amber codon in our expression plasmid for the 5′ nuclease domain (3) because the problem was alleviated by subcloning the 5′ nuclease mutations into an analogous expression construct that lacked the complete polymerase coding region.

Kinetic Measurements—Single-turnover measurements of 5′ nuclease cleavage were carried out using the substrates shown in Fig. 1a. The DNA oligonucleotides were labeled with 32P at either the 5′ or the 3′ end (see legend to Fig. 1). Reactions contained ~5 mM DNA substrate and the 5′ nuclease at a series of 7–8 concentrations (typically from 0.1 to 40 μM), chosen so as to bracket the $K_{m}$ value. All reactions were carried out at ambient temperature (23 °C) in a buffer containing 50 mM Tris-HCl, pH 7.6, 5 mM MgCl₂, 100 mM NaCl, and 5 μM ATP. The reactions catalyzed by the mutant 5′ nuclease derivatives (requiring sampling at time intervals of ≥10 s) were conducted as described previously (3). The reaction catalyzed by the wild-type 5′ nuclease was too fast for manual sampling, and was instead carried out on a rapid quench-flow instrument (KinTek Corp., Model RQF-3) by mixing a solution containing the DNA substrate with an equal volume of enzyme...
fractionation, appropriate fragments were quantitated by phosphorimaging. For each enzyme concentration, the percent of uncomplexed DNA was calculated by determining the ratio of the radioactivity in a DNA band within the protected region to that in a band outside the protected region, and comparing with the same ratio from a control lane without enzyme present. The calculated enzyme concentration at which 50% of the DNA is bound, which was determined by curve fitting. For this analysis to be valid, the DNA concentration in the reaction must be substantially below $K_{D}$, so that the total enzyme concentration approximates the concentration of free enzyme. This condition was fulfilled because of the relatively high $K_{D}$ values for the 5’ nuclease and its mutant derivatives.

**Phosphory Ethylation Interference**—The procedure was modified from published procedures (34, 35). The 5’ end-labeled 132-mer double-hairpin substrate (Fig. 1b) was used to analyze phosphate contacts close to the cleavage site. To ~2 pmol of the DNA substrate in 0.1 M NaCl, 50 mM MgCl$_2$, and 7 M urea, 0.1 mM ethylnitrosourea, pH 8.0, was added 0.1 mM of ethanol saturated with ethylnitrosourea at 50 °C. After incubation at 50 °C for about 25 min, the ethylnitrosourea was removed by five extractions with 1-ml portions of water-saturated ether. The ethylated DNA solution (~0.1 ml) was adjusted to 50 mM Tris, pH 7.5, 5 mM MgCl$_2$, and incubated at room temperature with wild-type 5’ nuclease (~200 pmol). One-third of the reaction mixture was quenched with an equal volume of formamide-dyes containing 30 mM EDTA, at time intervals chosen so as to obtain three samples having ~25, 50, and 75% product formation, respectively. The 36-mer product was separated from unreacted 132-mer substrate in each sample by electrophoresis on an 8% polyacrylamide gel containing 40% (v/v) formamide and 7 M urea. The DNA bands were located by staining with ethidium bromide, and quantitated as described in the legend to Fig. 3. To analyze phosphate contacts on the DNA strand opposite the cleavage site, the same procedure was carried out using the 112-mer, and analyzing samples from the pool of unreacted substrate on an 8% polyacrylamide gel containing 40% (v/v) formamide and 7 M urea.

**RESULTS**

**Mutant 5’ Nuclease Derivatives**—To determine which amino acid side chains of the 5’ nuclease are involved in binding the DNA substrate, we studied mutations in two groups of highly conserved residues (Table I). One was the nine carboxylates, shown to be crucial for 5’ nuclease cleavage in our previous study (3). We also made alanine substitutions at Arg$_{97}$, Arg$_{77}$, Tyr$_{77}$, Lys$_{78}$, and Arg$_{81}$ (Fig. 2A); in the structure of the T5 5’ nuclease, the equivalent residues define the base of the helical arch that is proposed to form part of the DNA-binding site (23) (Fig. 2B). The effects of all these mutations were studied in context of the separate 5’ nuclease domain (residues 1–323 of Pol I) so as to avoid ambiguities due to the DNA binding properties of the other domains of Pol I.

Cleavage of our standard (22 + 68)-mer substrate (Fig. 1) by the wild-type and mutant 5’ nuclease derivatives was measured under single-turnover conditions (excess enzyme) as a function of enzyme concentration, giving the enzyme-DNA dissociation constant ($K_D$) and the maximum cleavage rate ($k_c$), which reflects the rate of steps up to and including phosphoryl

**Measurement of $K_D$ by DNase I Protection**—Footprinting of the 112-mer double-hairpin oligonucleotide (Fig. 1b), as a function of 5’ nuclease concentration, was carried out as described previously (15, 32). After gel
transfer (Table II). All of the mutations tested caused substantial decreases in cleavage rate, ranging from ~20-fold (R20A and R81A) to ~5 × 10^3-fold (D115A). R20A and K78A were the only mutations that caused a significant decrease in DNA binding affinity; some of the carboxylate mutations, particularly D63A, D138A, and D185A, caused an increase in DNA binding affinity. For a few 5'-nuclease mutants, the DNA binding affinity was measured by a DNase I footprint titration; this showed good agreement with the kinetic determinations and had the added advantage that the measured binding constant is related to a binding site which can be visualized on a gel (Ref. 15 gives details of the 5'-nuclease footprint). Footprinting was clearly the preferred method for the D115A mutant protein, which had no measurable activity in the 5'-nuclease assay; conversely, footprinting was not appropriate for mutants with relatively high nuclease activity which degraded the DNA substrate during the experiment.

Circular dichroism (CD) measurements provided two lines of evidence that the 5'-nuclease mutations being studied did not cause any major changes in protein three-dimensional structure. First, the CD spectra of the mutant proteins were similar to that of wild-type, implying that the same secondary structure components are present in all. Second, the denaturation temperatures of the mutant proteins were all within a few degrees of the T_m for the wild-type 5'-nuclease (Table II). Mutations in surface residues on the proposed helical arch region gave T_m values very close to the wild-type value, as observed previously for mutations in surface residues on the polymerase domain of Klenow fragment (33). The T_m values for the carboxylate mutants were more variable; perhaps reflecting subtle changes in protein conformation and stability on altering negatively charged side chains (and metal ligands) in the central portion of the 5'-nuclease domain.

**Phosphory Ethylation Interference**—We used phosphate ethylation interference to define contacts on the DNA substrate that are important in binding to the 5'-nuclease. Our basic experimental design involved treatment of a labeled double-hairpin DNA substrate (Fig. 1b) with ethylnitrosourea under conditions that gave no more than one ethylation of backbone phosphodiester groups per DNA molecule. The pool of ethylated molecules was then used as the substrate for cleavage by the 5'-nuclease, and the labeled product was separated from uncleaved substrate by gel electrophoresis. Those molecules in which ethylation interfered with the 5'-nuclease reaction should be under-represented in the product molecules and enriched in the population of uncleaved substrate molecules (Fig. 3A). Chemical cleavage and gel electrophoresis then revealed the positions of ethylation in each population. Analysis of the non-cleaved substrate pool showed that phosphate ethylation was most inhibitory at the cleavage site and the immediately flanking positions (+1 and −1); weaker inhibition was seen at neighboring positions on the “top” strand (+2, −2, and −3) (Fig. 3C and E). The product pool was less informative because the full-length product band obscured the most relevant posi-

---

**Table I**

| Residue in E. coli Pol I | Conservation* |
|--------------------------|---------------|
| Asp13        | 47 Asp | 11 Asp | 4 Asp |
| Arg20        | 47 Arg | 10 Arg, 1 Ala | 1 Arg, 2 Gin, 1 Ser |
| Asp23        | 46 Asp, 1 Glu | 11 Asp | 4 Asp |
| Arg70        | 47 Arg | 11 Arg | 4 Arg |
| Tyr77        | 47 Tyr | 11 Tyr | 4 Tyr |
| Lys78        | 47 Lys | 11 Lys | 4 Lys |
| Arg81        | 47 Arg | 11 Arg | 4 Arg |
| Glu113       | 47 Glu | 11 Glu | 4 Glu |
| Asp115       | 47 Asp | 11 Asp | 4 Asp |
| Asp116       | 47 Asp | 11 Asp | 4 Asp |
| Asp138       | 47 Asp | 11 Asp | 4 Asp |
| Asp140       | 47 Asp | 9 Asp, 2 Gly | 4 Asp |
| Asp185       | 47 Asp | 9 Asp, 1 Lys, 1 Val |
| Asp188       | 47 Asp | 9 Asp, 2 Ser | 4 Asp |

*The alignment from which this analysis was derived is available on the web at pantheon.yale.edu/~ojoce/align.html.

Coding regions identified as members of the 5'-nuclease family on the basis of sequence similarity. In a few cases, e.g. Ref. 37, nuclease activity has been demonstrated.

---

**Table II**

| Enzyme | Kinetics* | Footprinting* | T_m (°C) |
|--------|-----------|---------------|----------|
|        | k_c [min⁻¹] | K_D |       |        | No Mg²⁺ | 5 mM Mg²⁺ |
| WT     | 47 ± 1     | 4.7 ± 0.7 (2) | 41 ± 1 (2) | 49       |
| R20A   | 2.8 ± 0.8  | 27 ± 1.8 (3) | ND       | 39       |
| R70A   | (4.3 ± 0.4) × 10⁻² | 4.0 ± 0.9 (2) | 1.8 ± 0.1 (2) | 42       |
| Y77A   | (4.0 ± 0.5) × 10⁻³ | 4.8 ± 0.6 (2) | 26 ± 3 (2) | 44       |
| K78A   | (4.3 ± 1.1) × 10⁻² | 11 ± 2 (2) | 5.5 ± 1.5 (4) | 40       |
| R81A   | 2.2 ± 0.1  | 5.5 ± 1.5 (4) | 7.6 ± 1.4 (3) | 44       |
| D15N   | (7.5 ± 0.9) × 10⁻⁴ | 0.86 ± 0.18 (3) | 2.7 ± 2.0 (5) | 38       |
| D65A   | (2.0 ± 0.4) × 10⁻² | 1.2 ± 0.1 (2) | 46 ± 1 (2) | 54       |
| E113A  | (1.4 ± 0.4) × 10⁻² | 1.2 ± 0.1 (2) | 46 ± 1 (2) | 54       |
| D115A  | 1.3 ± 10⁻⁶w | 1.2 ± 0.1 (2) | 46 ± 1 (2) | 54       |
| D116A  | (6.4 ± 1.6) × 10⁻³ | 0.3 ± 1.2 (3) | 43       |
| D138N  | (9.4 ± 1.9) × 10⁻⁴ | 0.30 ± 0.17 (3) | 39       |
| D140A  | (5.0 ± 1.0) × 10⁻⁴ | 4.6 ± 2.1 (5) | 50 ± 1 (2) | 55       |
| D185A  | (9.7 ± 1.6) × 10⁻³ | 0.18 ± 0.13 (3) | 0.1 ± 0.02 (2) | 37       |
| D188A  | (5.3 ± 1.0) × 10⁻³ | 1.4 ± 1.0 (4) | 41       |

* Kinetic measurements were made using the (22 + 68)-mer substrate (Fig. 1a).

DNase I protection was carried out using the 112-mer substrate (Fig. 1b). Small differences in binding affinity between the two substrates used in K_D measurements are to be expected because of their different sequences, and because the (22 + 68)-mer has a nick and the 112-mer has a single base gap upstream of the cleavage site (see also Ref. 15).

For the three proteins tested, the CD spectra in the presence and absence of Mg²⁺ were superimposable.

Use of the rapid-quench-flow instrument gave a more accurate measurement of the reaction rate for the wild-type 5'-3' exonuclease, and indicated that this rate is about 10-fold faster than that reported in our previous study.

The number of determinations is shown in parentheses.

ND, not determined; however, the high level of 5'-nuclease activity argues against any significant alteration to the structure.

The reaction was too slow to measure.
FIG. 3. Identification of positions on the DNA substrate where ethylation inhibits cleavage by the 5′ nuclease. The two double-hairpin substrates shown in Fig. 1b were used, and gave similar results. The 132-mer was preferable for studying the “top” strand around the cleavage site because its 35-nucleotide 5′ extension ensured that the relevant phosphate positions were sufficiently far from the labeled 5′ end that the two hydrolysis products (phosphorylated and non-phosphorylated) derived from each ethylation position (34) would have very similar gel mobilities. The 112-mer substrate, which has a shorter 5′-tail, gave better resolution for investigating the “bottom” strand opposite the cleavage site. A, analysis of phosphate positions on the top strand (132-mer substrate). Lane N shows the product of alkaline hydrolysis of the unfractonated ethylated DNA, and thus displays all possible ethylated phosphate positions. (A control hydrolys of non-ethylated DNA typically gave little or no signal.) The remaining lanes show the substrate and product pools after reaction with the 5′ nuclease for the indicated time intervals. The positions of ethylated phosphates are indicated relative to the 5′ nuclease cleavage site (designated “0”); negative numbers correspond to positions 5′ to the cleavage site and positive numbers correspond to positions 3′ to the cleavage site. B, analysis of phosphate positions on the bottom strand (112-mer substrate). The unfractonated ethylated DNA is compared with the unreacted substrate pool after 40 and 60 min reaction. Lane M contains markers generated by restriction enzyme digestion of the 112-mer (sizes are indicated) and G was generated by chemical cleavage at guanine residues (36). “*” indicates the 54-mer corresponding to the phosphate directly opposite the 5′ nuclease cleavage site. C, quantitation of the data in panel A. For each lane, the radioactivity present in each band was quantitated by PhosphorImager analysis, and was normalized to a band outside the ethylation “footprint.” The normalized value for each band was then divided by the corresponding value for the unfractonated ethylated DNA (N lane), so that a non-interfering position should give a value of 1, while interference will result in a value < 1 in the product pool or > 1 in the substrate pool. Because the product pool is more informative at low extents of reaction while the substrate pool is more informative as the reaction nears completion, the plot shows data from the product after 6 min reaction (C) and the unreacted substrate after 60 min reaction (D). The R20A and R70A mutant proteins were similar to the wild-type 5′ nuclease, with cleavage of the methylphosphonate substrate about 10–20-fold slower than the normal substrate. By contrast, the Y77A, K78A, and R81A proteins cleaved the methylphosphonate substrate more rapidly than they cleaved the unmodified substrate; the effect was largest for the K78A mutant.

All substrates with methylphosphonates 5′ to the cleavage site (including MP<sub>−2</sub>) gave essentially the same pattern of cleavage

Methylphosphonate Modifications—To investigate contacts between individual phosphodiester positions on the substrate and side chains on the 5′ nuclease, we used methylphosphonate-substituted DNA oligonucleotides in conjunction with mutations in the proposed helical arch region (R20A, R70A, Y77A, K78A, and R81A). Unlike phosphate ethylation, which adds a bulky alkyl group to the DNA phosphodiester backbone, the methylphosphonate modification removes a negative charge on the phosphodiester backbone without substantially increasing the size of the substituent. We tested the phosphate positions flanking the exonuclease cleavage site that corresponded to positions of inhibitory ethylations (−3, −2, −1, and +1 positions), using a series of oligonucleotides, each having a methyl group on a single backbone phosphate position (Fig. 1c). A methylphosphonate at the −7 position served as a negative control. Because the methylphosphonate modification was introduced by chemical synthesis, each oligonucleotide consisted of approximately equal amounts of the R<sub>p</sub> and S<sub>p</sub> methylphosphonate diastereomers.

Cleavage of the methylphosphonate substrates by the wild-type and mutant 5′ nucleases was measured initially at two different enzyme concentrations: an enzyme concentration at least 3-fold greater than K<sub>D</sub> (rate constant reflects k<sub>c</sub>) and an enzyme concentration less than 0.5 × K<sub>D</sub> (rate responds to k<sub>c</sub> / K<sub>D</sub>). At both concentrations the MP<sub>−2</sub> and MP<sub>−3</sub> analogs gave reaction rates indistinguishable from those of the MP<sub>−1</sub> and unmodified (22 + 68)-mer controls, and were therefore not studied further. The MP<sub>−1</sub> substrate was cleaved by the wild-type 5′ nuclease about 5-fold more slowly than the unmodified substrate, based on initial rates (Fig. 4A). Cleavage was noticeably biphasic, presumably reflecting different rates of cleavage of the two methylphosphonate diastereomers; however, the rates of the two phases were sufficiently different to allow determination of the kinetic parameters for the fast phase (Table III). For some of the mutant proteins the two phases were more similar in rate, and this precluded a detailed kinetic analysis. Nevertheless, some interesting trends were apparent from a comparison of the cleavage of unmodified and MP<sub>−1</sub> substrates by mutant 5′ nucleases derivatives (Fig. 4A). The R20A and R70A mutant proteins were similar to the wild-type 5′ nuclease, with cleavage of the methylphosphonate substrate about 10–20-fold slower than the normal substrate. By contrast, the Y77A, K78A, and R81A proteins cleaved the methylphosphonate substrate more rapidly than they cleaved the unmodified substrate; the effect was largest for the K78A mutant.

All substrates with methylphosphonates 5′ to the cleavage site (including MP<sub>−1</sub>) gave essentially the same pattern of cleavage
By contrast, with the MP analog, the predominant product, amounting to a few percent of the total, is the unmodified substrate, with the 11-mer as the major product. The yield of 10-mer was about 1.5 times the amount of 11-mer. A likely explanation is that the methylphosphonate substitution inhibited cleavage at the normal position, allowing a minor cleavage site to predominate. This idea is consistent with the observation that cleavage of the MP analog was 10-fold slower than for the unmodified (22-mer-OH). From the slope at low enzyme concentration, the Michaelis constant was determined to be 9.7 × 10⁻⁹ M⁻¹ min⁻¹, which is 10⁴-fold lower than for the shorter substrates.

Table III: Kinetics of cleavage of modified substrates by the 5' nucleotide

| Enzyme | Substrate | Length of 5' tail | kₐ | Kᵦ |
|--------|-----------|-----------------|-----|-----|
| Wild-type | 12-mer-OH | 0 | 92 ± 1 | 9.2 ± 1.7 |
|         | 12-mer    | 0 | 121 ± 5 | 10.1 ± 0.2 |
|         | 13-mer    | 1 | 82 ± 4 | 7.7 ± 3.1 |
|         | 14-mer    | 2 | 80 ± 8 | 3.9 ± 0.8 |
|         | 15-mer    | 3 | 67 ± 3 | 2.8 ± 0.4 |
|         | 22-mer    | 10 | 47 ± 1 | 4.7 ± 0.7 |
|         | MP 22-mer | 10 | 13.3 ± 0.06 | 3.9 ± 0.6 |

* The indicated oligonucleotides were annealed to the 68-mer (see Fig. 1a). Except for the 12-mer-OH, all were 5' phosphorylated.

** The kinetic determinations were carried out at least in duplicate, except for the determination of kₐ and Kᵦ for R81A cleavage of the 12-mer and 12-mer-OH, which came from a single experiment. However, the -10 fold difference in the rate of cleavage by R81A of the phosphorylated and nonphosphorylated 12-mer was confirmed independently.

* ND, not determined because the plot of rate versus enzyme concentration did not plateau even at the highest concentrations attainable.

products as the unmodified substrate, with the 11-mer as the predominant product (Fig. 4B).² By contrast, a methylphosphonate 3' to the cleavage site (MP₁ substrate) changed the position of cleavage by the wild-type 5' nuclease. With the unmodified (22 + 68)-mer, the major product was the 11-mer; this is attributed to cleavage between the two paired bases immediately following the 5' single-stranded tail (Fig. 5). The 10-mer was a minor product, amounting to a few percent of the total. By contrast, with the MP₁ substrate, the yield of 10-mer was about 1.5 times the amount of 11-mer. A likely explanation is that the methylphosphonate substitution inhibited cleavage at the normal position, allowing a minor cleavage site to predominate. This idea is consistent with the observation that cleavage of the MP₁ substrate was 40–50-fold slower than for the unmodified (22 + 68)-mer. The R20A mutation reversed the effect of the MP₁ substitution on the product distribution, so that cleavage generated about four times more 11-mer than 10-mer (Fig. 5).

² There were some minor differences; the MP₁ substrate gave a small amount of product longer than 11 nucleotides, presumably reflecting a preference for placing the methylphosphonate further from the cleavage site. Also, as would be expected, neither the MP₁ nor MP₂ substrates gave detectable amounts of 10-mer, normally seen as a minor cleavage product with the normal substrate.
FIG. 6. Effect of an upstream 3' flap on 5' nuclease cleavage. A, gel fractionation of the products of cleavage by the wild-type 5' nuclease of a series of substrates (shown in B) having different lengths of 3' flap. The strand with the 5' overhang was labeled at the 5' end and then annealed to a 2-fold excess of the template and a 4-fold excess of the appropriate upstream primer. Cleavage was allowed to proceed for the indicated times in a reaction mixture containing 10 nM DNA (labeled strand) and 50 µM 5' nuclease (wild-type or mutant) in 50 mM Tris-HCl, pH 7.9, 5 mM MgCl₂, 0.1 M NaCl. In these experiments, about 60% of the substrate remained uncleaved, presumably because of incomplete annealing of the labeled strand. The identity of the 15-mer cleavage product was confirmed using a chemically synthesized marker. B, product analysis of the reactions shown in A, as well as similar reactions carried out with the D188A and R20A mutant derivatives. The amount of each labeled product band, as a percentage of the total product, is plotted on a bar graph. On the left are indicated the substrate structures expected in the absence of any rearrangement. Using the reasoning explained in the text, we deduced the substrate configurations expected to give rise to the major products in each reaction, and these are shown to the right of the graphs. ▼ indicates the position of 5' nuclease cleavage.
other mutants were tested with the MP₃₁ substrate, D188A resembled R20A in giving predominantly the 11-mer product, but most had too little activity on the unannealed region at the 3’ end of the primer upstream of the 5’ nuclease cleavage site (Fig. 6). The best substrates were those designed with one or two unannealed bases at the primer 3’ end; these were cleaved slightly faster than the substrates with 0 (nicked substrate) or three unannealed bases (Fig. 6A). Cleavage of substrates with even longer 3’ flaps (4 and 7 nucleotides) was at least 10-fold slower than cleavage of the 0- or 3-nucleotide 3’ flap substrates and similar to that of a substrate without a primer upstream of the 5’ nuclease cleavage site. The sizes of the 5’ nuclease cleavage products indicated that some of the substrates, when bound to the 5’ nuclease active site, may have adopted base pairing configurations that were different from the simple arrangements originally assumed (listed on the left of Fig. 6B). Assuming that the 5’ nuclease always cuts between the two paired bases immediately adjacent to the 5’ tail, the substrate structures shown to the right of the bar graphs in Fig. 6B could be deduced, as follows (note that not all of the base pairs shown are complementary in the Watson-Crick sense). In the absence of an upstream primer, the wild-type 5’ nuclease gave approximately equal parts of the expected 15-mer, and the 16-mer corresponding to melting of an additional base pair of the duplex. With the nicked substrate, the products consisted of approximately equal amounts of 15- and 14-mer, the latter attributed to rearrangement to give a one-nucleotide 3’ flap. The substrates with 1- and 2-nucleotide 3’ flaps behaved as would be expected if no rearrangement took place, yielding almost exclusively the 15-mer on cleavage with the wild-type nuclease. By contrast, the substrate with the 3-nucleotide flap gave mainly 16-mer, implying that 5’ nuclease cleavage took place preferentially on molecules that had rearranged to give a 2-nucleotide 3’ flap. The 4-nucleotide flap substrate showed some rearrangement to a 3-nucleotide flap so that cleavage gave a mixture of 15- and 16-mer, whereas the 7-nucleotide flap showed less tendency to rearrange, yielding predominantly the 15-mer.

The R20A and D188A mutations influenced the pattern of cleavage products in different ways (Fig. 6B). On all four substrates tested, D188A gave longer products than wild-type, regardless of the presence of an upstream primer. The behavior of R20A was more complicated; this 5’ nuclease mutant gave shorter products than wild-type in the absence of an upstream primer, but the cleavage pattern was only subtly different from that of the wild-type enzyme on substrates having an upstream primer.

**Modifications to the 5’ Flap**—Since previous studies had shown eukaryotic FEN-1 to be tolerant of a wide range of modifications within the single-stranded 5’ tail, we tested the 30-nucleotide tail, or slower formation of the catalytically competent complex with the DNA branch point at the nuclease active site. The presence or absence of a phosphate at the 5’ end of a fully annealed downstream DNA strand (compare 12-mer with 12-mer-OH) did not affect cleavage by the wild-type 5’ nuclease, or by the R70A, Y77A, and K78A mutants. By contrast, the R81A mutant cleaved the phosphorylated substrate ~10-fold more slowly than the 5’-hydroxyl substrate; the kᵣ for the non-phosphorylated 12-mer was similar to the value for the standard (22 + 68)-mer, whereas the value for the phosphorylated 12-mer was unexpectedly low.

**Effect of Unpaired Bases at the Primer Terminus (a 3’ Flap)**—The wild-type 5’ nuclease was assayed on a series of substrates designed so that they differed only in the length of the unannealed region at the 3’ end of the primer upstream of the 5’ nuclease cleavage site (Fig. 6). The best substrates were those designed with one or two unannealed bases at the primer 3’ end; these were cleaved slightly faster than the substrates with 0 (nicked substrate) or three unannealed bases (Fig. 6A). Cleavage of substrates with even longer 3’ flaps (4 and 7 nucleotides) was at least 10-fold slower than cleavage of the 0- or 3-nucleotide 3’ flap substrates and similar to that of a substrate without a primer upstream of the 5’ nuclease cleavage site. The sizes of the 5’ nuclease cleavage products indicated that some of the substrates, when bound to the 5’ nuclease active site, may have adopted base pairing configurations that were different from the simple arrangements originally assumed (listed on the left of Fig. 6B). Assuming that the 5’ nuclease always cuts between the two paired bases immediately adjacent to the 5’ tail, the substrate structures shown to the right of the bar graphs in Fig. 6B could be deduced, as follows (note that not all of the base pairs shown are complementary in the Watson-Crick sense). In the absence of an upstream primer, the wild-type 5’ nuclease gave approximately equal parts of the expected 15-mer, and the 16-mer corresponding to melting of an additional base pair of the duplex. With the nicked substrate, the products consisted of approximately equal amounts of 15- and 14-mer, the latter attributed to rearrangement to give a one-nucleotide 3’ flap. The substrates with 1- and 2-nucleotide 3’ flaps behaved as would be expected if no rearrangement took place, yielding almost exclusively the 15-mer on cleavage with the wild-type nuclease. By contrast, the substrate with the 3-nucleotide flap gave mainly 16-mer, implying that 5’ nuclease cleavage took place preferentially on molecules that had rearranged to give a 2-nucleotide 3’ flap. The 4-nucleotide flap substrate showed some rearrangement to a 3-nucleotide flap so that cleavage gave a mixture of 15- and 16-mer, whereas the 7-nucleotide flap showed less tendency to rearrange, yielding predominantly the 15-mer.

The R20A and D188A mutations influenced the pattern of cleavage products in different ways (Fig. 6B). On all four substrates tested, D188A gave longer products than wild-type, regardless of the presence of an upstream primer. The behavior of R20A was more complicated; this 5’ nuclease mutant gave shorter products than wild-type in the absence of an upstream primer, but the cleavage pattern was only subtly different from that of the wild-type enzyme on substrates having an upstream primer.

**Modifications to the 5’ Flap**—Since previous studies had shown eukaryotic FEN-1 to be tolerant of a wide range of modifications within the single-stranded 5’ tail, we tested the 30-nucleotide tail, or slower formation of the catalytically competent complex with the DNA branch point at the nuclease active site. The presence or absence of a phosphate at the 5’ end of a fully annealed downstream DNA strand (compare 12-mer with 12-mer-OH) did not affect cleavage by the wild-type 5’ nuclease, or by the R70A, Y77A, and K78A mutants. By contrast, the R81A mutant cleaved the phosphorylated substrate ~10-fold more slowly than the 5’-hydroxyl substrate; the kᵣ for the non-phosphorylated 12-mer was similar to the value for the standard (22 + 68)-mer, whereas the value for the phosphorylated 12-mer was unexpectedly low.

**Effect of Unpaired Bases at the Primer Terminus (a 3’ Flap)**—The wild-type 5’ nuclease was assayed on a series of substrates designed so that they differed only in the length of the unannealed region at the 3’ end of the primer upstream of the 5’ nuclease cleavage site (Fig. 6). The best substrates were those designed with one or two unannealed bases at the primer 3’ end; these were cleaved slightly faster than the substrates with 0 (nicked substrate) or three unannealed bases (Fig. 6A). Cleavage of substrates with even longer 3’ flaps (4 and 7 nucleotides) was at least 10-fold slower than cleavage of the 0- or 3-nucleotide 3’ flap substrates and similar to that of a substrate without a primer upstream of the 5’ nuclease cleavage site. The sizes of the 5’ nuclease cleavage products indicated that some of the substrates, when bound to the 5’ nuclease active site, may have adopted base pairing configurations that were different from the simple arrangements originally assumed (listed on the left of Fig. 6B). Assuming that the 5’ nuclease always cuts between the two paired bases immediately adjacent to the 5’ tail, the substrate structures shown to the right of the bar graphs in Fig. 6B could be deduced, as follows (note that not all of the base pairs shown are complementary in the Watson-Crick sense). In the absence of an upstream primer, the wild-type 5’ nuclease gave approximately equal parts of the expected 15-mer, and the 16-mer corresponding to melting of an additional base pair of the duplex. With the nicked substrate, the products consisted of approximately equal amounts of 15- and 14-mer, the latter attributed to rearrangement to give a one-nucleotide 3’ flap. The substrates with 1- and 2-nucleotide 3’ flaps behaved as would be expected if no rearrangement took place, yielding almost exclusively the 15-mer on cleavage with the wild-type nuclease. By contrast, the substrate with the 3-nucleotide flap gave mainly 16-mer, implying that 5’ nuclease cleavage took place preferentially on molecules that had rearranged to give a 2-nucleotide 3’ flap. The 4-nucleotide flap substrate showed some rearrangement to a 3-nucleotide flap so that cleavage gave a mixture of 15- and 16-mer, whereas the 7-nucleotide flap showed less tendency to rearrange, yielding predominantly the 15-mer.

The R20A and D188A mutations influenced the pattern of cleavage products in different ways (Fig. 6B). On all four substrates tested, D188A gave longer products than wild-type, regardless of the presence of an upstream primer. The behavior of R20A was more complicated; this 5’ nuclease mutant gave shorter products than wild-type in the absence of an upstream primer, but the cleavage pattern was only subtly different from that of the wild-type enzyme on substrates having an upstream primer.

**Modifications to the 5’ Flap**—Since previous studies had shown eukaryotic FEN-1 to be tolerant of a wide range of modifications within the single-stranded 5’ tail, we tested the 30-nucleotide tail, or slower formation of the catalytically competent complex with the DNA branch point at the nuclease active site. The presence or absence of a phosphate at the 5’ end of a fully annealed downstream DNA strand (compare 12-mer with 12-mer-OH) did not affect cleavage by the wild-type 5’ nuclease, or by the R70A, Y77A, and K78A mutants. By contrast, the R81A mutant cleaved the phosphorylated substrate ~10-fold more slowly than the 5’-hydroxyl substrate; the kᵣ for the non-phosphorylated 12-mer was similar to the value for the standard (22 + 68)-mer, whereas the value for the phosphorylated 12-mer was unexpectedly low.
effect on the E. coli Pol I 5′ nuclelease of some fairly substantial changes to the 5′ tail (Fig. 7). Cutting was not significantly affected if the entire single-stranded tail had the wrong polarity, with a 5'-to-5′ linkage immediately adjacent to the cleavage site. Abasic deoxyribose spacers at either end of the 5′ tail also had very little effect, except that, when the abasic sites were immediately 5′ to the dinucleotide that defines the cut site, the cleavage rate was more sensitive to decreasing enzyme concentration (Fig. 7B), implying that the modification interfered with binding to the 5′ nuclelease. If the abasic residues were three nucleotides further away from the cleavage site, their effect on the 5′ nuclelease reaction was negligible (data not shown). By contrast, cleavage was completely blocked by duplex DNA within the 5′ flap strand. The substrate in which the 5′ strand had a terminal hairpin (Fig. 7A) was refractory to 5′ nuclelease cutting unless the hairpin had been pre-cleaved with EcoRI (data not shown).

**DISCUSSION**

Contacts on the DNA—Both the phosphate ethylation interference and methylphosphonate substitution approaches demonstrated that the most significant contacts to the phosphodiester backbone of the DNA substrate are the positions immediately adjacent to the 5′ nuclease cleavage site (~1 and +1). In the ethylation experiment, the inhibitory positions on the opposite DNA strand indicated that the enzyme contacts one face only of the DNA helix downstream of the cleavage site (Fig. 3E), and the position of the proposed contacts is entirely consistent with the conceptual model of T5 exonuclease bound to DNA (23) (Fig. 2B). Because ethylation increases the steric bulk of the modified phosphate, we believe that the inhibitory positions on the opposite DNA strand merely indicate a region of close approach between enzyme and substrate. By contrast, ethylation at the positions flanking the cleavage site may have compromised a binding interaction by removing a negative charge. Single methylphosphonate substitutions at the +1 and −1 positions provided further information on phosphate contacts without complications due to steric effects. The biphasic kinetics of cleavage of the MP−1 substrate shows that substitution of one of the non-bridging oxygens is mildly deleterious (~4-fold decrease in kcat) whereas substitution of the other causes a much larger decrease in reaction rate.

The rather small effects of removing the 5′ tail nucleotides are consistent with the data for methylphosphonate substitution at the −2 and −3 positions, and make biological sense since the 5′ nucleosome must be able to process DNA molecules having a variety of 5′ flap lengths. The insensitivity to the presence of a 5′-terminal phosphate at a nick might seem surprising, given the effect of a methylphosphonate linkage immediately 5′ to the cleavage site. However, a terminal phosphate differs from a phosphodiester linkage in overall negative charge, and may be bound somewhat differently to the active site when not attached to a single-stranded DNA tail. The 5′ nucleosome of Taq DNA polymerase is similarly insensitive to the phosphorylation status of a fully annealed (i.e. nicked) substrate (38), although eukaryotic FEN-1 has a strong preference for a 5′ phosphate (27, 39).

Comparison of DNA molecules with various lengths of 3′ flaps indicates that the 5′ nucleosome does indeed have a binding site for unpaired nucleotides at the 3′ end of a primer upstream of the cleavage site. The interpretation of these experiments hinges on the assumption that the DNA substrate is bound at the 5′ nucleosome active site in a consistent manner such that cleavage always takes place between two bases presented in a paired configuration adjacent to the 5′ single-stranded tail. The observed product(s) can then be used to infer the base pairing arrangement(s) of each substrate. We and others have found this approach to provide a convincing interpretation of the products observed in previous studies (14, 15). Our current data indicate that the 3′ flap-binding site is optimized for one or two nucleotides. The oligonucleotides designed to have a 1- or 2-nucleotide 3′ flap were the preferred substrates, and the predominant product in either case did not require rearrangement from the most stably base paired structure. Other substrates tended to rearrange toward these favored structures; the nicked substrate gave a substantial amount of product implying rearrangement to a substrate having a single nucleotide 3′ flap, and the substrate with a 3-nucleotide 3′ flap gave primarily the product derived after rearrangement to a 2-nucleotide 3′ flap. The formation of products derived from longer 3′ flaps indicated that these substrates can be accommodated within the 5′ nucleosome active site; the slower reaction rate implied that this arrangement is not optimal, while rearrangement to give a shorter flap would require too many base pairs to be sacrificed.

Protein Contacts to the DNA Phosphates—The kinetic data (Table II) indicated that side chains at either end of the putative helical arch (Arg20, Arg29, Tyr77, Lys78, and Arg81) play an important role in the 5′ nuclease reaction, as has been found in other homologs (19, 40–43). Additionally, three of our experiments provided data linking individual protein side chains with particular positions on the bound DNA substrate. In each case the evidence came from a comparison of the reaction of wild-type and mutant 5′ nucleases(s) with two different substrates. In general terms, an observation that does not conform to expectations can indicate proximity, and possibly a functionally significant interaction, of the mutated side chain with the modified position on the substrate.

Our data with the MP−1 methylphosphonate substrate provided evidence for an interaction of Lys78 and the surrounding region of the protein with the phosphodiester immediately 5′ to the cleavage site. The Y77A, K78A, and R81A mutant proteins cleaved the MP−1 substrate more rapidly than the normal substrate, reversing the trend seen with the wild type 5′ nuclease (Fig. 4). Preference of a mutant protein for a non-cognate substrate may seem hard to rationalize; the naïve expectation is that a protein mutation that removes a residue interacting with a particular phosphate oxygen should make the mutant protein insensitive to the replacement of that oxygen with a methyl group. However, if, following Fersht and colleagues (44), we take account of interactions with solvent (H2O) and consider the entire inventory of interactions made and broken when the enzyme interacts with its substrate, it can be seen that the combination of mutant protein and non-cognate substrate may be favored because the energetic cost of desolvating the interacting groups will be least. The effect on cleavage of the methylphosphonate substrate was greatest for the K78A mutant, suggesting that Lys78 makes the most important interaction with the −1 phosphodiester position, while Tyr77 and Arg81 have lesser roles, perhaps in orienting the substrate or the Lys78 side chain.

The comparison of the 5′-phosphorylated and 5′-hydroxyl nicked substrates provides further evidence placing Arg81 close to the 5′ side of the cleavage position. The R81A mutant 5′ nuclease cleaved the phosphorylated substrate much more slowly than the non-phosphorylated substrate, whereas the wild-type 5′ nuclease and other mutants, including K78A, cleaved both substrates at similar rates. A plausible explanation is that Arg81 would normally neutralize the charge on the terminal 5′ phosphate. As noted above, the 5′ phosphate on a nicked substrate carries a greater negative charge than a backbone phosphodiester and may be bound differently at the 5′...
nuclease active site. This could explain why R81A had the greatest effect with the 5’ phosphate, but K78A had the greatest effect in the MP₄₁ comparison.

Data from the MP₄₁ methylphosphonate substitution suggested that Arg²⁰ may be close to the phosphodiester on the 3’ side of the cleavage site. With the wild-type 5’ nuclease, the MP₄₁ modification decreased the cutting at the expected position to such an extent that a normally minor cleavage pathway became the dominant pathway. As shown in Fig. 5B, this second cleavage pathway can be explained by rearrangement of the DNA junction to an alternative (and less favored) base pairing arrangement. Cleavage of the MP₄₁ substrate by the R20A mutant protein occurred in the same position as cleavage of the normal substrate by the wild-type 5’ nuclease, suggesting two possible roles for Arg²⁰. One is that Arg²⁰ interacts with the +1 phosphodiester, providing the major discrimination against the MP₄₁ substitution. In the absence of the Arg²⁰ side chain this discrimination is lost and the perfectly base paired configuration of the substrate (left side of Fig. 5B) dominates the reaction, giving the 11-mer. Alternatively, Arg²⁰ may be part of the binding site for an unpaired 3’ flap; removal of this side chain would then disfavor the alternative base pairing arrangement shown on the right of Fig. 5B and tip the balance in favor of the 11-mer product. We prefer the former explanation because the R20A mutation did not have any unusual effect on the rate of cleavage of substrates with a 3’ flap; rather, the rates of cleavage by both the wild-type and R20A nuclease were stimulated to a similar degree by the presence of a 1- or 2-nucleotide 3’ flap (data not shown).

The Active Site Carboxylates—Our previous study showed that the nine invariant carboxylates are important for the 5’ nuclease reaction but did not establish their individual roles (3). The results described here argue against any of the carboxylates being involved in DNA binding. Three mutations, D63A, D138N, and D185A, caused an increase in DNA binding affinity; as in the polymerase active site of Klenow fragment (32), this may indicate positions where removal of a negatively charged side chain close to the path of the phosphodiester backbone enhances binding of the DNA substrate, although not necessarily in a catalytically appropriate conformation. The side chains of Asp⁶³, Asp¹³⁸, and Asp¹⁸⁵ (shown in red in Fig. 2A) may therefore indicate regions of close approach of the phosphodiester backbone.

The original question then remains: why do nine carboxylates appear to be crucially important in the 5’ nuclease reaction, even though the precedents of the polymerase and 3’-5’ exonuclease active sites suggest that a smaller number of acidic side chains should suffice to coordinate the metal ions necessary for a two-metal-ion phosphoryl transfer mechanism (17, 45–47)? Our data do not completely rule out the possibility that a subset of these residues might interact with DNA, presumably via coordinated metal ions. A DNA binding role could have gone undetected in our experiments either because removal of a single carboxylate ligand was not sufficient to prevent binding of a metal ion under our assay conditions, or because the putative protein-DNA interaction is felt in the transition state and does not contribute to ground-state binding. An alternative scenario, that a subset of the carboxylates may appear to be important in the 5’ nuclease reaction because they serve to maintain the appropriate geometry in the active site region, is consistent with the rather variable $T_m$ values for some of the carboxylate mutants and for the increased stability of the 5’ nuclease domain in the presence of Mg²⁺ (Table II). It might also explain why the phenotypes of the carboxylate mutants, while broadly consistent in the different nuclease that have been studied, are not always in complete agreement. Moreover, if the conformation of the active site region is rather delicately balanced and therefore easily influenced by external factors, then this might account for the inconsistencies in coordination geometry and metal-metal distances seen in the 5’ nuclease crystal structures (3, 6, 22–24, 26).

Does the 5’ Nuclease Play an Active Role in Strand Displacement?—When the R20A and D188A mutant 5’ nucleases cleaved a 5’ flap substrate, the distribution of products differed from that of the wild-type enzyme (Fig. 6). These differences were particularly evident on a DNA substrate that lacked an upstream primer, ruling out any explanations based on interaction with the 3’ end of the primer. The data therefore suggest that the amount of melting at the junction of the 5’ flap and the DNA duplex is not simply a property of the DNA substrate, but that the 5’ nuclease may play an active role in opening the downstream duplex. The R20A mutant tended to give shorter products than the wild-type nuclease, implying a decrease in strand displacement activity, which may be related to the proposed interaction between Arg²⁰ and the DNA 3’ to the cleavage site. Conversely, the D188A mutant gave longer products, suggesting enhanced strand displacement activity. The D188A protein also gave longer cleavage products with the MP₄₁ substrate, and we therefore infer that the MP₄₁ result can be attributed to the strand displacement activity of D188A rather than to a specific interaction of Asp¹⁸⁸ with the +1 phosphate position.

Interactions with the 5’ Flap and Mechanism of DNA Binding—The Pol I 5’ nuclease requires a free 5’ end on the unpaired single-stranded tail or flap, as illustrated by the inhibition of 5’ nuclease cleavage by a terminal hairpin in this work, and by the loop substrates described in our previous work (3). Based on similar observations with other structure-specific nucleases, it has been proposed that the 5’ tail must be threaded through the enzyme (10, 27), and the loop or arch structures seen crystallographically fit well with this idea, being large enough to bind single-stranded, but not duplex, DNA (24–26). In several publications (26–28), the proposed mechanism of DNA binding has the 5’ nuclease entering via the free 5’ end of the single-stranded overhang or flap, and sliding to the cleavage point at the junction between single- and double-stranded regions (like threading a bead onto a string). However, the data do not exclude other mechanisms of forming a catalytically competent enzyme-DNA complex, such as binding to the junction and then inserting the single-stranded tail through the loop or arch of the enzyme (like threading a needle).

We have shown that cleavage of a substrate with a 30-nucleotide 5’ tail is much less efficient than cleavage of shorter tails (Table III), and we believe this indicates a situation in which formation of the catalytically competent complex is rate-limiting. Although the less efficient cleavage of the 30-nucleotide tail could also be attributed to weaker (rather than slower) binding of the 5’ nuclease, we do not favor this explanation because it is hard to understand why the binding affinity of a cleavage site adjacent to a 30-nucleotide tail should be substantially different from that with a 10-nucleotide tail. If formation of the complex is indeed rate-limiting for the substrate with the 30-nucleotide tail, then this would favor the second model for enzyme-DNA binding (threading the 5’ strand through the

---

A puzzling example is provided by the D200N mutant in T4 RNase H, which had full activity (19), although mutations in the homologous residues in the 5’ nucleases of E. coli (D188A) and Mycobacterium tuberculosis (D202N) were much more deleterious (18). Additionally, three carboxylate substitutions in human FEN-1 appear to affect DNA binding (in contrast with our results); however, the same three mutations probably influence the protein conformation, as judged by changes in protein solubility (21).
enzyme after binding to the junction). Our reasoning is that sliding along a 30-nucleotide tail should not be dramatically slower than sliding along a 10-nucleotide tail, whereas, with the alternative threading model, one might anticipate a cut-off between short-tailed substrates, whose binding would be trivially simple, and longer substrates, whose binding could present considerable entropic problems.

Regardless of the mechanism by which the 5′ single-stranded tail might be threaded through the nuclease, this process cannot involve close monitoring of the structure of the tail, since our data indicate that cleavage is remarkably insensitive to structural modifications. Similarly, mammalian FEN-1 tolerates a variety of modifications (27, 28), and it is inhibited only by those modifications which cause a substantial increase in size (bound protein, or duplex DNA), or which decrease flexibility (some cis-platinum adducts). The requirement for flexibility also supports the “needle-threading” model described above. More puzzling is the observation that FEN-1 can cleave a substrate within the 5′ flap strand (28); this raises doubts about threading mechanisms in general, although it is difficult to envisage an alternative model that accounts for the requirement for a free 5′ end. In vivo the question of threading may be moot, however, since a substantial fraction of the intermediates encountered by the structure-specific nuclease, particularly during lagging strand replication, are likely to have very short 5′ overhangs of only one or two nucleotides (9, 48–50).

REFERENCES
1. Kornberg, A., and Baker, T. A. (1992) DNA Replication, 2nd Ed., W. H. Freeman and Co., San Francisco
2. Gutman, P. D., and Minton, K. W. (1993) Nucleic Acids Res. 21, 4406–4407
3. Xu, Y., Derbyshire, V., Ng, K., Sun, X. C., Grindley, N. D. F., and Joyce, C. M. (1995) J. Biol. Chem. 270, 284–302
4. Robins, P., Pappin, D. J. C., Wood, R. D., and Lindahl, T. (1994) J. Biol. Chem. 269, 28535–28538
5. Lieber, M. R. (1997) BioEssays 19, 233–240
6. Mueser, T. C., Nossal, N. G., and Hyde, C. C. (1996) Cell 85, 1101–1112
7. Harrington, J. J., and Lieber, M. R. (1994) Genes Dev. 8, 1344–1355
8. Kim, K., Biade, S., and Matsumoto, Y. (1998) J. Biol. Chem. 273, 8842–8848
9. Lundquist, R., and Oliveira, B. (1982) Cell 31, 53–60
10. Lyamichev, V., Brou, M. A. D., and Dahlberg, J. E. (1993) Science 260, 778–783
11. Harrington, J. J., and Lieber, M. R. (1994) EMBO J. 13, 1235–1246
12. Bhagwat, M., Hobbs, L. J., and Nossal, N. G. (1997) J. Biol. Chem. 272, 28523–28530
13. Murante, R. S., Rumbaugh, J. A., Barnes, C. J., Norton, J. R., and Bambara, R. A. (1996) J. Biol. Chem. 271, 25888–25897
14. Lyamichev, V., Brou, M. A. D., Varvel, V. E., and Dahlberg, J. E. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 6143–6148
15. Xu, Y., Grindley, N. D. F., and Joyce, C. M. (2000) J. Biol. Chem. 275, 20949–20955
16. Shen, B., Qin, J., Hosfield, D., and Tainer, J. A. (1998) Trends Biochem. Sci. 23, 171–173
17. Joyce, C. M., and Steitz, T. A. (1994) Annu. Rev. Biochem. 63, 777–822
18. Astatke, M., Grindley, N. D. F., and Joyce, C. M. (1995) J. Biol. Chem. 272, 28531–28538
19. DeMott, M. S., Shen, B., Park, M. S., Bambara, R. A., and Zigman, S. (1996) J. Biol. Chem. 271, 30545–3066
20. Puglisi, J. D., and Tinoco, I. (1989) Methods Enzymol. 180, 304–325
21. Shen, B., Nolan, J. P., Sklar, L. A., and Park, M. S. (1997) J. Biol. Chem. 272, 9176–9179
22. DeMott, M. S., Shen, B., Park, M. S., Bambara, R. A. (1999) Nucleic Acids Res. 37, 2332–2338
23. Kim, Y., Eom, S. H., Wang, J., Lee, D. S., Suh, S. W., and Steitz, T. A. (1995) Nature 376, 612–616
24. Ceska, T. A., Sayers, J. R., Stier, G., and Suck, D. (1996) Nature 382, 90–93
25. Hwang, K. Y., Baek, K., Kim, H.-Y., and Cho, Y. (1996) Nat. Struct. Biol. 3, 707–713
26. Bornarth, C. J., Ranali, T. A., Henriksen, L. A., Wahl, A. F., and Bambara, R. A. (1999) J. Biol. Chem. 274, 30377–30383
27. Bergh, N., and Hillen, W. (1991) Methods Enzymol. 203, 3–12
28. Polesky, A. H., Grindley, N. D. F., and Joyce, C. M. (1995) J. Biol. Chem. 265, 14579–14591
29. Astatke, M., Grindley, N. D. F., and Joyce, C. M. (1996) J. Biol. Chem. 270, 1945–1954
30. Rimphanitchayakit, V., and Grindley, N. D. F. (1991) in A Laboratory Guide to in Vitro Studies of Protein-DNA Interactions (Jost, J. P., and Saluz, H. P., eds) Vol. 5, pp. 111–120, Birkhauser Verlag Basel, Basel
31. Setlow, P. (1974) Methods Enzymol. 31, 1355–1365
32. Kaplan, D. L. (2000) Methods Enzymol. 6148–60
33. Polesky, A. H., Steitz, T. A., Grindley, N. D. F., and Joyce, C. M. (1990) J. Biol. Chem. 265, 14579–14591
34. Wissman, A., and Hillen, W. (1991) Methods Enzymol. 209, 365–379
35. Maxam, A. M., and Gilbert, W. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 560–564
36. Shafritz, K. M., Sandigursky, M., and Franklin, W. A. (1998) Nucleic Acids Res. 26, 2593–2597
37. Tong, L. J., Bennett, S. E., and Mosbaugh, D. W. (1990) Nucleic Acids Res. 18, 7317–7322
38. Antao, V. P., and Tinoco, I., Jr. (1992) Methods Enzymol. 209, 12–24
39. Wu, X., Li, J., Li, X., Haich, C. L., Burgers, P. M., and Lieber, M. R. (1996) Nucleic Acids Res. 24, 2036–2043
40. Merkens, L. S., Bryan, S. K., and Moses, R. E. (1995) Biochim. Biophys. Acta 1209, 243–248
41. Merkens, L. S., Bryan, S. K., and Moses, R. E. (1995) Biochim. Biophys. Acta 1269, 176–186
42. Merkens, L. S., Bryan, S. K., and Moses, R. E. (1995) Biochim. Biophys. Acta 1269, 176–186
43. Merkens, L. S., Bryan, S. K., and Moses, R. E. (1995) Biochim. Biophys. Acta 1269, 176–186
44. Merkens, L. S., Bryan, S. K., and Moses, R. E. (1995) Biochim. Biophys. Acta 1269, 176–186
