Novel 3’-Ribonuclease and 3’-Phosphatase Activities of the Bacterial Non-homologous End-joining Protein, DNA Ligase D*

Received for publication, April 13, 2005, and in revised form, May 11, 2005
Published, JBC Papers in Press, May 15, 2005, DOI 10.1074/jbc.M504002200

Hui Zhu and Stewart Shuman‡
From the Molecular Biology Program, Sloan-Kettering Institute, New York, New York 10021

Pseudomonas aeruginosa DNA ligase D (PaeLigD) exemplifies a family of bacterial DNA end-joining proteins that consist of a ligase domain fused to a polymerase domain and a putative nuclease module. The LigD polymerase preferentially adds single ribonucleotides at blunt DNA ends and, as we show here, is also capable of adding up to 4 ribonucleotides to a DNA primer-template. We report that PaeLigD has an intrinsic ability to resect the short tract of 3’-ribonucleotides of a primer-template substrate to the point at which the primer strand has a single 3’-ribonucleotide remaining. The failure to digest beyond this point reflects a requirement for a 2’-OH group on the penultimate nucleoside of the primer strand. Replacing the 2’-OH by a 2’-F, 2’-NH₂, 2’-OCH₃, or 2’-H abolishes the resection reaction. The ribonucleotide resection activity resides within a 187-amino acid N-terminal nuclease domain and is the result of at least two component steps: (i) the 3’-terminal nucleoside is first removed to yield a primer strand with a ribonucleoside 3’-PO₄ terminus, and (ii) the 3’-PO₄ is hydrolyzed to a 3’-OH. The 3’-ribonuclease and 3’-phosphatase activities are both dependent on a divalent cation, specifically manganese. PaeLigD preferentially remodels the 3’-ends of a duplex primer-template substrate rather than a single strand of identical composition, and it prefers DNA primer strands containing a short 3’-ribonucleotide tract to an all-RNA primer. The nuclease domain of PaeLigD and its bacterial homologs has no apparent structural or mechanistic similarity to previously characterized nuclease. Thus, we surmise that it exemplifies a novel phosphoesterase family, defined in part by conserved residues Asp-50, Arg-52, and His-84, which are essential for the 3’-ribonuclease and 3’-phosphatase reactions.

DNA double-strand breaks can be repaired either via homologous recombination or non-homologous end joining (NHEJ)1 (1). Homologous recombination and NHEJ pathways coexist in eukaryotic cells. Eukaryal NHEJ requires a DNA end-binding protein (Ku) and a specialized ATP-dependent DNA ligase (LigIV) (2, 3). Until recently, it had been thought that bacteria rely exclusively on homologous recombination to repair double-strand breaks. Although this is the case for Escherichia coli, the detection of homologs of Ku in a subset of bacterial proteomes raised interest in a putative bacterial NHEJ pathway (4–6). There is now direct evidence that Mycobacteria have a vigorous NHEJ system that requires Ku and a specialized polyfunctional ATP-dependent DNA ligase (LigD) (7, 8). Mycobacterial NHEJ is highly mutagenic (~50% error rate), even when repairing complementary 5’-overhang ends. An analysis of the recombination junctions from individual NHEJ events revealed the participation of several DNA end-remodeling activities, including template-dependent fill-in of 5’-overhangs, non-templated addition of single nucleotides at blunt ends, and nucleolytic resection (8). The findings that a LigD deletion suppresses the overall frequency of NHEJ and increases the fidelity of the LigD-independent residual repair events are consistent with LigD serving as a direct catalyst of error-prone repair in vivo (8).

Bacterial LigD proteins are multifunctional enzymes composed of an ATP-dependent DNA ligase domain fused to a polymerase domain and a putative nuclease domain (6–10). The domain order varies among bacterial LigD proteins, e.g. Mycobacterium tuberculosis (Mtu) and Mycobacterium smegmatis LigD consist of an N-terminal polymerase domain, a central putative nuclease domain, and a C-terminal ligase domain, whereas Pseudomonas aeruginosa (Pae) LigD is composed of an N-terminal putative nuclease domain, a central ligase domain, and a C-terminal polymerase domain (Fig. 1). The C-terminal segment of MtuLigD (amino acids 412–759) and the central segment of PaeLigD (amino acids 188–527) are autonomous ligase domains capable of DNA nick sealing and ATP-dependent autoadenylation in vitro (7, 9). The ligase domain includes the six defining nucleotidyl transferase motifs of the ATP-dependent DNA ligase family (7, 11).

The polymerase function of MtuLigD is localized to an autonomous N-terminal domain LigD-(1–299), whereas the PaeLigD polymerase activity resides within an autonomous C-terminal segment, LigD-(533–840) (8–10). The Pol domain of LigD has primary structure and functional similarities to members of the Pol X family of nucleic acid polymerases (12–14). The polymerase activities of Pae- and MtuLigD are manifest either as non-templated nucleotide addition to a blunt-ended duplex DNA primer or templated extension of a 5’-tailed duplex DNA primer-template (8–10). Non-templated blunt-end addition in vitro is limited to the incorporation of only 1 or 2 nucleotides at the primer terminus. During templated synthesis, the primer is elongated to the end of the template strand and is then further extended with a single non-templated nucleotide. Fill-in and addition of a single non-templated nucleotide are the molecular signatures of mycobacterial NHEJ in vivo at 5’-overhang double-strand breaks (DSB) and blunt-end DSB, respectively (8). It is notable that rNTPs are preferred over dNTPs as substrates for non-templated blunt-end addition. This property led to speculation that the initial insertions

* This work was supported by National Institutes of Health Grant GM63611. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
‡ An American Cancer Society Research Professor. To whom correspondence should be addressed. E-mail: s-shuman@ski.mskcc.org.
1 The abbreviations used are: NHEJ, non-homologous end joining; PaeLigD, Pseudomonas aeruginosa DNA ligase D; Mtu, Mycobacterium tuberculosis; DTT, dithiothreitol; DSB, double-strand break; rNMP, ribonucleoside monophosphate.

This paper is available on line at http://www.jbc.org
preceding the strand-sealing step of NHEJ of a blunt DSB might involve nRp incorporation and that the ability of LigD to use rNTPs as substrates might be advantageous for the repair of chromosomal double-strand breaks that arise in quiescent cells, insofar as the dNTP pool might be limiting when bacteria are not actively replicating.

There is less known at present about the nuclease function of bacterial LigD. Fig. 1 shows an alignment of the amino acid sequences of the putative nuclease domains of the LigD proteins from eight species of bacteria (P. aeruginosa, Agrobacterium tumefaciens, Bradyrhizobium japonicum, M. avium, Bradyrhizobium japonicum, M. loti, and N. farcinica (Nfa), M. maezii (Mma), and M. acetivorans (Mac)). The conservation of residues in the Nuc domain that were changed to alanine are highlighted in shaded boxes.

Fig. 1. Primary structure of the nuclease domain of Pseudomonas LigD and related proteins from bacteria and archaea. The LigD polypeptide is depicted in schematic form with the N terminal on the left and the C terminus on the right. The amino acid sequences of the N-terminal segment of PaeLigD (Pae) is aligned to the sequences of homologous polypeptides encoded by A. tumefaciens (Atu), B. bronchiseptica (Bbr), M. tuberculosis (Mtu), M. avium (Mav), M. loti (Mlo), and N. farcinica (Nfa), M. maezii (Mma), and M. acetivorans (Mac). Gaps in the alignment are indicated by dashes. Conserved residues are indicated by dots. The conserved residues of the Nuc domain were pooled and diluted with an equal volume of buffer B (50 mM Tris-HCl, pH 7.5, 1 mM DTT, 0.1% Triton X-100). An aliquot (40 μl) containing 50 μM each ATP, GTP, CTP, and UTP, and the primer-template was preannealed to 1 μM of [γ-32P]ATP, purified by native gel electrophoresis, and annealed to a 4-fold excess of an unlabeled primer-template. The primer-template reactions were quenched by adjusting the mixtures to 10 mM EDTA and 3 mM MgCl2 and subjected to gel electrophoresis. Activity was monitored by autoradiography and was expressed as the incorporation of radioactivity into the primer-template.
48% formamide. The products were resolved by electrophoresis through a 15-cm 18% polyacrylamide gel containing 7 μm urea in TBE (90 mM Tris borate, 2.5 mM EDTA). The products were visualized by autoradiography. The fraction of input primer extended was determined by scanning the gel with a Fujifilm BAS-2500 imaging apparatus.

Nuclease Assay—Reaction mixtures (10 μl) containing 50 mM Tris-HCl, pH 7.5, 5 mM DTT, 0.5 mM MnCl2, 0.5 pmol 5'-32P-labeled D10R2 or other primer-template as specified and enzyme were incubated at 37 °C for 20 min. The reactions were quenched by adjusting the mixtures to 7 mM EDTA and 31% formamide. The products were resolved by electrophoresis through a 40-cm 18% polyacrylamide gel containing 7 μm urea in TBE. The products were visualized by autoradiography and quantified by scanning the gel with a Fujifilm BAS-2500 imaging apparatus.

3'-Phosphatase Assay—The 5'-32P-labeled, 3'-phosphate-terminated D10R1-p and D9R1-p strands were prepared by digesting the 5'-32P-labeled D10R2 and D9R3 primer oligonucleotides with ribonuclease A. Reaction mixtures (100 μl) containing 50 pmol radiolabeled D10R2 or D9R3 oligonucleotide, 5 mM DTT, and 20 μg RNase A were incubated for 20 min at 37 °C. The mixtures were extracted with phenol-chloroform. The oligonucleotides were precipitated with ethanol in the presence of glycogen carrier and then annealed to a 4-fold excess of an unlabeled complementary 24-mer DNA strand to generate a blunt-5'-tailed molecule shown in Fig. 2A. When the primer contained two terminal ribonucleotides, the primer-utilization efficiency of RNA primer was reduced compared with the DNA primer, e.g. D11R1 (Fig. 2A). The 3'-phosphatase reaction mixtures (10 μl) containing 50 pmol Tris-HCl, pH 7.5, 5 mM DTT, 0.5 mM MnCl2, 0.25 pmol 5'-32P-labeled D10R1-p or D9R1-p primer-template and enzyme as specified were incubated at 37 °C for 20 min. The reactions were quenched by adjusting the mixtures to 7 mM EDTA and 31% formamide. The products were resolved by electrophoresis through a 40-cm 18% polyacrylamide gel containing 7 μm urea in TBE. The products were visualized by autoradiography.

RESULTS

Primer-template-directed RNA Polymerization by the Pol Domain of PaeLigD—Previous studies showed that rNTPs are preferred over dNTPs as substrates for non-templated blunt-end addition by PaeLigD (9). Here we queried whether the PaeLigD polymerase could incorporate ribonucleotides during a templated fill-in synthesis reaction. The DNA primer-template used was composed of a 5'-32P-labeled 12-mer DNA strand annealed to a complementary 24-mer strand to form the 5'-tailed molecule shown in Fig. 2A. Control experiments showed that the isolated PaeLigD Pol domain catalyzed dNTP-dependent extension of the primer strand until the overhang was filled in completely, after which an extra non-templated dNMP was added to the blunt 3'-end (Ref. 9 and data not shown). When dNTPs were replaced by rNTPs, the efficiency of primer utilization (defined as the percent of the primer strands elongated by at least one nucleotide) remained high, but the number of rNMPs added was reduced. At enzyme concentrations sufficient to elongate 85–95% of the DNA substrate, rNMP incorporation was limited to 1–3 nucleotides. Further increases in the enzyme concentration resulted in the accumulation of a predominant product with four added rNMPs. Only a very low fraction of the primers was extended beyond the n+4 template position. This result suggested that the LigD polymerase might be sensitive to the pentose sugar composition of the primer strand.

To address this point, we used a substrate consisting of a 5'-32P-labeled 12-mer RNA hybridized to the 24-mer DNA template strand (Fig. 2E). The efficiency of RNA primer utilization for RNA synthesis was reduced compared with the DNA primer, e.g. only 12–14% of the ends were extended by levels of polymerase that sufficed to extend 95% of the DNA primer (Fig. 2E versus A). Raising the enzyme concentration resulted in up to 49% primer utilization, but only 1 rNMP was added to the RNA primer at all levels of input enzyme tested. This finding shows that PaeLigD polymerase strongly prefers a DNA primer terminus versus an RNA primer.

The basis for the RNA interference effect was revealed by testing a series of 12-mer primers consisting of 11, 10, or 9 deoxyribonucleotides at their 5'-ends plus 1, 2, or 3 ribonucleotides at their 3'-ends. We observed that substitution of even a single ribonucleotide at the 3'-OH terminus (in the D11R1 substrate shown in Fig. 2B) reduced the efficiency of primer utilization compared with the all-DNA D12 primer. Moreover, most D11R1 primers that were extended at the lower enzyme concentration were elongated by 1 nucleotide, and the product distribution at the highest enzyme concentrations included a mixture of 13-, 14-, 15-, and 16-mer species, but nothing larger (Fig. 2B). When the primer contained two terminal ribonucleo-
tides (D10R2 in Fig. 2C), the primer utilization efficiency declined again, and the major product was the n + 1 species, even at high enzyme concentrations. Increasing the RNA tract to 3 nucleotides suppressed primer utilization even further, to the point that the D9R3 substrate phenocopied the all-RNA R12 primer-template construct (Fig. 2D). Together, these data explain the restricted addition of only 4 nucleotides at a DNA primer terminus, as follows: (i) productive use by the PaeLigD Pol domain of the primer 3’-OH terminus depends on contacts with deoxynucleotides at the three terminal positions of the primer strand; (ii) each serial rNMP incorporation event diminishes the ability of the ribo-extended primer to undergo the next-round of reaction with an rNTP; and (iii) after 3 ribonucleotides are added, activity is suppressed strongly and effectively limited to only one more cycle of extension under the conditions analyzed. Thus, the impediment to the RNA polymerase activity of LigD is imposed by the primer, not the rNTP substrate. Consistent with this model, we found that RNA tracts at the 3’-terminus also suppressed the efficiency of deoxynucleotide incorporation by the LigD Pol domain (data not shown). The implication is that LigD utilization of ribonucleotides for fill-in synthesis during NHEJ is plausible only for short repair tracts. This could include the fill-in events observed in vivo in M. smegmatis at double-strand breaks containing 4-nucleotide 5’-overhangs (8).

Removal of a 3’-Ribonucleotide by PaeLigD—An initial experiment showed that reaction of full-length PaeLigD with the D10R2 primer-template construct in the absence of nucleotides resulted in the shortening of the 5’-radiolabeled primer strand to yield a predominant product that migrated as an 11-mer (see below). No shorter products were detected when PaeLigD was reacted in parallel with the D12 or D11R1 primer-templates, suggesting that the reaction required a diribonucleotide 3’-terminus. To test whether this putative 3’-ribonucleotide was intrinsic to PaeLigD, we tracked the activity after sedimenting PaeLigD in a glycerol gradient. We showed previously that ligase and polymerase activities co sedimented with the PaeLigD protein, which was judged to be a monomer compared with internal standards (9). When sedimented by itself, the 97-kDa LigD polypeptide comprised a single discrete peak centered at gradient fractions 13–15 (Fig. 3A). The products of the reaction of the glycerol gradient fractions with the D10R2 primer template are shown in Fig. 3B. Fractions 13–15 contained the peak of activity that removed the 3’-terminal ribonucleotide. The peak fractions converted the input 12-mer D10R2 strand to a major species corresponding to an 11-mer D10R1, whereas the flanking fractions 11 and 17 generated a second more rapidly migrating species, which we will identify below as the terminal-phosphorylated 11-mer, D10R1-p. Based on cosedimentation, we would attribute the 3’-ribonucleotase activity to the recombinant PaeLigD protein.

An Autonomous 3’-Ribonuclease Domain of PaeLigD—We produced an N-terminal fragment, the LigD polypeptide spanning amino acids 1–187 (the putative nuclease domain) in E. coli, as a His6 fusion and purified the recombinant Nuc domain from a soluble bacterial lysate by Ni-agarose and cation exchange chromatography (Fig. 4A). We also produced in parallel a mutated version of the nuclease domain in which the invariant Hia-84 side chain was replaced by alanine. The H84A protein was purified by Ni-agarose chromatography. The wild-type and mutant LigD-(1–187) preparations were nearly homogeneous with respect to the ~29 kDa Pae polypeptide (Fig. 4A). The wild-type Nuc domain displayed the same 3’-processing activity on the D10R2 primer-template that was noted above for full-length PaeLigD (Fig. 4B). Moreover, the 3’-processing function was abolished by the H84A mutation, arguing that the observed activity was intrinsic to the Nuc domain (Fig. 4B).

The quaternary structure of the Nuc domain was examined by zonal velocity sedimentation in a 15–30% glycerol gradient (Fig. 5). Marker protein catalase (native size 248 kDa), bovine serum albumin (66 kDa), and cytochrome c (12 kDa) were included as internal standards. After centrifugation, the polypeptide compositions of the odd-numbered gradient fractions were analyzed by SDS-PAGE. The Nuc domain sedimented as a discrete peak (fractions 19–23) between bovine serum albumin and cytochrome c (Fig. 5A), as did the activity responsible for 3’ shortening of the D10R2 substrate (Fig. 5B). These results are consistent with a monomeric structure for the isolated Nuc domain.

Sequential Mechanism of 3’-Ribonucleotide Removal—A kinetic analysis of the reaction of the Nuc domain with the D10R2 primer-template is shown in Fig. 6A (and, in a separate experiment, in Fig. 8B). The 5’-labeled primer was rapidly converted to a fast migrating species (with apparent size of 10 nt) that comprised 52% of the total labeled material after 1–2 min and decayed steadily at 5, 10, 20, and 30 min. A second product, migrating as an 11-mer, comprised 15% of the total label after 1 min and increased steadily in abundance thereafter, concomitant with the decay of the “10-mer-sized” species. The kinetic data (plotted in Fig. 8C) support a precursor product relationship between the more rapidly migrating species
and the 11-mer end product. Because chain length obviously cannot increase with time in the absence of added nucleotides, we suspected that the initial faster migrating intermediate was an 11-mer D10R1-p molecule containing a phosphate at the processed 3'-end and that this intermediate was subsequently converted to an 11-mer D10R1 strand with a 3'-OH terminus. It is well established that a 3'-PO₄-terminated oligonucleotide migrates faster during denaturing PAGE than a 3'-OH polynucleotide of identical length and sequence. The kinetic data in Fig. 8 were fit using the CKS kinetic simulation program (version 1.0; IBM Corp.) to a unidirectional two-step reaction mechanism with rate constants of 0.02 and 0.0055 s⁻¹ for the ribonuclease and phosphatase steps, respectively.

If the proposed sequential mechanism is valid, then the PaeLigD Nuc domain should have a 3'-phosphatase activity.
uncoupled from prior resection of the 3'-terminal ribonucleotide.
To measure this activity, we generated a 5'-labeled D10R1-p strand by digesting D10R2 with RNase A. D10R1-p was then annealed to the complementary 24-mer DNA template strand to form the tagged duplex substrate shown in Fig. 4C. The D10R1-p strand (which comigrates with the initial intermediate generated during the D10R2 3'-processing reaction) was converted quantitatively by the wild-type Nuc domain to a more slowly migrating 11-mer species corresponding to D10R1-\textit{OH} (Fig. 4B). Note that the product of the 3'-phosphatase reaction comigrated with the 11-mer end product of the D10R2 processing reaction analyzed in the same gel (Fig. 4B). The conversion of D10R1-p to D10R1-\textit{OH} was abolished by the H84A mutation of the Nuc domain (Fig. 4B), which signifies that the 3'-phosphatase activity is intrinsic to LigD. This conclusion is supported by the findings that the 3'-phosphatase activity cosedimented with full-length LigD (Fig. 3C) and with the isolated Nuc domain (data not shown).

Further insights to the mechanism of 3' processing emerged from analysis of the reaction of the wild-type Nuc domain with the D9R3 primer-template (Fig. 4C), which was converted to two products that migrated faster (by 1-nucleotide steps) than the D10R2 reaction products analyzed in parallel (Fig. 4B). We surmise that the major end product is D9R1-\textit{OH} and the minor species is a D9R1-p reaction intermediate. The H84A mutation abolished formation of both products derived from the D9R3 substrate (Fig. 4B). To verify this reaction scheme, we generated a 5'-labeled D9R1-p strand by RNase A digestion of D9R3 and annealed the D9R1-p strand to the 24-mer DNA to form the substrate shown in Fig. 4C. The D9R1-p strand (which comigrates with the faster species generated during the D9R3 end resection reaction) was converted by the wild-type Nuc domain to a more slowly migrating species, D9R1-\textit{OH}, which comigrated with the major product of the D9R3 reaction (Fig. 4B). Again, the conversion of D9R1-p to D9R1-\textit{OH} was eliminated by the H84A mutation (Fig. 4B).

A kinetic analysis of the 3'-phosphatase reaction with the D10R1-p substrate is shown in Fig. 6C. The dephosphorylated product accumulated steadily with pseudo-first order kinetics and an apparent rate constant of 0.0046 s\(^{-1}\). This value agrees with the rate of 0.0055 s\(^{-1}\) estimated for the phosphatase step of the 3'-processing reaction with D10R2 substrate.

The kinetic analysis of the processing of the D9R3 substrate shown in Fig. 6B is complicated by the fact that one of the reaction intermediates, D9R2-p, comigrates with the end-product, D9R1-\textit{OH}. The results are compatible with the scheme depicted in Fig. 6B whereby the Nuc domain incises the phosphodiester of the terminal (rC)p(rC) dinucleotide to give the 11-mer D9R2-p, which is then converted to D9R1-p, which is in turn dephosphorylated to form the D9R1-\textit{OH} end product. A longer exposure of the gel revealed only trace amounts of a D9R2-\textit{OH} species at 5 and 10 min, which disappeared at later times (not shown).

**Divalent Cofactor Requirement**—The 3'-ribonucleotide resection activity of the \textit{Pae}LigD Nuc domain was optimal at pH 5.5–8.0 in 50 mM Tris acetate or Tris-HCl buffer (data not shown) and was strictly dependent on a divalent cation, specifically manganese (Fig. 4D). A manganese titration experiment showed that the 3'-ribonucleoside activity was optimal at 0.2–0.8 mM MnCl\(_2\) (not shown). Magnesium, zinc, and cadmium were ineffective at 0.5 mM concentration (Fig. 4D). Higher concentrations of magnesium (1, 2.5, 5, and 10 mM) were also ineffective (not shown). Cadmium, cobalt, and copper (0.5 mM) were capable of sustaining reduced activity that resulted in the formation of the initial D10R1-p product, little of which was converted to the D10R1-\textit{OH} end product (Fig. 4D). The 3'-phosphatase activity also required a divalent cation, again specifically manganese (Fig. 4D). Magnesium, zinc, and cadmium were inactive, whereas calcium, cobalt, and copper supported low levels of dephosphorylation (Fig. 4D).

**Requirement for the Template DNA Strand**—The dependence of the 3'-ribonucleoside activity on the DNA template strand was gauged by comparing product formation as a function of input Nuc domain for the D10R2 primer-template substrate versus the D10R2 single-stranded oligonucleotide (Fig. 7). The extent of product formation on the primer-template substrate increased with Nuc domain concentration. At enzyme levels sufficient to consume 20–70% of the input-labeled strand, the products consisted of a mixture of D10R1-p and D10R1-\textit{OH} strands (Fig. 7, A and D). At the highest level of enzyme, nearly all the substrate was converted to D10R1-\textit{OH}. Activity on the single-stranded D10R2 substrate was clearly diminished, and there was a strong skew in the product distribution toward accumulation of the D10R1-p intermediate, with little conversion to the D10R1-\textit{OH} end product (Fig. 7, B and D). The specific activity of the initial ribonucleoside excision reaction, gauged by the consumption of the D10R2 strand as function of enzyme in the linear range, was about 5-fold higher for the primer-template than for the primer alone. However, the specific activity in generating dephosphorylated D10R1 was about 35-fold higher for the primer-template than for the primer alone. These results showed that the \textit{Pae}LigD nuclease displays a strong bias for a 3' terminus within a duplex nucleic acid compared with a single-stranded oligonucleotide and hinted that the 3'-phosphatase might be especially sensitive to the secondary structure of the substrate.

The latter issue was addressed directly by comparing the 3'-phosphatase activity of the Nuc domain on a D10R1-p primer-template versus a D10R1-p single strand (Fig. 7C). The 3'-phosphatase-specific activity on the primer-template substrate was ∼200-fold greater than on the primer alone. Thus, the 3'-phosphatase reaction displays a fairly stringent requirement for the template DNA strand.

**Requirement for a 2'-OH at the Incised Phosphodiester**—At levels of input Nuc domain that sufficed to convert all of the input D10R2 primer-template to the D10R1 product, we observed virtually no shortening of either the all-DNA D12 primer-template or the D11R1 primer-template containing only a single 3'-ribonucleotide (Fig. 8A). These results indicate that a ribose sugar is required at the penultimate nucleotide position. The role of the 2'-OH at the penultimate position was probed using a series of chemically modified primer-templates in which the 12-mer 5'-PO\(_4\) strand consisted of 10 deoxynucleotides at the 5'-end, 1 ribonucleotide at the 3' terminus, and either a 2'-O-methyl, a 2'-amine, or a 2'-fluorine nucleotide at the penultimate position. The instructive finding was that all of the 2' modifications of the penultimate nucleoside abolished the 3'-ribonucleoside activity of \textit{Pae}LigD (Fig. 8A). These results argue for a direct catalytic role of the 2'-OH at the incised phosphodiester, which could entail (i) its action as a nucleophile to attack the phosphoryl group to yield a 2',3'-cyclic phosphate intermediate, which is subsequently hydrolyzed to a 3'-phosphate; (ii) coordination of the divalent cation cofactor; or (iii) assistance in activation of a nucleophilic water in its attack on the phosphorus center. The prospect that the 2'-OH could act only indirectly by affecting the sugar pucker of the penultimate nucleoside appears unlikely, given that a 2'-F is effective in promoting an RNA-like C3'-endo conformation, even more so than is a 2'-OCH\(_3\) (15).

The findings that LigD efficiently removes one or two ribonucleotides from the 3'-end of a DNA primer-template (Figs. 4A and 8A) raises the question of whether and how LigD would...
process an all-RNA strand in a primer-template substrate. We observed that the Nuc domain was less active in shortening the R12 primer-template than it was in resecting either the D10R2 or D9R3 substrates, as gauged by the extent of decay of the radiolabeled input strand (Fig. 8). Moreover, whereas the enzyme resected the D10R2 and D9R3 substrates to the last remaining 3'-H11-ribonucleotide, the products of the reaction with the R12 substrate were limited to only two major species, migrating at positions of an 11- and 10-mer, respectively (Fig. 8). (Note that the all-RNA 12-mer strand migrates more slowly than the DNA primer strands of the same length and base composition.) A kinetic analysis of the processing of the R12 primer-template in Fig. 8, B and C, with a parallel analysis of the reaction with D10R2 serving as the positive control. The “10-mer” RNA species was the first to appear and accumulated over 20 min to an extent of about 27% of the total labeled material. The longer “11-mer” species accumulated after an initial lag phase. We surmise that the initial product is a 3'-phosphorylated R11-p strand, which is subsequently dephosphorylated to a more slowly migrating R11-OH product. Only trace amounts of product shorter than the 10-mer were formed at the 20 and 30 min times (comprising 3–6% of the total label). Two relevant points emerge from this analysis: (i) the sequential reaction mechanism at the terminal (rN)p(rN) dinucleotide applies whether or not the 5'-ribonucleoside resection is DNA or RNA, and (ii) the initial increment of resection is a mononucleoside, even when there are multiple internal (rN)p(rN) linkages available. The product distributions of the D10R2 and R12 reactions are plotted in Fig. 8C. A kinetic simulation of the R12 data yielded estimated rate constants of 0.00055 and 0.0008 s⁻¹ for the first ribonuclease and phosphatase steps, respectively. Comparison to the simulation of the D10R2 kinetic data indicates that ribonuclease and phosphatase reactions are ~35- and 7-fold faster on the D10R2 primer than on the all-RNA strand.

**Mutational Inactivation of the 3'-Ribonuclease and 3'-Phosphatase**—The nuclease domain of PaeLigD displays primary structure similarity to other bacterial LigD proteins, but it does not score against any biochemically established nuclease families when subjected to a PSI-BLAST search at NCBI. Thus, we surmise that it exemplifies a novel phosphoesterase family. We showed above that activity is strictly dependent on His-84, which is conserved in all of the LigD Nuc-like polypeptides aligned in Fig. 1. To identify other candidate components of the active site, we extended the alanine scan to 4 other invariant amino acids, Asp-50, Arg-52, Glu-54, and Glu-82 (highlighted in shaded boxes in Fig. 1). The mutant Nuc domains were produced in E. coli and purified from soluble bacterial extracts. SDS-PAGE analysis revealed comparable purity for the wild-type and mutant proteins (Fig. 9A). The 3'-ribonuclease activities were assayed in parallel using the D10R2 primer-template substrate at a level of input enzyme (1.2 g) that was saturating for the wild-type Nuc domain; this allowed for ready detection of the most severe catalytic defects. We found that the D50A and R52A proteins (as well as the H84A mutant) were virtually inert in generating either the D10R1-p or D10R1-OH reaction products (Fig. 9B). On the other hand, the E54A protein generated the D10R1-OH end product to a similar extent as the wild-type Nuc domain (Fig. 9B). Whereas the E82A mutant was active in 3'-ribonucleoside resection, as gauged by consumption of the input strand, the product distribution was strongly skewed toward accumulation of the D10R1-p species (Fig. 9B), suggesting that the mutation selectively affected the 3'-phosphatase activity.
DISCUSSION

Pseudomonas LigD has a complex domain architecture composed of nuclease, ligase, and polymerase modules. It is homologous to mycobacterial LigD, which plays a critical role in an NHEJ pathway of DNA repair (8). Here we show that the Pol domain of PaeLigD is able to add ribonucleotides to a DNA primer-template, but this reaction is limited to incorporation of only a few rNMPs because the primer-template is rendered progressively less active as ribonucleotides accumulate at the 3'-end. We found previously that LigD prefers rNTPs when performing blunt-end additions in vitro, which is a signature feature of bacterial NHEJ in vivo (8, 9). The present study hints that ribonucleotides might also be utilized for short-patch fill-in synthesis prior to ligation. Bebenek et al. (16) have recently reported that yeast DNA polymerase IV (a Pol X family member implicated in yeast NHEJ) is also adept at adding a ribonucleoside to a DNA primer-template.

The major finding here is that PaeLigD has an intrinsic ability to resect the short tract of 3'-ribonucleotides of a primer-template substrate to the point at which the primer strand has a single 3'-ribonucleotide remaining. The 3'-resection activity is resident within an autonomous 187-amino acid N-terminal nuclease domain of PaeLigD. The 3'-ribonuclease activity is the composite of at least two component steps: (i) the 3'-terminal ribonucleoside is first removed to yield a primer strand with a 3'-phosphate group which is subsequently degraded by the 3'-phosphatase activity. A, aliquots (5 μg) of the Ni-agarose preparations of wild-type Nuc domain and the indicated Ala mutants were analyzed by SDS-PAGE. Polypeptides were visualized by staining the gel with Coomassie Blue dye. The positions and sizes (in kDa) of marker polypeptides are indicated on the left. B, nuclease reaction mixtures containing 0.5 pmol 32P-labeled D10R2 primer-template and 1.2 μg (~50 pmol) of wild-type (WT) or mutant protein as specified were incubated for 20 min at 37 °C. C, phosphatase reaction mixtures containing 0.25 pmol 32P-labeled D9R1-p primer-template and 1.2 μg of wild-type or mutant protein as specified were incubated for 20 min at 37 °C. D, a reaction mixture (100 μl) containing 12 μg (~500 pmol) of E82A protein and 5 pmol 32P-labeled D10R2 primer-template was incubated at 37 °C. Aliquots (10 μl) were withdrawn at the times indicated and quenched immediately with EDTA/formamide.

Direct measurement of 3'-phosphate removal from the D9R1-p primer-template at a level of input enzyme saturating for wild-type Nuc showed that activity was virtually abolished by the D50A, R52A, E82A, and H84A mutations, whereas the E54A change was benign (Fig. 8C). These results suggest that the active sites for the ribonuclease and phosphatase functions are likely to overlap, with both activities being reliant on Asp-50, Arg-52, and His-84. However, the finding that the loss of Glu-82 selectively impaired the 3'-phosphatase hints that the
ribonucleoside 3'-PO₄ terminus, and (ii) then the 3'-PO₄ is hydrolyzed to a 3'-OH. This sequential mechanism is supported by (i) kinetic evidence for a precursor-product relationship of the 3'-PO₄ and 3'-OH strands, (ii) the capacity of the Nuc domain to directly hydrolyze a primer-template with a ribonucleoside 3'-PO₄ end at a rate similar to the 3'-phosphatase step of the composite ribonucleotide resection reaction, and (iii) the fact that alterations of the substrate or the enzyme can selectively impair the 3'-phosphomonoesterase step and thereby result in accumulation of the 3'-PO₄ strand.

The 3'-ribonuclease and 3'-phosphatase activities are both dependent on a divalent cation, specifically manganese. Magnesium is conspicuously unable to support either activity. This metal specificity is notable given that the polymerase activity of the PaeLigD is also manganese-dependent (9) and in light of the recent report by Daly et al. (17) that the accumulation of intracellular manganese by Deinococcus facilitates the extreme resistance of this bacterium to ionizing radiation.

The 3'-nuclease activity of PaeLigD ceases when a single ribonucleotide remains at the 3' terminus of the primer-template. The failure to digest beyond this point reflects the stringent requirement for a 2'-OH group on the penultimate nucleoside of the primer strand. Replacing the 2'-OH by a 2'-F, 2'-NH₂, 2'-OCH₃, or 2'-H abolishes the phosphodiesterase reaction. A simple explanation for these modification interference effects is that the enzyme catalyzes scission of the phosphodiester via a 2',3'-cyclic phosphate intermediate. Alternatively, the 2'-OH could be required for coordination of the metal cofactor. Further studies will be required to discriminate the two models.

To our knowledge, there is no known nuclease activity with the properties we describe here for PaeLigD, to wit: (i) it catalyzes metal-dependent scission of the P-O5'-bond of a 3'-phosphodiester to yield a 3'-PO₄, (ii) it separately catalyzes metal-dependent hydrolysis of a P-O3'-bond of a phosphomonoester, (iii) it acts preferentially on a duplex primer-template substrate rather than a single strand of identical composition, and (iv) it prefers DNA primer strands containing a short 3'-OH terminus. Alternatively, this 3'-deoxynucleoside 3'-PO₄ and 3'-OH strands, (ii) the capacity of the Nuc domain to directly hydrolyze a primer-template with a ribonucleoside 3'-monophosphate, with no requirement for a 2'-OH (22, 23).

The 3'-phosphatase activity of LigD differs from that of the bacteriophage T4, mycobacteriophage, and baculovirus polynucleotide kinase-phosphatase enzymes, insofar as the viral enzymes can use magnesium as the divalent cation cofactor and they readily dephosphorylate single-stranded oligonucleotides (or 3'-nucleotide monophosphates) with no dependence on a template strand (24–26). The 3'-phosphatase domains of the viral polynucleotide kinase-phosphatase enzymes belong to the DxDxT superfamily of acylphosphatases (25–28), to which the Nuc domain of LigD bears no structural similarity.

Our results indicate that PaeLigD exemplifies a novel class of 3'-remodeling enzyme. The ribonucleotide resection activity is potentially relevant to LigD function in bacterial NHEJ, insofar as we find that the strand-joining activity of PaeLigD is stimulated by the presence of a single ribonucleotide at the 3'-OH terminus.²

REFERENCES

1. Cromie, G. A., Connelly, J. C., and Leach, D. R. F. (2001) Mol. Cell 8, 1163–1174
2. Lieber, M. R., Ma, Y., Pannicke, U., and Schwarz, K. (2003) Nat. Rev. Mol. Cell. Biol. 4, 712–726
3. Downs, J. A., and Jackson, S. P. (2004) Nat. Rev. Mol. Cell. Biol. 5, 367–378
4. Aravind, L., and Koonin, E. V. (2001) Genome Res. 11, 1365–1374
5. Doherty, A. J., Jackson, S. P., and Weller, G. R. (2003) PERSPECTS 560, 156–188
6. Weller, G. R., Kyesla, B., Roy, R., Tonkin, L. M., Scanlan, E., Della, M., Devine, S. K., Day, J. P., Wilkinson, A., di Fascia, F., Devine, K. M., Bowser, R. P., Jeggo, P. A., Jackson, S. P., and Doherty, A. J. (2002) Science 297, 1686–1689
7. Gong, C., Martins, A., Bongiorno, P., Glickman, M., and Shuman, S. (2004) J. Biol. Chem. 279, 20994–20996
8. Gong, C., Bongiorno, P., Martins, A., Stephanoh, N. C., Zhu, H., Shuman, S., and Glickman, M. S. (2005) Nat. Struct. Mol. Biol. 12, 304–312
9. Zhu, H., and Shuman, S. (2005) J. Biol. Chem. 280, 418–427
10. Della, M., Palmbo, P. L., Tseng, H. M., Tonkin, L. M., Daley, J. M., Topper, L. M., Pitcher, R. S., Tomkinson, A. E., Wilson, T. E., and Doherty, A. J. (2004) Science 306, 683–688
11. Zhu, H., and Lima, C. D. (2004) Curr. Opin. Struct. Biol. 14, 757–764
12. Augustin, M. A., Huber, R., and Kaiser, J. T. (2001) Nat. Struct. Biol. 8, 57–61
13. Ito, N., Nureki, O., Shirouzu, M., Yokoyama, S., and Hanaoka, F. (2003) Genes Cells 8, 913–923
14. Lips, G., Weiner, A. O., von Scheven, G., Buchen, C., and Kramer, P. (2004) Nat. Struct. Mol. Biol. 11, 157–162
15. Guschlbauer, W., and Jankowski, K. (1980) Nucleic Acids Res. 8, 1421–1433
16. Bebek, K., Garcia-Diaz, M., Patishall, S. R., and Kunel, T. A. (2005) J. Biol. Chem. 280, 20051–20058
17. Daly, M. J., Gaidamakova, E. K., Matrosov, V. V., Vasilenko, A., Zhai, M., Vankutewaran, A., Hess, M., Omelchenko, M. V., Kostandartinas, H. H., Makarov, K. S., Wackett, L. P., Frederickson, J. K., and Ghosal, D. (2004) Science 306, 1025–1028
18. Raines, R. T. (1996) Chem. Rev. 96, 1045–1065
19. Zuo, Y., and Deutscher, M. P. (2001) Nucleic Acids Res. 29, 1017–1026
20. Cheng, Z. F., and Deutscher, M. P. (2002) J. Biol. Chem. 277, 21624–21629
21. Cheng, Y., and Patel, D. J. (2004) J. Mol. Biol. 343, 305–312
22. Richardson, C. C., and Kornberg, A. (1984) J. Biol. Chem. 259, 224–250
23. Richardson, C. C., Leberman, I. R., and Kornberg, A. (1984) J. Biol. Chem. 239, 251–258
24. Cameron, V., and Uhlenbeck, O. C. (1977) Biochemistry 16, 5120–5126
25. Zhu, H., Yin, S., and Shuman, S. (2004) J. Biol. Chem. 279, 26358–26369
26. Martins, A., and Shuman, S. (2004) J. Biol. Chem. 279, 18220–18231
27. Wang, L. K., and Shuman, S. (2002) Nucleic Acids Res. 30, 1073–1080
28. Gichard, E. A., Pelletier, J., Wilson, G., and Stoddard, B. L. (2002) Structure 10, 1249–1260

² H. Zhu and S. Shuman, unpublished observation.
