Structures of Bacterial Polynucleotide Kinase in a Michaelis Complex with Nucleoside Triphosphate (NTP)-Mg²⁺ and 5'-OH RNA and a Mixed Substrate-Product Complex with NTP-Mg²⁺ and a 5'-Phosphorylated Oligonucleotide

Ushati Das, Li Kai Wang, Paul Smith, Annum Munir, Stewart Shuman

Molecular Biology Program, Sloan-Kettering Institute, New York, New York, USA

Clostridium thermocellum polynucleotide kinase (*CthPnk*), the 5'-end-healing module of a bacterial RNA repair system, catalyzes reversible phosphoryl transfer from a nucleoside triphosphate (NTP) donor to a 5'-OH polynucleotide acceptor, either DNA or RNA. Here we report the 1.5-Å crystal structure of *CthPnk*-D38N in a Michaelis complex with GTP-Mg²⁺ and a 5'-OH RNA oligonucleotide. The RNA-binding mode of *CthPnk* is different from that of the metazoan RNA kinase Clp1. *CthPnk* makes hydrogen bonds to the ribose 2'-hydroxyls of the 5' terminal nucleoside, via Gln51, and the penultimate nucleoside, via Gln83. The 5'-terminal nucleobase is sandwiched by Gln51 and Val129. Mutating Gln51 or Val129 to alanine reduced kinase specific activity 3-fold. Ser37 and Thr80 donate functionally redundant hydrogen bonds to the terminal phosphodiester; a S37A-T80A double mutation reduced kinase activity 50-fold. Crystallization of catalytically active *CthPnk* with GTP-Mg²⁺ and a 5'-OH DNA yielded a mixed substrate-product complex with GTP-Mg²⁺ and 5'-PO₄ DNA, wherein the product 5' phosphate group is displaced by the NTP γ phosphate and the local architecture of the acceptor site is perturbed.

Polynucleotide kinase (Pnk) proteins are a widely distributed class of cellular and virus-encoded nucleic acid repair enzymes that convert 5'-OH termini into 5'-PO₄ ends that can be sealed by RNA or DNA ligases (1–18). Pnk proteins are members of the P-loop phosphotransferase superfamily; they catalyze metal-dependent transfer of the γ phosphate of an NTP donor to a 5'-OH polynucleotide acceptor. They also execute the “reverse kinase” reaction in which a polynucleotide 5'-phosphate is transferred to an NDP. Pnk enzymes differ with respect to their nucleobase preferences for the NTP donor, varying from high specificity for GTP (19) to nonspecific utilization of any NTP as the substrate for the forward kinase reaction or any NDP as the substrate for the reverse reaction (9, 20). Pnk proteins also display distinctive polynucleotide substrate preferences *in vitro*, being either DNA specific (15–18), RNA specific (7–9), or nonselective for DNA versus RNA (1, 2, 10, 12). Recent interest in Pnk proteins has been sparked by the discoveries that inactivating mutations in human DNA and RNA kinase enzymes cause severe neurological developmental defects (21–24).

The Pnk proteins of bacteriophage T4 and the bacterium *Clostridium thermocellum* have been extensively characterized, biochemically and structurally (1–6, 12, 13, 20, 25, 26). Although both kinases can phosphorylate 5'-OH RNA and DNA ends *in vitro*, their biological functions are dedicated to RNA repair, inasmuch as the kinase modules in both cases are fused in *cis* to a 3'-end-healing RNA phosphoesterase and are coupled, functionally or physically, to an ATP-dependent RNA ligase. The T4 Pnk phosphatase (Pnpk) collaborates with T4 RNA ligase 1 to repair tRNA damage inflicted during the *Escherichia coli* antiviral response (27). *C. thermocellum* Pnpk (*CthPnpk*) is a component of a two-subunit RNA repair “machine,” the other subunit being Hen1 (28). The end-healing and end-sealing activities reside within three autonomous domain modules of the *CthPnpk* polypeptide: N-terminal kinase, central phosphatase, and C-terminal

ligase (12, 13, 29–31). Homologous Pnpk-Hen1 RNA repair systems are present in many diverse bacterial taxa (32).

Our structural and functional studies of *CthPnk* have illuminated apparently universal principles of the polynucleotide kinase mechanism while providing insights into the particular donor and acceptor specificity of the bacterial Pnk clade (13, 20, 26). To date, we have reported atomic resolution crystal structures of *CthPnk* as NTP-Mg²⁺ donor complexes with ATP, GTP, CTP, UTP, and dATP that define the phosphate donor site (13, 20). These structures revealed the principles of nucleobase-nonspecific NTP substrate utilization (shared with T4 Pnk), whereby the *CthPnk* establishes an identical network of ionic and hydrogen-bonding contacts to the α, β, and γ phosphates of whichever NTP donor is bound, while the interactions of the nucleoside moiety (in *anti* conformation) are limited to a π cation stack of the nucleobase on an arginine side chain (Arg116 in *CthPnk*).

To capture a structure of the Michaelis complex of *CthPnk*, containing an NTP donor, a metal cofactor, and a 5'-OH polynucleotide acceptor (26), we introduced a minimally perturbing modification of the enzyme (changing Asp38 to Asn) that precludes phosphoryl transfer. By this approach, we determined a crystal structure of *CthPnk*-D38N in a complex with GTP-Mg²⁺ and a 5'-OH oligodeoxynucleotide _{HO}CpCpTpGpT. In the Michaelis complex, the O5' nucleophile is situated 3.0 Å from the GTP γ phosphorus, where it is coordinated by Asn38 and is apical to the bridging β phosphate oxygen of the GDP leaving group. We

Received 8 August 2014 Accepted 20 September 2014

Published ahead of print 29 September 2014

Address correspondence to Stewart Shuman, s-shuman@ski.mskcc.org.

Copyright © 2014, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JB.02197-14

also solved the structure of *CthPnk*-D38N in a complex with GDP-Mg²⁺ and a 5'-PO₄ oligonucleotide, pCpCpTpGpT (26). In the product complex, the transferred phosphate has undergone stereochemical inversion and Asn38 coordinates the 5'-bridging phosphate oxygen of the oligonucleotide. The D38N enzyme is poised for catalysis but cannot execute because it lacks Asp38, which is thereby implicated as the essential general base catalyst that abstracts a proton from the 5'-OH during the kinase reaction. Asp38 serves as a general acid catalyst during the reverse kinase reaction by donating a proton to the O5' leaving group of the 5'-PO₄ strand. A mechanism of O5' activation by a conserved aspartate general base is also invoked for the catalysis of phosphoryl transfer by T4 Pnk, mammalian Pnk, and Clp1 (3–5, 17, 18, 33).

The _{HO}DNA acceptor binding mode of *CthPnk* is distinctive (26). The first two nucleotides of the acceptor strand bind in a deep groove on the enzyme surface so as to place the terminal 5'-OH group next to the NTP γ phosphate. The DNA-bound substrate and product complex structures indicate that the 5' end of the phosphoacceptor must be single stranded to access the *CthPnk* active site. Comparisons to the structures of T4 Pnk in complexes with _{HO}DNA oligonucleotides (25, 26) highlight the extreme divergence of their acceptor sites, starting from the second nucleotide of the DNA strands. From the perspective of the nucleic acid substrate, the salient difference is that there is no stacking of consecutive nucleobases in the *CthPnk* structure (26). In contrast, the T4 Pnk-DNA structures (25) reveal acceptor base stacking but with different patterns of stacking according to the sequence of the acceptor strand.

The goal of the present study was to capture a structure of *CthPnk*-D38N in a Michaelis complex with NTP-Mg²⁺ and a _{HO}RNA acceptor strand and thereby gauge (i) whether the bacterial kinase makes atomic contacts with the ribose hydroxyls (it does) and (ii) whether the binding mode is significantly different for DNA versus RNA (it is not). A secondary goal was to crystallize catalytically active *CthPnk* in the presence of NTP, Mg²⁺, and the 5'-OH acceptor, with a view to capturing a product complex formed *in situ*. Whereas the latter crystallization experiment was successful, to our surprise it yielded a mixed donor substrate-acceptor product complex, *CthPnk*-GTP-Mg²⁺-pDNA, with significant perturbations at the active site. This structure prompts speculation regarding the product dissociation steps comprising the bi-bi reaction pathway.

MATERIALS AND METHODS

***CthPnk* purification and mutagenesis.** The pET28b-His₁₀Smt3-*CthPnk*-(1–170) expression plasmid encoding the Pnk domain was described previously (13). Alanine mutations were introduced into the expression vector by quick-change PCR with *Pfu* DNA polymerase, and the Pnk inserts were sequenced to confirm the presence of the desired mutation and the absence of unwanted coding changes. The plasmids were transformed into *E. coli* BL21(DE3). Recombinant protein production was induced with IPTG, and the *CthPnk* proteins were purified as described previously (20). In brief, the His₁₀Smt3-*CthPnk* proteins were recovered from soluble bacterial lysates by Ni-agarose chromatography. The tag was removed by treatment with Smt3 protease Ulp1, and the tag-free *CthPnk* was separated from the cleaved His₁₀Smt3 tag by a second round of Ni-agarose chromatography. The *CthPnk* preparations were then adjusted to 10 mM EDTA, concentrated by centrifugal ultrafiltration, and gel filtered through a column of Superdex-200 equilibrated in 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 10% glycerol, 1 mM dithiothreitol

(DTT), 1 mM EDTA. The enzyme preparations were stored at –80°C. Protein concentrations were determined by using the Bio-Rad dye reagent with bovine serum albumin as the standard.

Crystallization, diffraction data collection, and structure determination. The tag-free kinase-inactive *CthPnk*-D38N mutant protein was purified as described previously (26). The protein solution (10.5 mg/ml) was adjusted to 2 mM GTP, 10 mM MgCl₂, and 1.1 mM 5'-OH RNA oligonucleotide (_{HO}CCUGU). Alternatively, the kinase-active *CthPnk* protein (7.8 mg/ml), purified as described previously (13), was adjusted to 2 mM GTP, 10 mM MgCl₂, and 1.1 mM 5'-OH DNA oligonucleotide (_{HO}CCTGT). After incubation for 10 min at 22°C, aliquots (2 μ l) of the protein solutions were mixed on coverslips with an equal volume of precipitant solution containing 100 mM sodium citrate (pH 5.0), 100 mM MgCl₂, 18 to 24% (vol/vol) PEG-6000. Crystals were grown at 22°C by hanging-drop vapor diffusion against a reservoir of the same precipitant solution. Single crystals were transferred to a precipitant solution with 15% glycerol prior to freezing the crystals in liquid nitrogen. Diffraction data were collected at NLSL beamline X25 equipped with a Pilatus 6M detector in 1,800 continuous increments of 0.2° each for the protein-RNA crystal and in 260 continuous increments of 1° each for the protein-DNA crystal. The data were integrated and reduced with MOSFLM and AIMLESS/SCALA. The *CthPnk*-RNA crystal diffracted to ~1.3 Å, but the data were anisotropic (mean $I/\sigma I$ dropped below 2.0 along the h axis at 1.5 Å). Thus, during scaling, we used the recommended cutoffs from the scaling logs and cut off the data at 1.5 Å. Similarly, for the *CthPnk*-DNA crystal, the data were anisotropic (mean $I/\sigma I$ dropped below 2.0 along the k axis at 1.8 Å), even though the crystal diffracted to ~1.7 Å. Thus, during scaling, we used the recommended cutoffs from the scaling logs and cut off the data at 1.8 Å. Diffraction statistics are compiled in Table 1. The crystals were in the space group P2₁2₁2₁ and had two *CthPnk* protomers per asymmetric unit. To solve the structure of the RNA complex, the diffraction data were refined against the *CthPnk*-GTP-Mg²⁺-_{HO}DNA structure from which the DNA and waters had been removed (26). Following the placement of the Pnk dimer by PHASER and rigid-body refinement, the active sites of both protomers revealed clear and continuous F_o-F_c density for the 5'-OH RNA oligonucleotide in the acceptor site. To solve the structure of the DNA complex, the diffraction data were refined against the *CthPnk*-GTP-Mg²⁺ structure from which the waters had been removed (20). In this case, F_o-F_c density for a 5'-phosphate-terminated DNA was observed in the A protomer only. The models were iteratively rebuilt by hand in COOT (34) and refined in PHENIX (35). The final models had excellent geometry, no Ramachandran outliers, and no large F_o-F_c difference Fourier peaks (Table 1).

Accession numbers. The coordinates for the GTP-Mg²⁺-_{HO}RNA and GTP-Mg²⁺-pDNA complexes have been deposited in the RCSB protein structure database (PDB ID codes 4QM6 and 4QM7).

RESULTS

Michaelis complex of *CthPnk* with GTP-Mg²⁺ and a 5'-OH RNA strand. We grew crystals from a mixture of 0.6 mM *CthPnk*-D38N with 2 mM GTP, 10 mM MgCl₂, and 1.1 mM 5'-OH-terminated oligoribonucleotide _{HO}CpCpUpGpU and solved the structure at a 1.5-Å resolution (Table 1). The asymmetric unit comprised two *CthPnk* protomers organized as a kinase homodimer (13). The electron density maps revealed occupancy of the donor and acceptor sites of both protomers by GTP-Mg²⁺ and RNA oligonucleotide, respectively. The tertiary structure of the kinase-GTP-Mg²⁺-RNA complex is shown in Fig. 1A. The first four ribonucleotides of the 5'-OH RNA strand were modeled into F_o-F_c electron density, in which the nucleobases and ribose 2'-hydroxyls were well defined (Fig. 1B). A detailed stereo view of the active site is shown in Fig. 2A. The GTP donor and amino acids that contact the β and γ phosphates and the Mg²⁺ ion are depicted as stick models with

TABLE 1 Crystallographic data and refinement statistics^a

Parameter	<i>CthPnk</i> -GTP-Mg ²⁺ - _{HO} RNA	<i>CthPnk</i> -GTP-Mg ²⁺ -pDNA
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Unit cell dimensions (Å) at 130K	<i>a</i> = 52.3; <i>b</i> = 74.4; <i>c</i> = 119.6	<i>a</i> = 46.2; <i>b</i> = 67.2; <i>c</i> = 118.9
Diffraction data quality		
Resolution (Å)	39.9–1.5 (1.58–1.50)	23.1–1.8 (1.84–1.80)
Radiation source	NSLS X25	NSLS X25
Wavelength	1.00 Å	1.10 Å
Processing software	MOSFLM/SCALA	MOSFLM/AIMLESS
<i>R</i> _{sym} , % ^b	7.2 (19.9)	12.0 (32.2)
Unique reflections	75,553 (10,863)	34,969 (2,031)
Mean redundancy	11.9 (11.6)	8.7 (7.8)
Completeness, %	100.0 (99.9)	99.7 (99.4)
Mean <i>I</i> /σ <i>I</i>	22.2 (10.5)	13.4 (6.8)
Refinement and model statistics (<i>F</i> > 0)		
Resolution (Å)	39.4–1.5 (1.52–1.50)	23.1–1.8 (1.83–1.80)
Completeness, %	99.9 (99.8)	99.9 (98.6)
<i>R</i> _{free} / <i>R</i> _{work} , % ^c	17.5/15.3 (20.3/16.5)	18.8/16.3 (17.3/18.5)
RMSD, bonds/angles	0.007 Å/1.20°	0.014 Å/1.53°
Protomers/ASU ^d	2	2
Ramachandran plot	99.1% favored, no outliers	99.1% favored, no outliers
B factors, Å ² (overall/Wilson)	20.7/12.7	24.3/18.5
PDB ID	4QM6	4QM7
Model contents		
Protein residues	342	342
Heteroatoms	2 GTP, 2 Mg ²⁺ , 1 phosphate, 5 sodium, 2 oligonucleotides	2 GTP, 2 Mg ²⁺ , 1 oligonucleotide
Water	567	357

^a Standard definitions are used for all parameters. Numbers in parentheses refer to data in the highest-resolution bin. The refinement and geometric statistics come from PHENIX.

^b *R*_{sym} output as *R*_{merge} by SCALA or AIMLESS.

^c *R*_{free} sets consisted of 8% of data chosen at random against which structures were not refined.

^d ASU, asymmetric unit.

green carbons: the amino acids that contact the RNA acceptor strand are rendered as stick models with beige carbons.

As in previous *CthPnk*-NTP structures (13, 20), the GTP phosphate donor binds within a groove formed by the P-loop (¹⁵GSS GSGKST²³) and an overlying “lid” composed of two α helices and the connecting ¹²⁰RTDRQVE¹²⁶ peptide (Fig. 1A). The GTP α and β phosphates are engaged by a network of hydrogen bonds from Thr23 and the P-loop main-chain amides. As shown in Fig. 2A, the γ phosphate is anchored by the lid residues Arg120 and Arg123 and the P-loop Ser17. The P-loop Lys21 and the catalytic Mg²⁺ bridge the GTP β and γ phosphates. The P-loop Ser22 is a constituent of an octahedral Mg²⁺ coordination complex, which also includes nonbridging β and γ phosphate oxygens of GTP and three waters. The structural and mutational data for *CthPnk* suggest a catalytic mechanism whereby the P-loop lysine and Mg²⁺ stabilize the transition state on the NTP γ phosphorus, assisted by the lid arginines (12, 13, 20, 26). The same mechanism involving the metal and basic residues in the P-loop and lid applies to T4 Pnk (3–6) and the metazoan RNA-specific kinase Clp1 (33).

The RNA phosphoacceptor binding site. The O5′ nucleophile of the RNA acceptor strand is poised 3.1 Å from the GTP γ phosphorus and apical to the bridging β phosphate oxygen of the GDP leaving group (O5′-P-O angle = 161°) (Fig. 1A and 2A). Asn38 Nδ coordinates the O5′ atom (2.9 Å) in lieu of the native Asp38 carboxylate, thereby trapping the Michaelis complex with RNA for want of a general base. Figure 3 shows a superposition of the _{HO}RNA (_{HO}CpCpUpGp) and _{HO}DNA (_{HO}CpCpTpGp) acceptors

from the Michaelis complexes of *CthPnk*-D38N-GTP-Mg²⁺ with _{HO}RNA and _{HO}DNA ligands of identical nucleobase sequence. The path of the strands and the conformations of the first three nucleotides are virtually identical for the RNA and DNA substrates. The fourth guanosine nucleotide deviates slightly, via rotation at the C5-O5-5′P linkage. Consequently, the 3′-phosphate of the guanosine nucleotide (the last group modeled into electron density) is offset by 2.5 Å in the two crystal structures. The Pnk homodimers in the DNA-bound and RNA-bound Michaelis complexes align with a root mean square deviation (RMSD) of 0.22 Å over 342 Cα atoms, signifying no difference in protein conformation according to the nucleic acid substrate. We conclude that *CthPnk*, which is adept at phosphorylating either RNA or DNA *in vitro*, does so via a single polynucleotide binding mode.

That said, the present structure of the RNA Michaelis complex does reveal two RNA-specific contacts. The first is a hydrogen bond from Gln51 Oε to the ribose 2′-OH of the 5′-OH nucleoside (Fig. 2A and 3). This interaction of Gln51 is in addition to its sugar-nonspecific stacking on the first cytosine nucleobase of RNA and DNA ligands (Fig. 3). The second RNA-specific contact is a bifurcated hydrogen bond from the ribose 2′-OH of the second nucleoside to the Gln83 side chain amide (Fig. 3). This is in addition to the hydrogen bond from Gln83 Nε to the C2′pU³ phosphate that is observed in both RNA and DNA complexes.

Other pertinent features of 5′-OH acceptor recognition in the RNA Michaelis complex include (i) anchoring of the first phosphodiester by Arg41, Ser37, and Thr80 side chains and a Thr80

FIG 1 Kinase tertiary structure and electron density for _{HO}RNA and pDNA oligonucleotides. (A) Stereo view of the tertiary structure of the *CthPnk*-D38N-GTP-Mg²⁺-_{HO}RNA complex, shown as a ribbon trace with magenta β strands, cyan helices, and beige intervening loops and turns. The N and C termini of the polypeptide are indicated. The GTP in the donor site and the 5'-OH RNA in the acceptor site are depicted as stick models with green and gray carbons, respectively. Mg²⁺ is depicted as a magenta sphere. The lid domain that engages the NTP donor is indicated. (B) Stereo view of the F_o-F_c electron density map (red) of the 5'-OH RNA _{HO}C¹pC²pU³pG⁴p of the *CthPnk*-D38N-GTP-Mg²⁺-_{HO}RNA complex contoured at 3.0 σ. (C) Stereo view of the F_o-F_c electron density map (red) of the 5'-PO₄ DNA pC¹pC²pT³p of the *CthPnk*-GTP-Mg²⁺-pDNA complex contoured at 3.0 σ. The oligonucleotides are shown as stick models. The nucleobases and 5' termini are labeled.

main-chain amide; (ii) a salt bridge between Asn38 and the essential Arg41 side chain that orients the general base; (iii) sandwiching of the first nucleobase between Val129 and Gln51; (iv) a network of direct (via His133) and water-mediated (via Lys132) contacts to the O2, N3, and N4 atoms of the first cytosine nucleobase; (v) engagement of the U³pG⁴ phosphate by hydrogen bonds from Tyr128 and the C¹ nucleobase; and (vi) π stacking interactions of Phe58 and Tyr128 on the C² and G⁴ nucleobases, respectively (Fig. 2A).

Mutational analysis of the phosphoacceptor site. Initial mutagenesis of amino acids in the phosphoacceptor site that contact the 5'-terminal _{HO}N¹p nucleotide showed that alanine substitutions for Asp38 and Arg41, which tether the 5'-OH and 3'-phos-

phate, respectively, abolished kinase activity (13). In contrast, alanines in lieu of Ser37 or Thr80, which also contact the _{HO}N¹p phosphate (Fig. 2A), elicited only a modest (3-fold) decrease in kinase specific activity compared to wild-type Pnk (26). Here, we tested the impact of the S37A-T80A double mutation, thinking that the hydrogen bonds from these hydroxyamino acids to the first phosphodiester might be functionally redundant. We assayed the extent of label transfer from 100 μM [γ-³²P]ATP to 10 μM 10-mer _{HO}CCTGTATGAT acceptor as a function of input protein (Fig. 4). (Note that the nucleobase sequence of the 5' half of the acceptor substrate was identical to that of the ligands used for cocrystallization of the Michaelis complexes.) Specific activities of the wild-type and mutant kinase were derived by linear regression curve fitting in Prism and are plotted in bar graph format in Fig. 4. The S37A-T80A double mutant was 2% as active as wild-type Pnk (Fig. 4), implying that Ser37 and Thr80 do indeed make functionally redundant contacts to the acceptor strand.

The 5' nucleobase is sandwiched by van der Waals contacts with Val129 and Gln51 (Fig. 2A). Here we found that V129A and Q51A mutations reduced kinase-specific activity to one-third of the wild-type level (Fig. 4). Thr54 makes van der Waals contacts to the second ribose sugar and the third nucleobase, plus a hydrogen bond from Thr54-Oγ to Arg41 (Fig. 2A). We found that replacing Thr54 with alanine reduced kinase specific activity to 41% of the wild-type Pnk (Fig. 4). The modest effects of subtracting the Val129, Gln51, and Thr54 side chains that stack on the first and third nucleobases are concordant with the 4-fold decrement seen when His133 was changed to alanine (26) but contrast with the more drastic 12-fold decrease in specific activity when the π stack of Phe58 on the second nucleobase was subtracted by alanine mutation (26).

A mixed substrate-product complex of active *CthPnk* with GTP-Mg²⁺ and pDNA. We grew crystals of a mixture of 0.4 mM catalytically active *CthPnk* with 2 mM GTP, 10 mM Mg²⁺, and 1.1 mM _{HO}DNA oligonucleotide _{HO}CpCpTpGpT. The crystals belonged to the space group P2₁2₁2₁, with one *CthPnk* homodimer in the asymmetric unit. Whereas both *CthPnk* protomers had GTP-Mg²⁺ in the phosphate donor site, F_o-F_c density for DNA oligonucleotide was seen only in the acceptor site of the A protomer, into which we modeled the trinucleotide C¹pC²pT³p. The electron density map showed that the DNA oligonucleotide contained a 5'-phosphate terminus (Fig. 1C), signifying that the input _{HO}DNA strand had undergone phosphorylation by *CthPnk* *in situ*. Thus, the structure we solved was not that of a GDP-Mg²⁺-pDNA product complex. Rather, the presence of GTP-Mg²⁺ in the donor site indicated that the immediate GDP-Mg²⁺ product of the kinase reaction *in situ* must have dissociated from the enzyme and been replaced by fresh GTP-Mg²⁺ that was in excess in the crystallization solution. The 1.8-Å refinement statistics and model contents of the *CthPnk*-GTP-Mg²⁺-pDNA structure are provided in Table 1.

A stereo view of the active site of the *CthPnk*-GTP-Mg²⁺-pDNA complex is shown in Fig. 2B, in the same orientation and color scheme as the active site of the _{HO}RNA Michaelis complex depicted in Fig. 2A. The GTP-Mg²⁺ and its contacts are effectively identical in the two structures, except for the lid Arg123 side chain, which is reoriented in the GTP-pDNA complex so that it contacts both the GTP γ phosphate and the oligonucleotide 5'-PO₄. Figure 5 depicts the nucleotide and oligonucleotide ligands in the kinase product complex, obtained by cocrystallization of *CthPnk*-D38N

FIG 2 Active sites of the GTP-Mg²⁺-_{HO}-RNA complex (A) and GTP-Mg²⁺-pDNA complex (B). Stereo views are shown. GTP and amino acids comprising the NTP donor site are shown as stick models with green carbons. Mg²⁺ is a magenta sphere. Waters in the metal coordination complex are red spheres. The _{HO}-RNA and pDNA oligonucleotides are rendered as stick models with gray carbons. Amino acids comprising the oligonucleotide site are depicted as stick models with beige carbons. Electrostatic and hydrogen-bonding interactions are denoted by black dashed lines and van der Waals contacts by beige dashed lines. The magenta dashed line in panel A indicates the distance (3.1 Å) between the O5' nucleophile of the _{HO}-RNA acceptor strand and the GTP γ phosphorus.

with GDP-Mg²⁺ and pDNA (Fig. 5, top), aligned to those in the mixed substrate-product complex of active *CthPnk* solved presently (Fig. 5, bottom). In the D38N-pDNA product complex, which freezes the enzyme immediately after what would be the phosphoryl transfer step, the Mg²⁺ and Lys21 bridge the gap between the GDP β phosphate and DNA 5'-phosphate groups,

maintaining the atomic contacts they make to the β and γ phosphate oxygens of the GTP substrate (Fig. 5). The GDP β phosphate and the DNA 5' phosphate are separated by 4 Å and are poised in what we construe to be both the product state of the forward kinase reaction and the Michaelis complex for the reverse kinase reaction.

The “extra” DNA 5' phosphate in the *CthPnk*-GTP-Mg²⁺-pDNA structure clashes with the GTP γ phosphate, which occupies its normal position in the donor site. The upshot of the clash is that the DNA 5' phosphate is displaced via rotation about the O5'-C5'-C4'-C3' torsion angle of the sugar, so that the DNA 5' phosphorus moves 4.9 Å (forward and to the right in the view in Fig. 5), placing the DNA 5'-phosphorus 5.8 Å away from the GTP γ phosphorus. The extra 5'-phosphate repels and reorients the Asp38 catalytic residue so that it no longer contacts Arg41 and now points away from the active site (Fig. 2B and 5). The position of Arg41 and its contacts with the first internucleotide phosphodiester of the DNA are unaffected by the extra 5' phosphate (Fig. 2B and 5).

The other salient change triggered by the extra phosphate is an outward displacement of the loop peptide (⁴⁶DDENDQ⁵¹) of the α3-loop-α4 module that lines the oligonucleotide binding site (Fig. 6). The dashed line in Fig. 6 shows how the loop in the GTP-Mg²⁺-pDNA complex moves by 4.8 Å at the Glu48 Cα atom relative to its position in the GTP-Mg²⁺-_{HO}-DNA complex. The Gln51 side chain, which is a direct constituent of the acceptor interface, is not in contact with pDNA in the mixed GTP-pDNA product complex but rather is disordered and could not be mod-

FIG 3 Comparison of _{HO}-RNA and _{HO}-DNA acceptors. Stereo view of the aligned 5'-OH RNA _{HO}-C¹pC²pU³pG⁴p acceptor (stick model with cyan carbons) and 5'-OH DNA _{HO}-C¹pC²pT³pG⁴p acceptor (stick model with gray carbons) from their respective Michaelis complexes with GTP-Mg²⁺. The Gln51 and Gln83 side chains in the RNA complex (stick models with beige carbons) make hydrogen bonds (red dashed lines) to the C¹ and C² ribose 2'-hydroxyl groups, respectively.

FIG 4 Structure-guided mutagenesis of the phosphoacceptor site. Kinase reaction mixtures (10 μ l) containing 50 mM Tris-HCl (pH 7.0), 10 mM MgCl₂, 5 mM DTT, 100 μ M [γ -³²P]ATP, 100 pmol of a 10-mer 5'-OH DNA oligonucleotide phosphoacceptor (H₂CCTGTATGAT), and serial 2-fold dilutions of wild-type or mutant kinases (range, 0.78 to 50 ng) were incubated for 30 min at 45°C. The reactions were quenched by adding an equal volume of 90% formamide, 50 mM EDTA, 0.01% bromophenol blue-xylene cyanol. The mixtures were analyzed by electrophoresis (at 7 W constant power) through a 15-cm 20% polyacrylamide gel containing 8 M urea in 45 mM Tris-borate, 1.2 mM EDTA. The ³²P-labeled DNAs were visualized and quantified by scanning the gel with a Fujifilm BAS-2500 imager. The extents of DNA phosphorylation were plotted as a function of input protein, with each datum being the average of three titration experiments $\pm$ the standard error of the mean (SEM). The kinase specific activities of the wild-type and mutant Pnk proteins were derived by linear regression curve fitting in Prism and are plotted in bar graph format. The error bars indicate the standard deviation of the curve fit.

eled into density beyond the β carbon (Fig. 6). Collectively, these changes indicate that replacing GDP-Mg²⁺ with GTP-Mg²⁺ weakens the interaction of *CthPnk* with the 5'-phosphorylated kinase reaction product.

DISCUSSION

The present study provides new insights into how *CthPnk* engages DNA versus RNA substrates. The Michaelis complex structures of *CthPnk* bound to NTP-Mg²⁺-_{HO}RNA and NTP-Mg²⁺-_{HO}DNA have identical polypeptide conformations, and the paths that the DNA and RNA single strands traverse are much the same, especially with respect to the three 5'-terminal nucleotides. The RNA Michaelis complex unveils two contacts to the ribose 2'-OH groups of the terminal and penultimate sugars (via Gln51 and Gln83, respectively), in addition to the sugar-nonspecific interactions of Gln51 and Gln83 with both DNA and RNA acceptor strands. Mutating Gln51 to alanine elicited a 3-fold decrease in kinase specific activity with a DNA substrate (Fig. 4), signifying that its sugar-nonspecific contacts play a modest role in 5'-OH acceptor recognition. Mutating Gln83 to alanine had no impact on kinase specific activity (26). Our structural studies rationalize available biochemical evidence that *CthPnk*, like T4Pnk, does not discriminate between single-strand RNA and DNA substrates *in vitro*, notwithstanding that the bacterial and phage kinases are RNA repair enzymes.

Previously, we discussed the distinctive single-strand DNA

FIG 5 Exchange of GTP-Mg²⁺ for GDP-Mg²⁺ in the kinase product complex. The structures of the *CthPnk*-D38N-GDP-Mg²⁺-pDNA product complex (PDB code 4MDE) and the wild-type *CthPnk*-GTP-Mg²⁺-pDNA complex were superimposed and offset vertically. The nucleotide, pDNA product, and Lys21, Arg41, and Asn38/Asp38 side chains are shown as stick models. Mg²⁺ is a magenta sphere. The 5' phosphorus atoms are colored differently in the two structures to highlight the displacement of the 5' phosphate when fresh GTP binds in the donor site.

binding modes of *CthPnk* versus T4 Pnk (26). The recently reported structure of the *Caenorhabditis elegans* RNA kinase Clp1 (*CeClp1*) bound to ADP-Mg²⁺ and a tetranucleotide 5'-OH RNA acceptor strand (33), together with the present structure of the *CthPnk*-GTP-Mg²⁺-_{HO}RNA complex, allows the first comparison of RNA binding styles in Pnk proteins from distantly related taxa. Figure 7 shows the *CeClp1* and *CthPnk* RNA complexes, aligned to superimpose the donor site nucleotides, magnesium ions, RNA 5'-OH groups, and aspartate general base residues that coordinate the O5' nucleophile. The figure highlights the different trajectories and conformations of the RNA acceptor strands. With respect to trajectory, the two RNA acceptors diverge immediately, by virtue of rotation of the RNA around the C4'-C5' bond of the terminal nucleoside, so that the first internucleotide phosphodiester in *CeClp1* projects downward in the view in Fig. 7, whereas the first phosphodiester projects upward in *CthPnk*. The single-stranded RNA in *CeClp1* adopts a right-handed helical conformation with 3'-endo sugar pucker akin to that of one strand of duplex RNA, by virtue of stacking interactions between the sequential RNA nucleobases, aided by π stacking of a tryptophan side chain on the 5'-terminal G¹ nucleobase and ionic interactions of Arg297 with the first and second internucleotide phosphodiesters (Fig. 7). There are no enzymatic contacts between *CeClp1* and the second, third, or fourth nucleobase. As noted by Dikfidan et al. (33), the helical single-strand RNA acceptor in the *CeClp1* acceptor site readily overlies one strand of a blunt-ended double-stranded nucleic acid, either DNA or RNA, with no steric clash. This nicely accounts for why Clp1 discriminates in favor of single-strand RNA acceptors versus single-strand DNA but has similar 5'-kinase activity on blunt-ended RNA and DNA substrates (33).

FIG 6 Structural perturbations in the acceptor site of the GTP-Mg²⁺-pDNA complex. The structures of the *CthPnk*-D38N-GTP-Mg²⁺-_{H₂O} DNA Michaelis complex (PDB code 4MDF; cyan) and the wild-type *CthPnk*-GTP-Mg²⁺-pDNA complex (green) were superimposed. The DNA ligands and selected side chains are shown as stick models in a stereo view highlighting changes in side chain position and orientation caused by the DNA 5'-phosphate. The blue line denotes the 4.8 Å displacement of the Glu48 C α atom of the loop segment connecting two α helices that form the acceptor-binding site.

In stark contrast anent *CthPnk*, the 5'-OH end of the nucleic acid must be single stranded to enter the acceptor site; the three terminal ribose sugars of the _{H₂O} RNA acceptor have a 2'-endo pucker, and the four nucleobases do not stack on one another but are

displayed out from the sugar-phosphate backbone (Fig. 7), where they pack against the surface of the enzyme via π stacking of C² on Phe58 and G⁴ on Tyr128 or van der Waals contacts between C¹ and Val129 and U³ with Thr54 (Fig. 2A).

FIG 7 Distinct RNA binding modes in *CeClp1* versus *CthPnk*. The *CeClp1* and *CthPnk* structures in complexes with 5'-OH RNA tetranucleotides were aligned so as to superimpose the donor site nucleotides (ADP in *CeClp1*; GTP in *CthPnk*), magnesium ions (magenta spheres), RNA 5'-OH groups, and catalytic aspartate side chains and then offset vertically. The figure highlights the completely different trajectories and conformations of the RNA acceptor strands.

CthPnk catalyzes a bisubstrate, biproduct reaction. With respect to substrate binding, our structures make it clear that *CthPnk* can engage NTP-Mg²⁺ in the donor site in the absence of a 5'-OH acceptor nucleic acid (13, 20). Although it is not known whether the kinetic mechanism is random bi-bi or ordered bi-bi, the ensemble of *CthPnk* structures prompts us to speculate along the following lines. The mixed substrate-product complex structure of the wild-type *CthPnk* with GTP-Mg²⁺ and pDNA captures the enzyme in a state after the phosphoryl transfer reaction, in which the pDNA product remains in the acceptor site but the GDP-Mg²⁺ product has dissociated from the donor site and fresh GTP-Mg²⁺ has bound in its place. The presence of both the GTP γ phosphate and the DNA 5'-phosphate causes a steric clash and electrostatic repulsion. The GTP γ phosphate is the “winner” in this clash of phosphates, occupying the same position as in the Michaelis complex. The product oligonucleotide 5'-phosphate is the “loser,” being displaced relative to its position in the true NDP-Mg²⁺-pDNA product complex, concomitant with displacement of the catalytic aspartate and the ⁴⁶DDENDQ⁵¹ loop that lines the oligonucleotide binding site. The structure suggests a model for ordered product dissociation whereby NDP-Mg²⁺ dissociates prior to 5'-phosphate nucleic acid. This is consistent with the inference that the NDP product is less avidly bound in the active site than NTP substrate, owing to the loss of multiple enzymatic contacts with the γ phosphate (from Ser17, Lys21, and Arg123) (26). Because these erstwhile contacts to the γ phosphate are transferred to the oligonucleotide 5'-phosphate in the product complex (26), one expects the 5'-phosphorylated product to be more tightly engaged by the kinase than the 5'-OH substrate. In that event, the entry of fresh NTP-Mg²⁺ into the donor site could trigger dissociation of the 5'-phosphorylated product by inducing the destabilizing distortions seen in the mixed substrate-product crystal structure.

ACKNOWLEDGMENTS

We thank Annie Heroux, Neil Whalen, and Agata Jacewicz for their assistance with data collection.

This research was supported by NIH grant GM42498.

REFERENCES

- Richardson CC. 1965. Phosphorylation of nucleic acid by an enzyme from T4 bacteriophage-infected *Escherichia coli*. Proc. Natl. Acad. Sci. U. S. A. 54:158–165. <http://dx.doi.org/10.1073/pnas.54.1.158>.
- Novogrodsky A, Hurwitz J. 1966. The enzymatic phosphorylation of ribonucleic acid and deoxyribonucleic acid: phosphorylation at 5'-hydroxyl termini. J. Biol. Chem. 241:2923–2932.
- Wang LK, Shuman S. 2001. Domain structure and mutational analysis of T4 polynucleotide kinase. J. Biol. Chem. 276:26868–26874. <http://dx.doi.org/10.1074/jbc.M103663200>.
- Wang LK, Shuman S. 2002. Mutational analysis defines the 5' kinase and 3' phosphatase active sites of T4 polynucleotide kinase. Nucleic Acids Res. 30:1073–1080. <http://dx.doi.org/10.1093/nar/30.4.1073>.
- Wang LK, Lima CD, Shuman S. 2002. Structure and mechanism of T4 polynucleotide kinase—an RNA repair enzyme. EMBO J. 21:3873–3880. <http://dx.doi.org/10.1093/emboj/cdf397>.
- Galburt EA, Pelletier J, Wilson G, Stoddard BL. 2002. Structure of a tRNA repair enzyme and molecular biology workhorse: T4 polynucleotide kinase. Structure 10:1249–1260. [http://dx.doi.org/10.1016/S0969-2126\(02\)00835-3](http://dx.doi.org/10.1016/S0969-2126(02)00835-3).
- Shuman S, Hurwitz J. 1979. 5' hydroxyl polyribonucleotide kinase from HeLa cell nuclei. J. Biol. Chem. 254:10396–10404.
- Weitzer S, Martinez J. 2007. The human RNA kinase hCp1 is active on 3' transfer RNA exons and short interfering RNAs. Nature 447:222–226. <http://dx.doi.org/10.1038/nature05777>.
- Jain R, Shuman S. 2009. Characterization of a thermostable archaeal polynucleotide kinase homologous to human Clp1. RNA 15:923–931. <http://dx.doi.org/10.1261/rna.1492809>.
- Zhu H, Yin S, Shuman S. 2004. Characterization of polynucleotide kinase/phosphatase enzymes from mycobacteriophages Omega and Cjw1 and vibriophage KVP40. J. Biol. Chem. 279:26358–26369. <http://dx.doi.org/10.1074/jbc.M403200200>.
- Martins A, Shuman S. 2004. Characterization of a baculovirus enzyme with RNA ligase, polynucleotide 5' kinase and polynucleotide 3' phosphatase activities. J. Biol. Chem. 279:18220–18231. <http://dx.doi.org/10.1074/jbc.M313386200>.
- Martins A, Shuman S. 2005. An end-healing enzyme from *Clostridium thermocellum* with 5' kinase, 2',3' phosphatase, and adenylyltransferase activities. RNA 11:1271–1280. <http://dx.doi.org/10.1261/rna.2690505>.
- Wang LK, Das U, Smith P, Shuman S. 2012. Structure and mechanism of the polynucleotide kinase component of the bacterial PNKP-Hen1 RNA repair system. RNA 18:2277–2286. <http://dx.doi.org/10.1261/rna.036061.112>.
- Wang LK, Schwer B, Englert M, Beier H, Shuman S. 2006. Structure-function analysis of the kinase-CPD domain of yeast tRNA ligase (Trl1) and requirements for complementation of tRNA splicing by a plant Trl1 homolog. Nucleic Acids Res. 34:517–527. <http://dx.doi.org/10.1093/nar/gkj441>.
- Jilani A, Ramotar D, Slack C, Ong C, Yang XM, Scherer SW, Lasko DD. 1999. Molecular cloning of the human gene, PNKP, encoding a polynucleotide kinase 3'-phosphatase and evidence for its role in repair of DNA strand breaks caused by oxidative damage. J. Biol. Chem. 274:24176–24186. <http://dx.doi.org/10.1074/jbc.274.34.24176>.
- Karimi-Busheri F, Daly G, Robins P, Canas B, Pappin DJC, Sgouros J, Miller GG, Fakhrai H, Davis EM, Le Beau MM, Weinfeld. 1999. M. Molecular characterization of human DNA kinase. J. Biol. Chem. 274: 24187–24194.
- Garces F, Pearl LH, Oliver AW. 2011. The structural basis for substrate recognition by mammalian polynucleotide kinase 3' phosphatase. Mol. Cell 44:385–396. <http://dx.doi.org/10.1016/j.molcel.2011.08.036>.
- Coquelle N, Haval-Shahriari Z, Bernstein N, Green R, Glover JNM. 2011. Structural basis for the phosphatase activity of PNKP on single- and double-stranded DNA substrates. Proc. Natl. Acad. Sci. U. S. A. 108: 21022–21027. <http://dx.doi.org/10.1073/pnas.1112036108>.
- Remus BS, Shuman S. 2014. Distinctive kinetics and substrate specificities of plant and fungal tRNA ligases. RNA 20:462–473. <http://dx.doi.org/10.1261/rna.043752.113>.
- Das U, Wang LK, Smith P, Shuman S. 2013. Structural and biochemical analysis of the phosphate donor specificity of the polynucleotide kinase component of the bacterial PNKP · Hen1 RNA repair system. Biochemistry 52:4734–4743. <http://dx.doi.org/10.1021/bi400412x>.
- Shen J, Gilmore EC, Marshall CA, Haddadin M, Reynolds JJ, Eyaid W, Bodell A, Barry B, Gleason D, Allen K, Ganesh VS, Chang BS, Grix A, Hill RS, Topcu M, Caldecott KW, Barkovich AJ, Walsh CA. 2010. Mutations in PNKP cause microcephaly, seizures and defects in DNA repair. Nat. Genet. 42:245–249. <http://dx.doi.org/10.1038/ng.526>.
- Reynolds JJ, Walker AK, Gilmore EC, Walsh CA, Caldecott KW. 2012. Impact of PNKP mutations associated with microcephaly, seizures and developmental delay on enzyme activity and DNA strand break repair. Nucleic Acids Res. 40:6608–6619. <http://dx.doi.org/10.1093/nar/gks318>.
- Schaffer AE, Eggens VR, Caglayan AO, Reuter MS, Scott E, Coufal NG, Silhavy JL, Xue Y, Kayserli H, Yasuno K, Rosti RO, Abdellateef M, Caglar C, Kashner PR, Cazemier JL, Weterman MA, Cantagrel V, Cai N, Zweier C, Altunoglu U, Satkin NB, Aktar F, Tuysuz B, Yalcinkaya C, Caksen H, Bilguvar K, Fu XD, Trotta CR, Gabriel S, Reis A, Gunel M, Baas F, Gleeson JG. 2014. CLP1 founder mutation links tRNA splicing and maturation to cerebellar development and neurodegeneration. Cell 157:651–663. <http://dx.doi.org/10.1016/j.cell.2014.03.049>.
- Karaca E, Weitzer S, Pehlivan D, Shiraishi H, Gogakos T, Hanada T, Jhangiani SN, Wiszniewski W, Withers M, Campbell IM, Erdin S, Isikay S, Franco LM, Gonzaga-Jauregui C, Gambin T, Gelowani V, Hunter JV, Yesil G, Koparir E, Yilmaz S, Brown M, Briskin D, Hafner M, Morozov P, Farazi TA, Bernreuther C, Glatzel M, Trattig S, Friske J, Kronnerwetter C, Bainbridge MN, Gezdirici A, Seven M, Muzny DM, Boerwinkle E, Ozon M, Clausen T, Tuschl T, Yuksel A, Hess A, Gibbs RA, Martinez J, Penninger JM, Lupski JR. 2014. Human CLP1 mutations alter tRNA biogenesis, affecting both peripheral and central nervous system function. Cell 157:636–650. <http://dx.doi.org/10.1016/j.cell.2014.02.058>.
- Eastberg JH, Pelletier J, Stoddard BL. 2004. Recognition of DNA substrates by bacteriophage T4 polynucleotide kinase. Nucleic Acids Res. 32: 653–660. <http://dx.doi.org/10.1093/nar/gkh212>.
- Das U, Wang LK, Smith P, Jacewicz A, Shuman S. 2014. Structures of bacterial polynucleotide kinase in a Michaelis complex with GTP · Mg²⁺ and 5'-OH oligonucleotide and a product complex with GDP · Mg²⁺ and 5'-PO₄ oligonucleotide reveal a mechanism of general acid-base catalysis and the determinants of phosphoacceptor recognition. Nucleic Acids Res. 42:1152–1161. <http://dx.doi.org/10.1093/nar/gkt936>.
- Amitsur M, Levitz R, Kaufman G. 1987. Bacteriophage T4 anticodon nuclease, polynucleotide kinase, and RNA ligase reprocess the host lysine tRNA. EMBO J. 6:2499–2503.
- Chan CM, Zhou C, Huang R. 2009. Reconstituting bacterial RNA repair and modification in vitro. Science 326:247. <http://dx.doi.org/10.1126/science.1179480>.
- Smith P, Wang LK, Nair PA, Shuman S. 2012. The adenylyltransferase domain of bacterial PNKP defines a unique RNA ligase family. Proc. Natl. Acad. Sci. U. S. A. 109:2296–2301. <http://dx.doi.org/10.1073/pnas.1116827109>.
- Wang LK, Smith P, Shuman S. 2013. Structure and mechanism of the 2',3' phosphatase component of the bacterial PNKP-Hen1 RNA repair system. Nucleic Acids Res. 41:5864–5873. <http://dx.doi.org/10.1093/nar/gkt221>.
- Wang P, Chan CM, Christensen D, Zhang C, Selvadurai K, Huang RH. 2012. Molecular basis of bacterial protein Hen1 activating the ligase activity of bacterial protein PNKP for RNA repair. Proc. Natl. Acad. Sci. U. S. A. 109:13248–13253. <http://dx.doi.org/10.1073/pnas.1209805109>.
- Jain R, Shuman S. 2010. Bacterial Hen1 is a 3' terminal RNA ribose 2'-O-methyltransferase component of a bacterial RNA repair cassette. RNA 16:316–323. <http://dx.doi.org/10.1261/rna.1926510>.
- Dikfidan A, Loll B, Zeymer C, Magler I, Clausen T, Meinhart A. 2014. RNA specificity and regulation of catalysis in the eukaryotic polynucleotide kinase Clp1. Mol. Cell 54:975–986. <http://dx.doi.org/10.1016/j.molcel.2014.04.005>.
- Emsley P, Cowtan K. 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. D 60:2126–2132. <http://dx.doi.org/10.1107/S0907444904019158>.
- Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger T, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC. 2002. PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr. D 58:1948–1954. <http://dx.doi.org/10.1107/S0907444902016657>.