Structure and mechanism of the 2',3' phosphatase component of the bacterial Pnkp-Hen1 RNA repair system

Li Kai Wang, Paul Smith and Stewart Shuman*

Molecular Biology Program, Sloan-Kettering Institute, New York, NY 10065, USA

Received February 13, 2013; Revised March 6, 2013; Accepted March 8, 2013

ABSTRACT

Pnkp is the end-healing and end-sealing component of an RNA repair system present in diverse bacteria from many phyla. Pnkp is composed of three catalytic modules: an N-terminal polynucleotide kinase, a central 2',3' phosphatase and a C-terminal ligase. The phosphatase module is a Mn$^{2+}$-dependent phosphodiesterase-monoesterase that dephosphorylates 2',3'-cyclic phosphate RNA ends. Here we report the crystal structure of the phosphatase domain of Clostridium thermocellum Pnkp with Mn$^{2+}$ and citrate in the active site. The protein consists of a core binuclear metallophosphoesterase fold (exemplified by bacteriophage λ phosphatase) embellished by distinctive secondary structure elements. The active site contains a single Mn$^{2+}$ in an octahedral coordination complex with Asp187, His189, Asp233, two citrate oxygens and a water. The citrate fills the binding site for the scissile phosphate, wherein it is coordinated by Arg237, Asn263 and His264. The citrate invades the site normally occupied by a second metal (engaged by Asp233, Asn263, His323 and His376), and thereby dislocates His376. A continuous tract of positive surface potential flanking the active site suggests an RNA binding site. The structure illuminates a large body of mutational data regarding the metal and substrate specificity of Clostridium thermocellum Pnkp phosphatase.

INTRODUCTION

RNA repair enzymes capable of sealing 2',3'-cyclic phosphate and 5'-OH ends are present in diverse taxa in all phylogenetic domains of life (1). Repair pathways anchored by ‘classic’ ATP-dependent RNA ligases entail the joining of 3'-OH and 5'-PO$_4$ termini. To satisfy the ligase specificity for 3'-OH/5'-PO$_4$ ends, the original broken 2',3'-cyclic phosphate and 5'-OH ends must be ‘healed’ before they can be sealed. Healing entails two discrete reactions: hydrolysis of the 2',3'-cyclic phosphate by a phosphoesterase enzyme to form a 3'-OH, and phosphorylation of the 5'-OH by a polynucleotide kinase enzyme to form a 5'-PO$_4$.

The Pnkp-Hen1 RNA repair system of Clostridium thermocellum (Cth) is encoded in an operon-like gene cassette (Figure 1A). Pnkp-Hen1 systems are present in diverse bacteria from 10 different phyla. CthPnkp is the end-healing and end-sealing component and is composed of three catalytic domains: N-terminal kinase, central phosphatase and C-terminal ligase (2). The kinase module catalyzes phosphoryl transfer from ATP to the 5'-OH RNA end. The phosphatase domain releases P$_i$ from 2'-PO$_4$, 3'-PO$_4$ or 2',3'-cyclic phosphate ribonucleotides (Figure 1B). The ligase domain reacts with ATP to form a covalent enzyme–AMP adduct, just as classic RNA ligases do during strand joining, but it is unable per se to seal RNA strands (2,3). The sealing function of the Pnkp ligase domain is activated by the N-terminal half of the Hen1 protein (4). The C-terminal half of Hen1 is an autonomous Mn$^{2+}$-dependent 3'-terminal ribose 2'-O-methyltransferase that installs a 2'-OCH$_3$ mark at the RNA repair junction before ligation, and thereby immunizes the sealed RNA against recurrent endonuclease cleavage (Figure 1B) (5–8).

The physical and functional properties of the Pnkp-Hen1 RNA repair system are distinctive vis à vis others that have been characterized (1,9,10). Thus, it is of interest to delineate the structure of the Pnkp-Hen1 holoenzyme and, en route, its constituent catalytic domains, to understand the principles of catalysis and substrate recognition, the relatedness of the Pnkp and Hen1 domains to other nucleic acid repair enzymes and the logic of the domain organization and its implications for orchestrating the healing and sealing steps.

To date, crystal structures have been solved for the isolated CthHen1 C-terminal methyltransferase domain.

*To whom correspondence should be addressed. Tel: +1 212 639 7145; Fax: +1 212 717 3623; Email: s-shuman@ski.mskcc.org

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

© The Author(s) 2013. Published by Oxford University Press. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
(a monomer) bound to the reaction product S-adenosylhomocysteine (6), the isolated CthPnkp C-terminal ligase domain (a monomer) in a complex with the ATP substrate and as the covalent ligase–AMP intermediate (3), and the ligase–AMP/C15Mg2+/Hen1(N) heterodimer (4). The structure of the ligase domain is distinctive, being composed of a classic ligase adenylyltransferase fold that is embellished by a unique α-helical insert module and a unique C-terminal α-helical module (3). The insert module of the ligase domain mediates the interaction with the N-terminal domain of Hen1, which is itself a novel tertiary structure (4).

The N-terminal 170-amino acid segment of CthPnkp is an autonomous polynucleotide kinase module. We recently reported the crystal structure of the kinase domain (a homodimer) bound to ATP·Mg2+ and ADP·Mg2+, reflective of substrate and product complexes, respectively, and used these structures to guide a mutational analysis of the kinase active site (11). Our findings highlighted structural and functional similarities to another RNA repair enzyme, phage T4 polynucleotide kinase (12,13), while revealing unique structural features of the bacterial polynucleotide kinase clade.

The 2',3'-phosphatase component of CthPnkp has been the subject of extensive biochemical characterization and mutational analyses (2,14–17). The enzyme is a Ni2+/Mn2+-dependent phosphodiesterase and monoesterase, active on nucleotides (2',3'-cyclic AMP, 3' AMP and 2' AMP) and generic substrates (bis-p-nitrophenyl phosphate, p-nitrophenyl phosphate and p-nitrophenyl phenylphosphonate). The phosphodiesterase and monoesterase reactions rely on overlapping but different ensembles of active site functional groups. The enzyme is remarkably plastic, insofar as the CthPnkp phosphatase can be transformed toward narrower metal and substrate specificities via mutations of the active site. For example, certain changes transform CthPnkp into a Mn2+-dependent phosphodiesterase devoid of monoesterase activity (15–17). The RNA 2',3'-cyclic phosphodiesterase activity of CthPnkp is thought to be most pertinent to RNA repair in light of the fact that 2',3'-cyclic phosphate termini are the predominant products of several known RNA damage pathways (17).

The primary structure of the CthPnkp phosphatase domain identifies it as a binuclear metallophosphoesterase—a vast enzyme superfamily distributed widely among taxa. A prototypal member is bacteriophage λ phosphatase (λ-Pase), which has been characterized structurally and biochemically (16,18–21). λ-Pase uses Mn2+ to catalyze phosphoester hydrolysis with a variety of substrates, including phosphopeptides, phosphoproteins, nucleoside 2',3'-cyclic phosphates and generic organic phosphomonoesters and diesters such as p-nitrophenyl phosphate and bis-p-nitrophenyl phosphate. Yet, the physiological substrate(s) and biological function of λ-Pase remain obscure. To our knowledge, only one other metallophosphoesterase enzyme besides CthPnkp plays a dedicated role in RNA metabolism, that being the RNA debranchoing enzyme Dbr1 (22). Here we report the crystal

Figure 1. The Pnkp-Hen1 RNA repair pathway. (A) The adjacent co-oriented ORFs encoding the Hen1 and Pnkp polypeptides comprise a bacterial RNA repair cassette conserved in taxa from 10 different bacterial phyla. Pnkp is a trifunctional RNA repair enzyme composed of 5'-OH polynucleotide kinase, 2',3'-phosphatase and ligase domains. Hen1 is composed of a C-terminal methyltransferase catalytic domain (MTase) fused to a unique N-terminal module that binds to the Pnkp ligase domain. (B) 2',3'-cyclic phosphate and 5'-OH ends are substrates for healing and sealing by Pnkp and Hen1. The 5' end is phosphorylated by the Pnkp kinase, and the 2',3'-cyclic phosphate is removed by the Pnkp phosphatase. The Hen1 methyltransferase installs a 2'-OCH3 mark at the terminal ribonucleoside before ligation of the ends. The ligation reaction is mediated by a heterodimer of the Pnkp ligase domain and the Hen1 N-terminal domain. The repair junction with the methyl mark is then resistant to scission by transesterifying endoribonucleases.
structure of the CthPnkp phosphatase domain, which verifies its membership in the metallophosphoesterase superfamily, while highlighting structural features that distinguish it from λ-Pase.

METHODS

Purification of CthPnkp-(171–424)

A pET-based expression plasmid encoding the phosphatase domain fused to an N-terminal His$_{10}$Smt3 module and a C-terminal His$_6$-tag was transformed into E. coli BL21(DE3). A 4-l culture derived from a single kanamycin-resistant transformant was grown at 37°C in Luria–Bertani medium containing 60 µg/ml kanamycin until the A$_{600}$ reached 0.6. The culture was then adjusted to 0.3 mM isopropyl-β-D-thiogalactoside and 2% (v/v) ethanol, and incubated for 17 h at 17°C with continuous shaking. Cells were harvested by centrifugation, and the pellet was stored at −80°C. All subsequent procedures were performed at 4°C. Thawed bacteria were resuspended in 130 ml of lysis buffer (50 mM Tris–HCl, pH 7.5, 1.2 M NaCl, 25 mM imidazole, 10% glycerol). Lysozyme, PMSF, benzamidine and Triton X-100 were added to final concentrations of 1 mg/ml, 1 mM, 1 mM and 0.2%, respectively. After mixing for ~50 min, the resulting lysate was sonicated to reduce viscosity, and insoluble material was removed by centrifugation for 45 min at 17 000 rpm in a Sorvall SS34 rotor. The soluble extract (130 ml; 2.7 g of protein) was applied to a 13-ml column of Ni-nitrotriacetic acid–agarose (Qiagen) that had been equilibrated with lysis buffer. The column was washed sequentially with 30 ml of 3 M KCl in buffer A (50 mM Tris–HCl, pH 7.5, 200 mM NaCl, 10% glycerol) and 30 ml of 100 mM imidazole in buffer A. The bound protein was serially step-eluted with buffer A containing 200 mM and 300 mM imidazole. The elution profile was monitored by SDS–PAGE. The peak phosphatase-containing fractions were pooled (185 ml; 550 mg of protein). A 60-ml aliquot containing 180 mg of protein was supplemented with 0.36 mg of His-tagged Smt3-specific protease Ulp1 and 1.5 ml of 200 mM EDTA. The mixture was then dialyzed overnight against 41 of 20 mM Tris–HCl, pH 7.5, 100 mM NaCl, 1 mM EDTA, and 10% glycerol. SDS–PAGE analysis of the dialysates indicated that the His$_{10}$Smt3 module had been cleaved from the phosphatase domain. The dialysate was adjusted to 50 mM imidazole and then applied to a 7-ml phosphocellulose column that had been equilibrated with 20 mM Tris–HCl, pH 7.5, 70 mM NaCl, 50 mM imidazole, and 10% glycerol. The phosphatase was recovered in the flow-through fractions, free of the His$_{10}$Smt3 protein, which was retained on the column. The phosphatase was concentrated by centrifugal ultrafiltration to 8.3 ml of an 11.5 mg/ml solution. Two separate aliquots (4 ml each) were gel-filtered through a Superdex-200 column equilibrated with 20 mM Tris–HCl, pH 7.5, and 70 mM NaCl, 10% glycerol; the A$_{280}$ profile revealed a single peak of included protein at a position consistent with a monomer. The peak fractions were pooled and concentrated by centrifugal ultrafiltration to 3.4 ml of a 24 mg/ml solution and stored at −80°C.

Crystallization

The phosphatase solution was adjusted to 10 mM MnCl$_2$ and incubated for 10 min before aliquots (2 µl) were mixed on a cover slip with an equal volume of precipitant solution containing 100 mM Na-citrate, pH 5.6, 33.5% (w/v) PEG-3350, and 15% DMSO. Crystals were grown at 22°C by hanging-drop vapor diffusion against a reservoir of the same precipitant solution. Rod-shaped crystals appeared after 1 d; single crystals were frozen directly in liquid nitrogen.

Diffraction data collection and structure determination

The structure of CthPnkp-(171–424) was solved using single-wavelength anomalous diffraction (SAD) data from a single crystal as follows. Diffraction data were collected at NSLS beamline X25 using a Pilatus-6M detector at a wavelength of 1.846 Å intended to maximize anomalous scattering from the manganese atoms that we presumed were bound to protein. Data at 2.9-Å resolution were collected in three continuous sweeps (1° increments) at three different crystal orientations to give a total of 933 images that were processed in HKL2000/SCALEPACK to give the highly redundant phasing dataset described in Table 1 (top right column). A second higher-resolution dataset (1.6 Å) used for refinement was collected on a second crystal at a wavelength of 1.1 Å. Data from two continuous sweeps comprising 185 1° increments and 200 0.5° increments were processed using XDS (23) to give the refinement dataset described in Table 1 (top left column). Both crystals belonged to space group P2$_1$ with a unit cell volume consistent with two CthPnkp-(171–424) protomers per asymmetric unit, assuming a solvent content of 40%. The crystals were isomorphous. The phasing dataset was truncated at a diffraction maximum of 2.9 Å owing to the crystal-to-detector distance used; the refinement dataset was complete to a diffraction limit of 1.6 Å.

Phases for CthPnkp-(171–424) crystals were obtained via PHENIX.AUTOSOL (24). Within AUTOSOL, the HYSS subroutine found two manganese sites; their coordinates were consistent with the translational non-crystallographic symmetry (NCS) observed via native Patterson analysis. Using phases from these initial two sites, the PHASER subroutine found four additional low-occupancy sites. Initial phases using all six sites were improved via density modification in the RESOLVE subroutine using NCS averaging. The initial model built by the AUTOBUILD subroutine contained 326 residues in two protomers and revealed the low-occupancy sites identified by PHASER to correspond to sulfur atoms of Cys191 and Met251 of each protomer. This initial model was then refined against the high-resolution dataset, extended further in AUTOBUILD and iteratively rebuilt by hand in COOT (25) and refined in PHENIX.REFINE. The final model had $R_{free}/R_{work}$ values of 20.8%/16.9%, with excellent geometry, no Ramachandran outliers and no large $\text{Fo}-\text{Fc}$ difference Fourier peaks (Table 1).
RESULTS AND DISCUSSION

Crystallization of the phosphatase domain of CthPnkp

The N-terminal segment of CthPnkp, spanning residues 1–425, is active as a polynucleotide kinase and a 2',3'-phosphoesterase (2). The central segment CthPnkp-(171–424) comprises an autonomous phosphatase domain (14). Here we grew crystals of CthPnkp-(171–424) by hanging-drop vapor diffusion after mixing a sample of the protein solution containing 0.8 mM phosphatase and 10 mM MnCl2 with an equal volume of precipitant solution containing 100 mM Na-citrate, pH 5.6, 33.5% PEG-3350, and 15% DMSO. The crystals diffracted well (up to 1.6 Å resolution) and belonged to space group P21. SAD phasing located two manganese atoms and four sulfur atoms in the asymmetric unit, which contained two phosphatase protomers. The refined model of the phosphatase at 1.6 Å resolution (Rwork/Rfree = 16.9%/20.8%; Table 1) consisted of a continuous polypeptide chain with excellent geometry and no Ramachandran outliers. A single Mn2+ and a citrate anion were present in the active site of each protomer. The A and B protomers superimposed with a root-mean-square deviation (RMSD) of 0.53 Å at 254 Cα positions. The phosphatase protomers are monomeric in the crystal.

Overview of the phosphatase structure

Two views of the tertiary structure of the phosphatase are depicted in Figure 2A; helices are colored cyan, and β strands are magenta. The secondary structure elements are shown above the amino acid sequence in Figure 2D. The phosphatase fold consists of a central β sandwich composed of a six-strand β sheet with topology β1↓-β7↓-β8↑-β9↑-β11↑-β10↓ (on the left side of the sandwich in Figure 1A) and a five-strand β sheet with topology β6↑-β5↑-β2↑-β12↓-β13↑ (on the right side of the sandwich). The folding topology is diagrammed in
Figure 2. Tertiary structure of the *CthPnkp* phosphatase domain and comparison with *λ* phosphatase. (A) The tertiary structure of protomer B of the *CthPnkp* phosphatase (*CthPase*) is shown in two orientations. The N and C termini and selected secondary structure elements are labeled in the left panel. Citrate (stick model) and Mn²⁺ (magenta sphere) are bound in the active site. (B) The structure of *λ* phosphatase (protomer C from pdb 1G5B) with sulfate and Mn²⁺ in the active site was superimposed on the *CthPase* structure (the view at right in panel A) and then offset horizontally. (C) The surface models of *CthPase* and vacuum electrostatics were generated in Pymol. The citrate in the active site is shown as a stick model. The middle view is looking down on the active site. The left/right views are rotated clockwise/counterclockwise to highlight the stalk-like projection of the Cth-loop from the enzyme surface flanking the active site. (D) The aligned primary structures of the *Cth* and *λ* phosphatases are shown, with amino acids colored according to secondary structure (blue for helices and magenta for β strands). Positions of side-chain identity/similarity are denoted by black circle above the sequence. Gaps in the alignment are indicated by dashes. The secondary structure elements of *Cth* phosphatase are shown above the amino acid sequence, with β strands rendered as arrows and α helices as cylinders. (E) Folding topology of the *CthPnkp* phosphatase.
Figure 2A. The β sheets are flanked by α helices (Figure 2A). Citrate and Mn$^{2+}$ are bound in a pocket at the base of a deep canyon formed by the 17-amino acid loop connecting α helices 6 and 7 (Figure 2A and C).

A DALI search (26) with the CthPnkp phosphatase identified numerous members of the binuclear metallophosphoesterase superfamily as homologs. Top DALI hits were a putative phosphatase from Trypanosoma brucei (pdb 2QJC; Z score, 23.8; RMSD of 2.6 Å over 207 C$\alpha$ positions), a putative diadenosine tetraphosphate hydrolyase from Shigella flexneri (pdb 2DFJ; Z score, 22.1; RMSD of 2.9 Å over 206 C$\alpha$ positions) and bacteriophage λ phosphatase (pdb 1G5B; Z score, 21.9; RMSD of 2.3 Å over 199 C$\alpha$ positions). The superimposed tertiary structures of CthPnkp and λ-Pase are shown side-by-side in Figure 2A (right panel) and B. Their primary and secondary structures are aligned in Figure 2D. The bacterial and phage enzymes have a common core fold, comprising the central β sandwich and several of the flanking α helices.

There are three noteworthy differences between the bacterial and phage phosphatases: (i) the presence in CthPnkp of a 17-amino acid β-hairpin module, composed of strands β3 and β4 and the intervening loop (Figure 2A), that has no counterpart in λ-Pase; (ii) the ‘Cth-loop’ and helix α7 that have no equivalent in λ-Pase (Figure 2A and B); and (iii) the presence in Pnkp of a C-terminal β13 strand at the outer edge of the five-strand β sheet that is absent in λ-Pase (Figure 2D and E).

The Cth-loop forms a distinctive stalk-like protrusion from the enzyme surface next to the active site (Figure 2C). The loop is rich in acidic side chains (Glu346, Glu350, Asp352, Glu353, Asp361) that impart a negative surface electrostatic potential (colored red in Figure 2C) to this part of the enzyme. The Cth-loop is present and conserved in length and primary structure in 10 other Pnkp enzymes from diverse taxa representing 10 different bacterial phyla (Figure 3B). A stereo view of the structure of the z6-loop–z7 segment is shown in Figure 3A, with conserved side chains and their interactions highlighted. The base of the loop is stabilized by (i) van der Waals contacts between residues Tyr341 and Trp362 and between Glu346 and Tyr360, (ii) a hydrogen bond from Tyr360 to Asp365 and between Glu346 and Tyr360, (ii) a hydrogen bond from Tyr360–N to Glu346–O and from Tyr341–O to Asp363–O, (ii) a hydrogen bond from Tyr360 to Val358–O, (ii) a hydrogen bond from Tyr360 to Glu346 and Tyr360, (ii) a hydrogen bond from Tyr360 to Asp352 and (iii) a triad of hydrogen bonds between the Asp352 carboxylate and the main-chain amide nitrogens of Tyr354, Gly355 and Leu365 (Figure 3A). We speculate that the Cth-loop may undergo a conformational switch on binding to an RNA substrate, whereby its concave surface that forms the canyon wall next to the active site closes over the poly-nucleotide to clamp it in place. Alternatively, the Cth-loop might facilitate relocation of the healed RNA 3’ end to other enzymatic active sites (e.g. methyltransferase or ligase) during RNA repair. The distinctive β3–β4 hairpin module is disposed on the inferolateral surface of the enzyme remote from the active site (Figure 2A). The inward-facing side chains of the β3 and β4 strands are hydrophobic (Ile206, Tyr218, Val220) and make van der Waals contacts with hydrophobic residues emanating from helices α1 and α2 (Tyr192, Leu199, Leu247, Met251). The outward-facing amino acids of the β3–β4 hairpin are mainly hydrophilic (Glu207, Glu210, Arg291, Thr221), consistent with them being solvent exposed in the context of the Pnkp-Hen1 complex.

The C-terminal β13 strand (highlighted in yellow in Figure 4A) and flanking amino acids are an integral element of the Pnkp phosphatase fold. In particular, residues Arg405, Tyr406, Glu410, Tyr420 are conserved among bacterial Pnkp proteins (Figure 4B) and engage in a wide network of hydrogen bonding and electrostatic interactions that tether the β12–β13 hairpin to other secondary structure elements of the phosphatase (Figure 4A).

The central panel in Figure 3C shows that citrate (a mimetic of the scissile phosphate) occupies a deep pocket at the canyon’s base. The negative electrostatic surface of the pocket reflects its function in binding the metal cations essential for catalysis. The citrate sits above the Mn$^{2+}$ atom in the active site. The enzyme surface to the right of the active site pocket is notable for its positive electrostatics (colored blue). The positive potential extends around the enzyme surface as the protein is rotated counterclockwise (the rightmost view in Figure 2C), suggestive of an RNA-binding track. We speculate that RNA end healing by the Pnkp phosphatase entails anchoring of the RNA 2’,3’ phosphate anion in the metal-bound pocket while the flanking 3’-terminal RNA segment wraps around the blue enzyme surface.

The phosphatase active site

The F$_o$–F$_c$ maps revealed electron density for citrate in the active site; the composite omit map for this citrate and nearby elements is shown in Figure 5A. The strong anomalous peak over the neighboring metal ion validated its assignment as manganese (Figure 5A). A detailed view of the active site architecture and the atomic contacts to citrate and Mn$^{2+}$ is shown in Figure 5B. Two of the sites in the Mn$^{2+}$ coordination complex are occupied by citrate oxygens. The other four sites are filled by His189-Ne, Asp187-Oδ2, Asp233-Oδ2 and a water coordinated by Asp392-Oδ2. The atomic contacts of Mn$^{2+}$ in the citrate-bound CthPnkp phosphatase are virtually identical to those of the corresponding Mn$^{2+}$ in λ-Pase (shown in Figure 5C). The λ-Pase structure captures the enzyme with a sulfate anion in the active site (21) in a state that mimics a product complex of the phosphomonoesterase reaction. The sulfate O1 atom coordinated between the two Mn$^{2+}$ ions in the λ-Pase structure corresponds to the metal-bound OH nucleophile in the Michaelis complex before phosphohydrolase chemistry (exemplified by the structure of Mycobacterium tuberculosis Rv0805; pdb 2HY1) (27).

The citrate C6 atom (colored yellow in Figure 5B) superimposes on the sulfate sulfur atom in λ-Pase (Figure 5C). Atomic contacts to the citrate O3, O5 and O7 atoms made by the Arg237 Asn263 and His264 side chains in CthPnkp hew to those made by λ-Pase residues Arg53, Asn75 and His76, respectively, with the sulfate
anion. Thus, we surmise that Arg237, Asn263 and His264 engage the scissile phosphate during catalysis by the CthPnkp phosphatase and, together with the metal(s), stabilize the transition state and promote expulsion of the ribose hydroxyl leaving group.

The salient feature of the CthPnkp active site vis-à-vis C21-Pase is that filling the phosphate binding site with citrate precludes occupancy of the second metal site. Specifically, the citrate C2, C1, O1 and O2 atoms encroach on the second Mn$^{2+}$ site (Figure 5A and B). This site in CthPnkp, which can be inferred with high confidence from the C21-Pase structure, would be formed by Asp233, His323, His376 and Asn263$^{\text{Od}}$. When citrate is bound, its O1 atom makes a hydrogen bond with His323$^{\text{Ne}}$ (Figure 5B). Also, the citrate C1, O1 and O2 carboxylate dislocates the His376-Thr377 dipeptide in the β9–β10 loop, displacing the His376-N$^{\text{δ}}$ atom 3.2 Å away from its position in the C21-Pase Mn$^{2+}$ coordination complex (see His376 in Figure 5B versus His186 in Figure 5C).

In effect, the citrate$^{\text{Mn}^{2+}}$ structure of CthPnkp exemplifies an inhibitor complex.

**Structural insights to mutational data**

Initial mutational analysis of the CthPnkp phosphatase by alanine scanning identified Asp187, His189, Asp233, Arg237, Asn263, His264, His323, His376 and Asp392 as essential for its Ni$^{2+}$-dependent ribonucleotide-3'-phosphomonoesterase and ribonucleotide-2'-phosphomonoesterase activities (14). Introducing conservative substitutions for these residues (e.g. Asn or Glu for Asp, Asn or Gln for His, Gln or Asp for Asn) failed to restore the ribonucleotide phosphomonoesterase activities, except for the replacement of His264 by glutamine (14). Subsequent studies focused on the Mn$^{2+}$-dependent 2',3'-cyclic phosphodiesterase activity of CthPnkp, which is the most pertinent to its 3' end-healing function during RNA repair (17). Mutations of Asp187 and Asp233 abolished 2',3'-cyclic phosphodiesterase activity (16), consistent with
Asp187 and Asp233 interactions with one or both of the catalytic Mn$^{2+}$ ions in the active site (Figure 5). Mutations of His264 reduced cyclic phosphodiesterase activity by an order of magnitude (16), highlighting the importance of its contacts to the scissile phosphate inferred from the citrate complex (Figure 5). Indeed, this phosphate-binding histidine was found to be a defining determinant of the ability of other binuclear metallophosphoesterases to act as 2',3'-cyclic nucleotide phosphodiesterases (27).

Mutations of active site components His189 and Arg237 exert remarkable effects on the cyclic phosphodiesterase activity of CthPnkp that are worth discussing in light of the crystal structure. Arg237 makes a bidentate contact with the citrate that presumably reflects its interaction with the scissile phosphate (even allowing for adoption of an alternative arginine rotamer as seen the C15-Pase/C21-SO$_4$ complex). R237A, R237K and R327Q mutations increased the $K_m$ for 2',3'-cAMP by 3- to 6-fold, thereby underscoring the importance of the phosphate contacts for substrate binding affinity. Yet, Arg237 was found to make little contribution to the maximal rate of 2',3'-cyclic phosphate hydrolysis, insofar as the alanine change had no effect on $k_{cat}$ and the R237K and R237Q changes actually increased $k_{cat}$ by 3-fold and 2-fold, respectively (16). The same trends were seen for the effects of R237A, R237K and R237Q on Mn$^{2+}$-dependent hydrolysis of the non-nucleotide phosphodiester substrate bis-p-nitrophenyl phosphate, i.e. these changes increased the $K_m$ for substrate by 2- to 3-fold while increasing $k_{cat}$ by 2-fold (R237K), 10-fold (R237A) or 20-fold (R237Q) (15). We suspect that weakening the arginine/PO$_4$ interactions increases the product off-rate, thereby explaining the observed increases in $k_{cat}$ for the mutant enzymes. The most impressive effect of the arginine mutations was to abolish the Mn$^{2+}$-dependent phosphomonoesterase activity of CthPnkp, thereby altering the reaction outcome by limiting it to a single step of phosphodiester hydrolysis instead of the sequential phosphodiesterase and monooesterase reactions performed by the wild-type CthPnkp (15). It is conceivable that the greater negative charge on the associative transition state of the phosphomonoester versus the diester demands neutralization by the arginine.

Mutating the Mn$^{2+}$ ligand His189 to aspartate also has dramatic effects on CthPnkp activity. The H189D change abolishes Ni$^{2+}$-dependent phosphomonoesterase and diesterase activities and the Mn$^{2+}$-dependent

Figure 4. Structural role of the distinctive $\beta$13 strand. (A) Stereo view highlighting the hydrogen bonding and electrostatic interactions of the C-terminal $\beta$13 strand (strand colored yellow) and flanking residues that tether the $\beta$12-$\beta$13 hairpin to other secondary structure elements of the phosphatase. (B) The primary structure of the C-terminal $\beta$10-$\beta$11-$\beta$12-$\beta$13 segment of CthPnkp phosphatase is aligned to the equivalent segments of Pnkp proteins from bacteria representing 10 phyla (see Figure 3 legend). Positions of side-chain identity/similarity in all of the aligned Pnkps are denoted by black circle above the sequence. The active site His376 residue is denoted by vertical line.

Nucleic Acids Research, 2013, Vol. 41, No. 11 5871
monooesterase activity, but spares the Mn\textsuperscript{2+}-dependent phosphodiesterase functions of CthPnkp (15). H189D has no effect on the K\textsubscript{m} for 2',3'-cAMP, but increases the k\textsubscript{cat} of the cyclic phosphodiesterase reaction by 4-fold (16). The H189D mutation fundamentally reprograms the tRNA repair pathway of CthPnkp, causing it to convert a broken RNA 2',3'-cyclic phosphate end to an RNA 3'-OH,2'-PO\textsubscript{4} product instead of the RNA 3'-OH,2'-OH product generated by the wild-type enzyme (17). Altering the outcome of end healing can redirect the subsequent end sealing steps of RNA repair toward ligases with different end specificities (17).

Concluding remarks

The present study provides new insights to the structure and mechanism of the phosphatase domain of bacterial Pnkp by capturing the enzyme in a complex with Mn\textsuperscript{2+} and citrate (a mimic of the scissile phosphate). The core fold and the amino acid constituents of the active site are similar to those of \( \lambda \)-Pase, the prototypal binuclear metallophosphoesterase. The core fold is embellished by distinctive Cth-specific loop and \( \beta \)-hairpin modules, with the Cth-loop being a good candidate to aid in RNA substrate recognition by virtue of its proximity to the active site. A continuous tract of positive surface potential flanking the CthPnkp active site suggests a role for this surface in RNA binding. It is noteworthy that \( \lambda \)-Pase lacks an analogous positive surface flanking its active site (not shown), which is in keeping with the idea that the physiological substrate for \( \lambda \)-Pase (although unknown) is probably not RNA.

The crystallized phosphatase domain of bacterial Pnkp is a monomer. In contrast, the crystallized kinase domain is a homodimer. The kinase–kinase homodimer interface can now be presumed to be the principal driver of the homodimeric quaternary structure of the isolated full-length CthPnkp protein (2) and of Pnkp in the context of the Pnkp-Hen1 RNA repair complex, which is a \((\text{Pnkp})\text{2}(\text{Hen1})\text{2}\) tetramer (5). The Hen1 methyltransferase domain is a monomer (6, 7), and the ligase-Hen1(N) complex is a 1:1 heterodimer (4). The parallel orientation of the kinase domain protomers in the kinase homodimer in the crystal structure, combined with aforementioned knowledge of the structures of all other component modules, supports a model for the \((\text{Pnkp})\text{2}(\text{Hen1})\text{2}\) tetramer (11) in which the central phosphatase domains reside on the same face of the kinase dimer and serve as linkers to pairs of LIG-Hen1 heterodimers. The next challenge will be to crystallize the complete Pnkp-Hen1 complex, the structure of which can be solved based on the now full suite of high-resolution crystal structures of the component domains.

To our knowledge, the CthPnkp phosphatase is the first example of a structurally defined binuclear metallophosphoesterase dedicated to RNA repair or modification. Other RNA end-healing enzymes belong to unrelated enzyme superfamilies (discussed in 28). To our knowledge, the only other binuclear metallophosphoesterase dedicated to RNA modification is the intron-debranching enzyme Dbr1 that hydrolyzes the 2'-5'-phosphodiester linkage of the excised lariat intron product of eukaryal pre-mRNA splicing. Dbr1, like CthPnkp, is a Mn\textsuperscript{2+}-dependent phosphodiesterase, and mutational analysis of Dbr1 highlights the essentiality of many of the conserved side chains that comprise the active sites of CthPnkp and \( \lambda \)-Pase (22).
REFERENCES

1. Popow, J., Schleifer, A. and Martinez, J. (2012) Diversity and roles of tRNA ligases. Cell Mol. Life Sci., 69, 2657–2670.
2. Martins, A. and Shuman, S. (2005) An end-healing enzyme from Clostridium thermocellum with 5′ kinase, 2′,3′ phosphatase, and adenylytransferase activities. RNA, 11, 1271–1280.
3. Smith, P., Wang, L.K., Nair, P.A. and Shuman, S. (2012) The adenylytransferase domain of bacterial Pnkp defines a unique RNA ligase family. Proc. Natl Acad. Sci. USA, 109, 2296–2301.
4. Wang, P., Chan, C.M., Christensen, D., Zhang, C., Selvadurai, K. and Huang, R.H. (2012) Molecular basis of bacterial protein Hen1 activating the ligase activity of bacterial protein Pnkp for RNA repair. Proc. Natl Acad. Sci. USA, 109, 13248–13253.
5. Chan, C.M., Zhou, C. and Huang, R. (2009) Reconstituting bacterial RNA repair and modification in vitro. Science, 326, 247.
6. Chan, C.M., Zhou, C., Brunzelle, J.S. and Huang, R.H. (2009) Structural and biochemical insights into 2′-O-methylation at the 3′-terminal nucleotide of RNA by Hen1. Proc. Natl Acad. Sci. USA, 106, 17699–17704.
7. Jain, R. and Shuman, S. (2010) Bacterial Hen1 is a 3′ terminal RNA ribose 2′-O-methyltransferase component of a bacterial RNA repair cassette. RNA, 16, 316–325.
8. Jain, R. and Shuman, S. (2011) Active site mapping and substrate specificity of bacterial Hen1, a manganese-dependent 3′ terminal RNA ribose 2′-O-methyltransferase. RNA, 17, 429–438.
9. Nandakumar, J., Schwer, B., Schaffrath, R. and Shuman, S. (2008) RNA repair: an antidote to cytotoxic eukaryal RNA damage. Mol. Cell, 31, 278–286.
10. Huang, R.H. (2012) Unique 2′-O-methylation by Hen1 in eukaryotic RNA interference and bacterial RNA repair. Biochemistry, 51, 4087–4095.
11. Wang, L.K., Das, U., Smith, P. and Shuman, S. (2012) Structure and mechanism of the polynucleotide kinase component of the bacterial Pnkp-Hen1 RNA repair system. RNA, 18, 2277–2286.
12. Wang, L.K., Lima, C.D. and Shuman, S. (2002) Structure and mechanism of T4 polynucleotide kinase – an RNA repair enzyme. EMBO J., 21, 3873–3880.
13. Galbur, E.A., Pelletier, J., Wilson, G. and Stoddard, B.L. (2002) Structure of a tRNA repair enzyme and molecular biology Workhorse: T4 polynucleotide kinase. Structure, 10, 1249–1260.
14. Keppetipola, N. and Shuman, S. (2006) Mechanism of the phosphatase component of Clostridium thermocellum polynucleotide kinase-phosphatase. RNA, 12, 73–82.
15. Keppetipola, N. and Shuman, S. (2006) Distinct enzymic functional groups are required for the phosphomonoesterase and phosphodiesterase activities of Clostridium thermocellum polynucleotide kinase-phosphatase. J. Biol. Chem., 281, 19251–19259.
16. Keppetipola, N. and Shuman, S. (2007) Characterization of the 2′,3′ cyclic phosphodiesterase activities of Clostridium thermocellum polynucleotide kinase-phosphatase and bacteriophage λ phosphatase. Nucleic Acids Res., 35, 7721–7732.
17. Keppetipola, N., Nandakumar, J. and Shuman, S. (2007) Reprogramming the tRNA splicing activity of a bacterial RNA repair enzyme. Nucleic Acids Res., 35, 3624–3630.
18. Cohen, P.T. and Cohen, P. (1989) Discovery of a protein phosphatase activity encoded in the genome of bacteriophage λ. Biochem. J., 260, 931–934.
19. Barik, S. (1993) Expression and biochemical properties of a protein serine/threonine phosphatase encoded by bacteriophage λ. Proc. Natl Acad. Sci. USA, 90, 10633–10637.
20. Zhuo, S., Clemens, J.C., Hakes, D.J., Barford, D. and Dixon, J.E. (1993) Expression, purification, crystallization, and biochemical characterization of a recombinant protein phosphatase. J. Biol. Chem., 268, 17754–17761.
21. Voigtli, W.C., White, D.J., Reiter, N.J., Rusan, F. and Rosenzweig, A.C. (2000) Structure of the bacteriophage λ Ser/Thr protein phosphatase with sulfate ion bound in two coordination modes. Biochemistry, 39, 15365–15374.
22. Khalid, F., Damha, M., Shuman, S. and Schwer, B. (2005) Structure-function analysis of yeast RNA debranching enzyme (Dbr1), a manganese-dependent phosphodiesterase. Nucleic Acids Res., 33, 6349–6360.
23. Kabsch, W. (2010) XDS. Acta Crystallogr., D66, 125–132.
24. Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Saccotelli, I.C., Sauter, N.K. and Terwilliger, T.C. (2002) PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr., D58, 1948–1954.
25. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr., D60, 2126–2132.
26. Holm, L., Kaariainen, S., Rosenström, P. and Schenkel, A. (2008) Searching protein structure databases with Dalilite v.3. Bioinformatics, 24, 1780–1781.
27. Keppetipola, N. and Shuman, S. (2008) A phosphate-binding histidine of binuclear metallophosphodiesterase enzymes is a determinant of 2′,3′ cyclic nucleotide phosphodiesterase activity. J. Biol. Chem., 283, 30942–30949.
28. Das, U. and Shuman, S. (2013) Mechanism of RNA 2′,3′-cyclic phosphate end-healing by T4 polynucleotide kinase-phosphatase. Nucleic Acids Res., 41, 355–365.