Molecular Cell, Vol. 17, 657–670, March 4, 2005, Copyright ©2005 by Elsevier Inc. DOI 10.1016/j.molcel.2005.02.012

The Molecular Architecture of the Mammalian DNA Repair Enzyme, Polynucleotide Kinase

Nina K. Bernstein,[1] R. Scott Williams,[1] Melissa L. Rakovszky,[1] Diana Cui,[1] Ruth Green,[1] Feridoun Karimi-Busheri,[2] Rajam S. Mani,[2] Sarah Galicia,[3] C. Anne Koch,[4] Carol E. Cass,[2] Daniel Durocher,[3] Michael Weinfeld,[2] and J.N. Mark Glover[1,*]

[1]Department of Biochemistry
4–74 Medical Sciences Building
University of Alberta
Edmonton, Alberta T6G 2H7
Canada

[2]Experimental Oncology, Cross Cancer Institute, and Department of Oncology
University of Alberta
Edmonton, Alberta T6G 1Z2
Canada

[3]Samuel Lunenfeld Research Institute
Mount Sinai Hospital
600 University Avenue
Toronto, Ontario M5G 1X5
Canada

[4]Department of Radiation Oncology
Princess Margaret Hospital
610 University Avenue
Toronto, Ontario M5G 2M9
Canada

Summary

Mammalian polynucleotide kinase (PNK) is a key component of both the base excision repair (BER) and nonhomologous end-joining (NHEJ) DNA repair pathways. PNK acts as a 5′-kinase/3′-phosphatase to create 5′-phosphate/3′-hydroxyl termini, which are a necessary prerequisite for ligation during repair. PNK is recruited to repair complexes through interactions between its N-terminal FHA domain and phosphorylated components of either pathway. Here, we describe the crystal structure of intact mammalian PNK and a structure of the PNK FHA bound to a cognate phosphopeptide. The kinase domain has a broad substrate binding pocket, which preferentially recognizes double-stranded substrates with recessed 5′ termini. In contrast, the phosphatase domain efficiently dephosphorylates single-stranded 3′-phospho termini as well as double-stranded substrates. The FHA domain is linked to the kinase/phosphatase catalytic domain by a flexible tether, and it exhibits a mode of target selection based on electrostatic complementarity between the binding surface and the phosphothreonine peptide.

Introduction

Scission of the DNA backbone is a common form of damage that can arise not only as a direct result of an initial genotoxic event, but also as an intermediate product in several DNA repair pathways. As a result, all organisms have evolved highly specialized DNA repair polymerases and ligases to efficiently process strand breaks. All DNA polymerases and ligases characterized to date are highly selective for the type of DNA ends that can be utilized; all require 5′-phosphate and 3′-hydroxyl DNA termini. This requirement, however, creates further complications for repair, as many DNA-damaging agents and upstream repair processes often generate DNA ends that are incompatible with the polymerases and ligases. DNA backbone breaks with 5′-hydroxyl termini can result from ionizing radiation, DNase II action, and antineoplastic agents such as camptothecin. Sources of 3′-phosphate at breaks include ionizing radiation, DNase II action, and Tdp1 cleavage of camptothecin-trapped topoisomerase I DNA adducts (Wang, 1996) or 3′-phosphoglycolates at some double-strand breaks (Zhou et al., 2005). Furthermore, 3′-phosphate termini are transiently generated during repair via the base excision repair (BER) pathway, the primary mechanism by which small base lesions and single-strand breaks are repaired (Izumi et al., 2003; Wiederhold et al., 2004). Mammalian polynucleotide kinase (PNK) (Pheiffer and Zimmerman, 1982; Habraken and Verly, 1983) possesses dual catalytic activities, a 5′ DNA kinase and a 3′ phosphatase, and is the principal enzyme responsible for restoring 5′-phosphate and 3′-hydroxyl at DNA-strand breaks.

Although mammalian PNK has long been suspected to play a key role in DNA repair (Teraoka et al., 1975), only recently have conclusive experiments been performed to implicate this enzyme in the repair of both single-strand breaks (SSBs) and double-strand breaks (DSBs). PNK has been shown to stimulate SSB repair in both in vitro reconstitution experiments (Karimi-Busheri et al., 1998; Whitehouse et al., 2001; Wiederhold et al., 2004) and in vivo studies (Loizou et al., 2004). PNK has also been shown to be an integral component of the nonhomologous end-joining (NHEJ) pathway, which is the major route for the repair of double-strand breaks in mammalian cells (Chappell et al., 2002; Koch et al., 2004). Consistent with a central role for PNK in both of these processes, knockdown of endogenous PNK in a human cell line leads to an increase in spontaneous mutations and enhanced sensitivity to a broad range of genotoxic agents (Rasouli-Nia et al., 2004). In *S. pombe*, deletion of Pnk1, a homolog of PNK, produces similar effects (Meijer et al., 2002). Not only does PNK play a role in the repair of strand breaks, it functions in the repair of base lesions as part of the BER pathway. The recently discovered DNA glycosylases/AP lyases NEIL1 and NEIL2 initiate AP endonuclease-independent BER and interact with PNK in a multiprotein BER complex (Hazra et al., 2002; Rosenquist et al., 2003; Wiederhold et al., 2004). These enzymes excise abasic (AP) sites leaving breaks with 3′-phosphate termini. Although AP-endonuclease was previously thought to process 3′-phosphates, it has now been shown to do so much

*Correspondence: mark.glover@ualberta.ca

more slowly than PNK, leaving PNK as the major cellular 3′ phosphatase (Wiederhold et al., 2004).

Many DNA repair enzymes appear to be actively recruited to specific repair complexes to ensure that potentially mutagenic intermediates along the repair pathway are not released prematurely (Mitra et al., 2002). PNK contains, in addition to its catalytic kinase and phosphatase domains, an N-terminal FHA domain (Caldecott, 2003), which has recently been shown to target the enzyme to sites of BER and NHEJ repair. XRCC1 and XRCC4, central components of the BER and NHEJ pathways, respectively, are phosphorylated at threonine residues by CK2 (Koch et al., 2004; Loizou et al., 2004). The FHA domain of PNK specifically recognizes these phosphorylated forms of XRCC1 and XRCC4, directing PNK to the site of repair. The PNK FHA exhibits distinct preferences for peptide sequences N-terminal to the phosphothreonine. In contrast, several other well-characterized FHA domains show specificity for the sequences C-terminal to the phosphothreonine. The PNK FHA is highly similar to the FHA domain of aprataxin (APTX), a protein associated with the neurological disorder ataxia-oculomotor apraxia and likely also involved in DNA repair (Moreira et al., 2001; Gueven et al., 2004). Indeed, aprataxin is found in vivo in complexes with XRCC1/Lig III and XRCC4/Lig IV, in parallel with but exclusive of PNK. A role for aprataxin in regulating PNK activity has therefore been proposed (Luo et al., 2004; Clements et al., 2004).

Another well-known bifunctional polynucleotide kinase/phosphatase, T4 PNK, has been extensively characterized structurally and functionally. In addition to a battery of mutagenesis and activity experiments, crystal structures have been determined for the kinase domain of T4 PNK, as well as for the full-length enzyme, with and without bound oligonucleotide substrates (Galburt et al., 2002; Wang et al., 2002a; Eastberg et al., 2004). Aside from the similar catalytic activities, T4 and mammalian PNK display profound differences. T4 PNK is a tetramer in vivo, whereas mammalian PNK is a monomer (Midgley and Murray, 1985; Mani et al., 2001). The T4 kinase is quite promiscuous in its substrate choice, accepting RNA, DNA, as well as mono and oligonucleotides (Kleppe and Lillehaug, 1979). The mammalian kinase is specific for DNA and requires a minimal substrate length of 8 nucleotides (Karimi-Busheri and Weinfeld, 1997). Furthermore, while the T4 kinase clearly prefers protruding 5′ ends, the mammalian kinase shows a preference for recessed DNA ends (Karimi-Busheri and Weinfeld, 1997). Finally, T4 PNK lacks the FHA domain. These contrasts are consistent with these enzymes' different biological roles. T4 PNK does not play a role in DNA repair, but instead participates in the repair of a specifically cleaved bacterial lysine tRNA as part of a phage-encoded system to evade the host cell defense (Amitsur et al., 1987). T4 and mammalian PNKs share virtually no sequence similarity, aside from the core elements of the kinase and phosphatase active sites, the P loop and the DxDGT motifs, respectively. Moreover, the kinase and phosphatase are reversed in sequence in the two enzymes.

We have determined the crystal structure of full-length mouse PNK (mPNK). In agreement with domain-mapping experiments and sequence analysis, the protein comprises three domains: kinase, phosphatase and FHA. We have refined the definition of substrate preferences for the kinase and phosphatase using activity and binding assays. Based on the biochemical data, our structures, and structures of related proteins, we have proposed the modes of substrate binding to both the kinase and phosphatase domains of mPNK. Furthermore, to characterize the role of PNK in multiprotein assemblies, we have determined the crystal structure of the mouse PNK FHA domain in complex with the XRCC4-derived phosphopeptide. While the mode of phosphothreonine binding is highly conserved in PNK relative to the other FHA domains of known structure, the elements of peptide recognition differ significantly. This structure suggests that the PNK FHA has an adaptable peptide binding specificity that allows it to recognize highly negatively charged CK2-phosphorylated targets within XRCC1 and XRCC4.

Results and Discussion

Structural Overview of PNK

The crystal structure of full-length mouse PNK was determined by Se-MAD methods and refined at a resolution of 2.8 Å (Table 1). The structure reveals that PNK is folded into three compact domains, an N-terminal FHA domain (Ser 6–Ser 110), the phosphatase domain (Gly 145–Glu 336), and the kinase domain (Phe 340–Glu 521) (Figures 1A and 1B). The kinase and phosphatase together form the catalytic domain. We observe two NCS-related catalytic domains per asymmetric unit, but only a single FHA domain with weaker electron density. The second FHA domain is not visible in the electron density and is presumably completely disordered. Since the 30-residue linker between the FHA domain and the catalytic domain is undefined in the structure, we cannot assign the observed FHA to either of the NCS-related catalytic domains. The interface between the FHA domain and the closest phosphatase domain, which buries only a small region of polar surface (less than 850 $Å^2$), likely results from crystal packing and probably does not represent interactions that are stable in solution. The flexible organization of the FHA and catalytic domains relative to one another is consistent with results of limited proteolysis experiments (Figure 1C). Trypsin digestion rapidly converts the full-length PNK to a pair of fragments: the catalytic domain (40 kDa) and the FHA domain (14 kDa). SDS-PAGE analysis of PNK crystals confirmed that they contained full-length PNK and not proteolyzed fragments (Figure 1C). The readily cleavable linker between the FHA and catalytic domains accounts for the frequent recovery of PNK as a 40 kDa protein from tissue extracts, (Bosdal and Lillehaug, 1985; Karimi-Busheri and Weinfeld, 1997). The flexibility of the FHA domain relative to the catalytic domain supports the idea that FHA has no direct role in catalysis by PNK, and instead serves as a targeting module.

In contrast, the kinase and phosphatase domains contact one another through a large interface involving the interdomain linker (Leu 337–Ala 339) and the C-terminal residues Gln 518–Gly 522, which pack against helix 5 and the β12-α4 loop of the phosphatase (Figure 1A).

Table 1. Crystallographic Data Collection, Phasing, and Refinement Statistics

Data Collection		mPNK		FHA
Space group		$P2_12_12$		$C222_1$
Cell dimensions	a (Å)	95.34		51.16
	b (Å)	169.51		123.00
	c (Å)	76.89		136.71
	λ1 (Remote)	λ2 (Peak)	λ3 (Inflection)	
Wavelength (Å)	1.0199	0.9793	0.9797	1.1159
Resolution range (Å)	30–2.8	30–2.9	30–3.1	50–2.2
Observations	113165	101608	84819	51266
Unique reflections	31529	28314	23521	18994
Data coverage total/final shell[a] (%)	99.0 (91.2)	99.9 (99.9)	99.9 (100)	89.4 (70.2)
<I/σI> total/final shell	13.1 (2.4)	13.3 (3.8)	10.9 (3.2)	8.8 (3.4)
R_{sym} total/final shell (%)[b]	9.5 (44.8)	9.5 (32.7)	11.5 (37.6)	5.1 (19.4)

Phasing Statistics	mPNK
Resolution range (Å)	30.0–3.1
No. of selenium sites	8/12
FOM – solve	0.3
FOM – resolve	0.61

Refinement Statistics		mPNK		FHA
Resolution range (Å)		30–2.8		50–2.2
R_{work}/R_{free} (%)[c]		0.218/0.252		0.211/0.243
No. of refined atoms	Protein	6690		2350
	Water	78		188
	Ligands (SO_4^{2-})	60	phosphopeptide	181
Rms deviations	Bonds (Å)	0.005		0.005
	Angles (°)	0.919		0.862
Average B factors (Å^2)	Protein	26.0		27.3
	Water	27.7		28.8
	Ligands (SO_4^{2-})	65.8	phosphopeptide	42.8
Ramachandran	Most favored	678 (93.3%)		235 (92.9%)
	Allowed	49 (6.7%)		18 (7.1%)
	Generously allowed	0		0
	Disallowed	0		0

Data from wavelength λ1 were used during crystallographic refinement. R_{free} calculated with 5% of all reflections excluded from refinement stages using the native data set. No I/σI cutoff was used in the refinement.
[a] Final shell: λ1: 2.90 – 2.80 Å, λ2: 3.00 – 3.10 Å, λ3: 3.21 – 3.10 Å, FHA: 2.24 – 2.20.
[b] $R_{sym} = \Sigma\,|(I_{hkl}) - \langle I\rangle|/\Sigma(I_{hkl})$ where I_{hkl} is the integrated intensity of a given reflection.
[c] $R_{work} = \Sigma_h|F_o(h) - F_c(h)|/\Sigma_h|F_o(h)|$, where $F_o(h)$ and $F_c(h)$ are observed and calculated structure factors.

The intimate association explains why the catalytic domains could not be separated by proteolysis (Figure 1C). Despite the tight association of the kinase and phosphatase domains, the two PNK molecules in the crystallographic asymmetric unit display a moderate degree of flexibility (10° rotation) in domain orientation. While the flexibility does not perturb either active site, it may enable PNK to adjust to different DNA-strand break structures.

Although the kinase and phosphatase domains are reversed in order in the sequence of mammalian and T4 PNK, in the three-dimensional structure their spatial orientation is similar. The phosphatase is located on the same side of the kinase in both molecules, but a significant shift between the domains alters the relative orientation of the kinase and phosphatase active sites. In T4 PNK the active sites of one monomer point in opposite directions, whereas in mammalian PNK, the two active sites face approximately the same direction. However, the tetrameric organization of T4 PNK stabilizes the relative orientation of the kinase and phosphatase and allows the substrate a choice between several combinations of active sites. Thus, a tRNA substrate was proposed to interact simultaneously with kinase and phosphatase sites facing each other across the tetramer interface, contributed by two monomers (Galburt et al., 2002). Mammalian PNK, on the other hand, is a monomer, where only the double interdomain linkage stabilizes the domain orientation. Because in mPNK there is only a single kinase and phosphatase active site, it is convenient for them to face in the same direction, presumably from where all substrates approach.

The Kinase Domain

Substrate Specificity

Early experiments on PNK revealed major differences in substrate preference between the mammalian kinase and the T4 kinase, consistent with the known biological functions of the two enzymes (Karimi-Busheri et al., 1998). T4 PNK was shown to prefer 5′-overhanging substrates over either blunt or 5′-recessed termini, whereas mammalian PNK was shown to preferentially phosphorylate nicks and small gaps compared to sin-

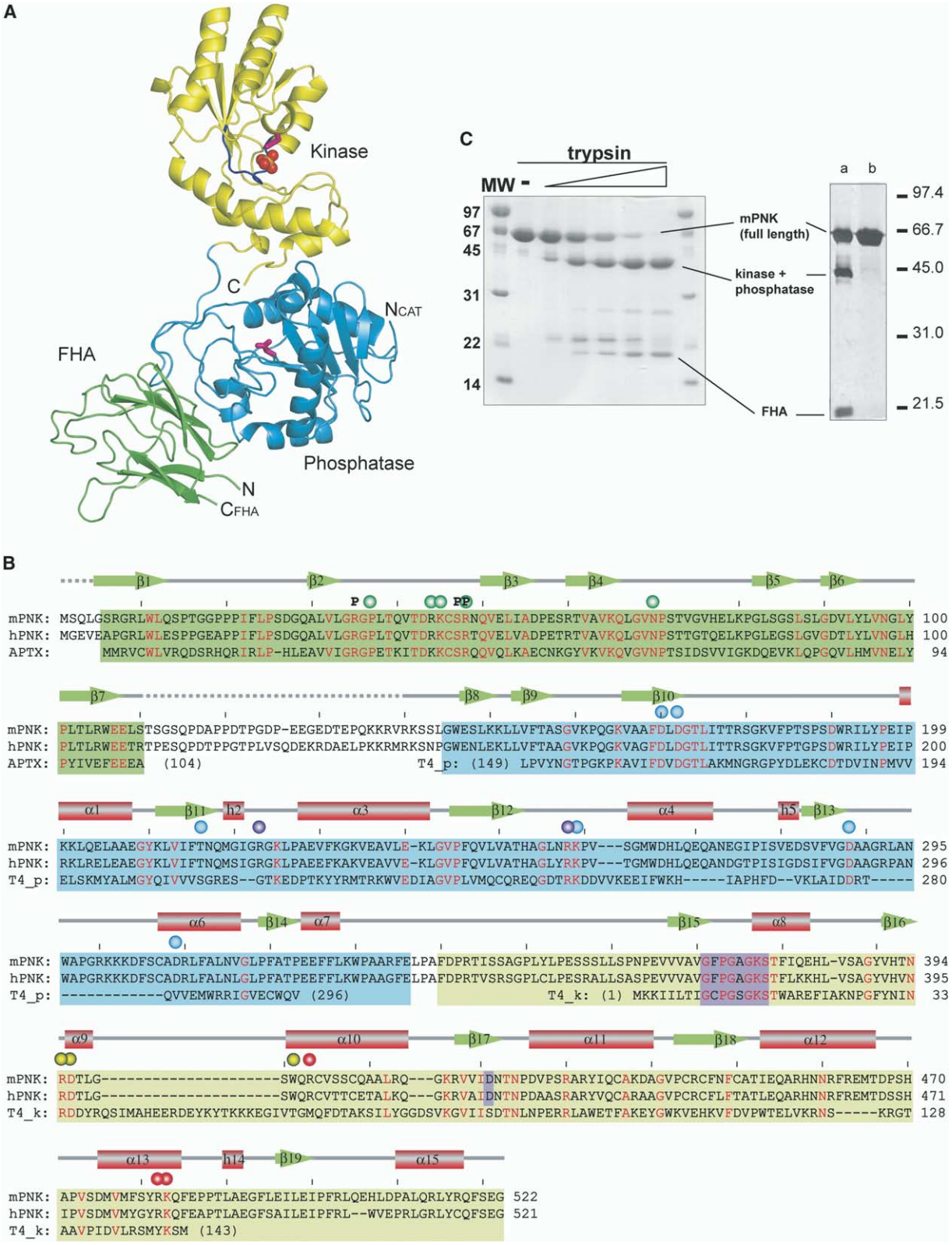

Figure 1. Structure of Mouse PNK

(A) Ribbon diagram of mPNK, with kinase in yellow, phosphatase in blue and FHA domain in green. Catalytic side chains (Asp 170 and Asp 396 in the phosphatase and kinase, respectively) are in pink, the ATP binding P loop is in navy blue, and the sulfate bound at the P loop is in orange and red spheres.

gle-stranded substrates (Lillehaug et al., 1976; Karimi-Busheri and Weinfeld, 1997).

To further define the kinase substrate preferences of PNK, we assessed the ability of mPNK to phosphorylate 5′-hydroxyl groups in a series of DNA substrates with two possible sites of phosphorylation, a 5′-hydroxyl at a blunt, double-stranded end, and a second 5′-hydroxyl at an internal site, corresponding to a gap or a nick (Figure 2A). At low concentration, mPNK exhibited a clear preference for the 5′ recessed hydroxyl over the blunt-ended substrate (Figure 2B), as well as over single-stranded DNA (data not shown). However, as the enzyme concentration increased, the selectivity for the recessed 5′-hydroxyl was lost. We next explored the optimal structure around the 5′ recessed hydroxyl, varying the size of the gap in our model substrate (Figure 2C). Our results show that varying the gap from 0 to 10 nucleotides (Figure 2, substrates a–e) did not have any noticeable effect on the amount of phosphorylation at the internal 5′-hydroxyl. In addition, two-stranded substrates with a 3′ single-stranded overhang (Figure 2, substrates f–h) exhibited the same extent of phosphorylation as matched nicked or gapped substrates, suggesting that the optimal kinase substrate is a recessed 5′-hydroxyl. Finally, we examined the minimal 3′ overhang required by mPNK (Figure 2D). A set of DNA duplexes with 3′ overhangs of 0, 3, 5, and 8 nucleotides (Figure 2, substrates i–m) was tested for 5′ phosphorylation. We found that selectivity for the recessed 5′-hydroxyl appeared with 3′ overhangs of between 3 and 5 nucleotides.

To directly probe the affinity of mPNK for DNA substrates, we utilized an assay which monitors the change of intrinsic, mPNK tryptophan fluorescence upon binding to DNA substrates (Mani et al., 2003). Binding affinities were measured for the duplex substrate containing a 5 nucleotide 3′ overhang, as well as for the constituent 13- or 8-mer oligonucleotides. The mPNK-DNA interaction was accompanied by a partial quenching of fluorescence with no change in the emission maximum. Binding affinity (K_d) and stoichiometry were determined from best fits to plots of fluorescence quenching as a function of DNA concentration (Figure 2E). In each case, the data indicate that mPNK binds its DNA substrate at a single site with 1:1 stoichiometry. The overhanging duplex DNA was bound with an apparent K_d of 0.36 ± 0.04 μM, while both the 13- and 8-mer DNAs were bound with significantly higher K_ds of 1.0 ± 0.1 and 1.4 ± 0.1 μM, respectively.

Taken together, these experiments demonstrate that mammalian PNK selectively binds and phosphorylates 5′-hydroxyl termini recessed by more than 3 nucleotides. The length of the gap following the 5′-hydroxyl appears to be immaterial. Indeed, the overhanging DNA may be entirely single-stranded with no decrease in recessed 5′-hydroxyl phosphorylation.

Structure of the Kinase Domain

The 189-residue kinase domain of mPNK is larger than its similarly folded counterpart in T4 PNK, which contains only 148 residues. This domain, which belongs to the adenylate kinase family, consists of a 5-stranded parallel β sheet, common to GTPases and P loop kinases (Leipe et al., 2003), flanked by helices on both sides. In addition, three helices (α12, 13, and 15) lie between the α/β sandwich and the phosphatase domain. The T4 kinase is smaller, containing only a 4-stranded parallel sheet with comparable topology.

The kinase active site is located in a long cleft, with the ATP binding site at one end and the substrate binding site at the other (Figure 3A). While we have not yet been able to crystallize a nucleotide-bound form of mPNK, we can define the ATP binding site by the conserved Walker A and B motifs present in various P loop kinases (Leipe et al., 2003). The Walker A motif, the P loop (Gly 371–Ser 378), binds the β- and γ-phosphates of ATP and adopts a conformation essentially identical to that found in T4 PNK. In the Walker B motif, Asp 421 forms a hydrogen bond with Ser 378 for proper positioning of the two motifs and coordinates Mg^{2+}, which is not observed in our structure but is required for kinase activity. This motif is generally less conserved among the P loop NTPases, and is replaced by a serine in T4 PNK. The ATP binding site is completed by the lid subdomain (helices 12 and 13), which folds over the P loop. In our structure of mPNK a sulfate anion is located at the position corresponding to the ADP β-phosphate in the T4 structure (Galburt et al., 2002). The contacts made by the sulfate and the ADP β-phosphate with the P loop are the same in both structures, but the lid differs slightly in orientation. Sequence and structural analyses of lid subdomains of P loop kinases identified two conserved arginines (Rx(2-3)R) in the first lid helix (α12 in mPNK), which interact with the bound ATP, stacking with the adenine and forming salt bridges with β- and γ- phosphates (Leipe et al., 2003). In T4 PNK Arg 122 and Arg 126 perform these roles. In mPNK, however, the corresponding residues in sequence, Arg 457 and Arg 461, are on the opposite face of helix 12, pointing away from the nucleotide binding site. Arg 463 from helix 12 forms hydrogen bonds with the bound sulfate (Figure 3A). When ADP from the T4 PNK structure is superimposed, however, the side chain of Arg 463

(B) Sequence comparison of mouse PNK (mPNK) with human PNK (hPNK). The FHA, phosphatase and kinase domains are highlighted in green, blue and yellow, respectively. In addition, the FHA domain is aligned with the human aprataxin FHA domain (APTX), and the phosphatase and kinase domains are structurally aligned with T4 PNK phosphatase and kinase (T4p and T4k). mPNK secondary structure is shown above the sequence with every tenth residue of mPNK marked (|). Helices in the secondary structure are labeled sequentially, with α helices indicated by 'α' and 3_{10} helices by 'h'. In the FHA domain phosphothreonine- and peptide binding residues are marked with "P" and green spheres, respectively. In the phosphatase, residues conserved at the active site are labeled with blue spheres, and residues predicted to bind the substrate backbone, with purple spheres. In the kinase, ATP binding motifs (Walker A and B) are highlighted in purple. Yellow spheres mark the catalytic Asp and residues predicted to bind the substrate backbone phosphate, and red spheres mark positive surface patches predicted to interact with distal backbone phosphates.

(C) On the left, limited trypsin proteolysis of mPNK. On the right, SDS-PAGE gel of mPNK crystal (lane b) *vs.* a partially degraded mPNK in solution (lane a). All figures showing structure were generated with PyMOL (http://www.pymol.org).

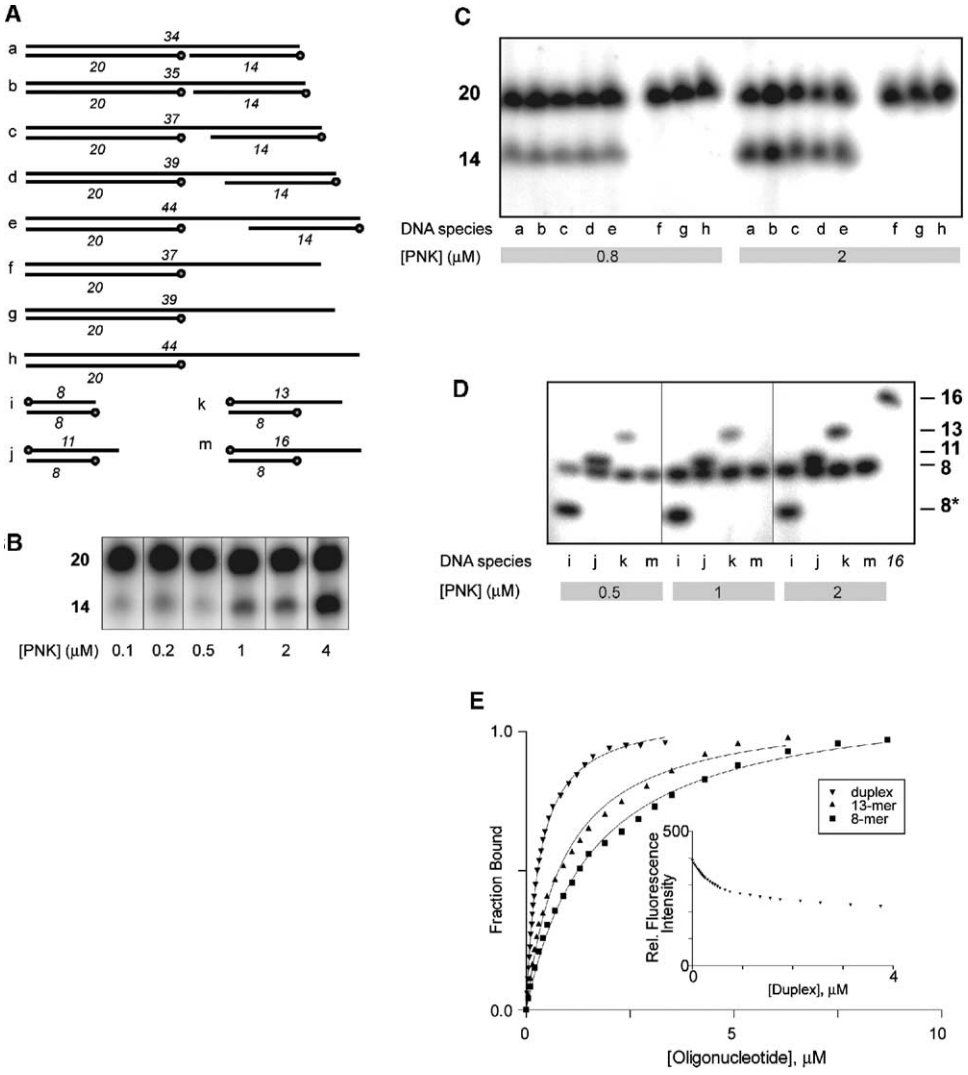

Figure 2. Elucidation of Kinase Substrate Specificity

(A) Substrates tested in kinase activity and binding assays. Circles at termini represent 5′-OH, and numbers above or below the individual strands represent oligonucleotide length. (See Supplemental Data for sequences.)

(B) Phosphorylation of a fixed concentration (5 μM) of substrate (b) by increasing concentrations of mPNK.

(C) Phosphorylation of model substrates with variable gap size (a–e) and the corresponding 5′-recessed OH (f–h) at 0.5 μM by two concentrations of mPNK.

(D) Determination of minimal length of 3′ overhang. Substrates (i–m) are phosphorylated by increasing concentrations of mPNK. Strand 8* runs faster on the denaturing gel due to sequence differences.

(E) Fluorescence titration of mPNK with substrate (k) (duplex) or its constituent oligonucleotides (13-mer and 8-mer). The inset figure shows a sample raw data plot for the duplex.

clashes with the ribose ring, and would need to be rotated upon ATP binding. Alternately, ATP may adopt different conformations in mammalian and T4 PNK or the lid may refold upon ATP binding. CD studies of human PNK showed a significant decrease in α-helical structure upon ATP binding (Mani et al., 2003).

The substrate binding end of the active site cleft also shows gross differences between mammalian and T4 PNK, not surprising in light of their divergent substrate preferences. The substrate binding site is more open in mPNK than in T4 PNK. In the phage enzyme an insertion (Gln 39–Ile 59) folded into a longer loop after helix 2 (α9 in mPNK) and a longer helix 3 (α10 in mPNK), extends over the substrate binding site, enclosing the active site in a tunnel narrow enough to admit only single-stranded DNA (Figure 3B). In mPNK the active site is more accessible to solvent and to substrate, which is generally larger than the optimal T4 kinase substrate.

A Model for the Kinase/Substrate Complex

Based on our kinase assay experiments and previous work (Karimi-Busheri and Weinfeld, 1997), we can define a minimal preferred kinase substrate as a DNA duplex with 8 base pairs upstream of the 5′-hydroxyl that is recessed by 4–5 nucleotides. In an attempt to better

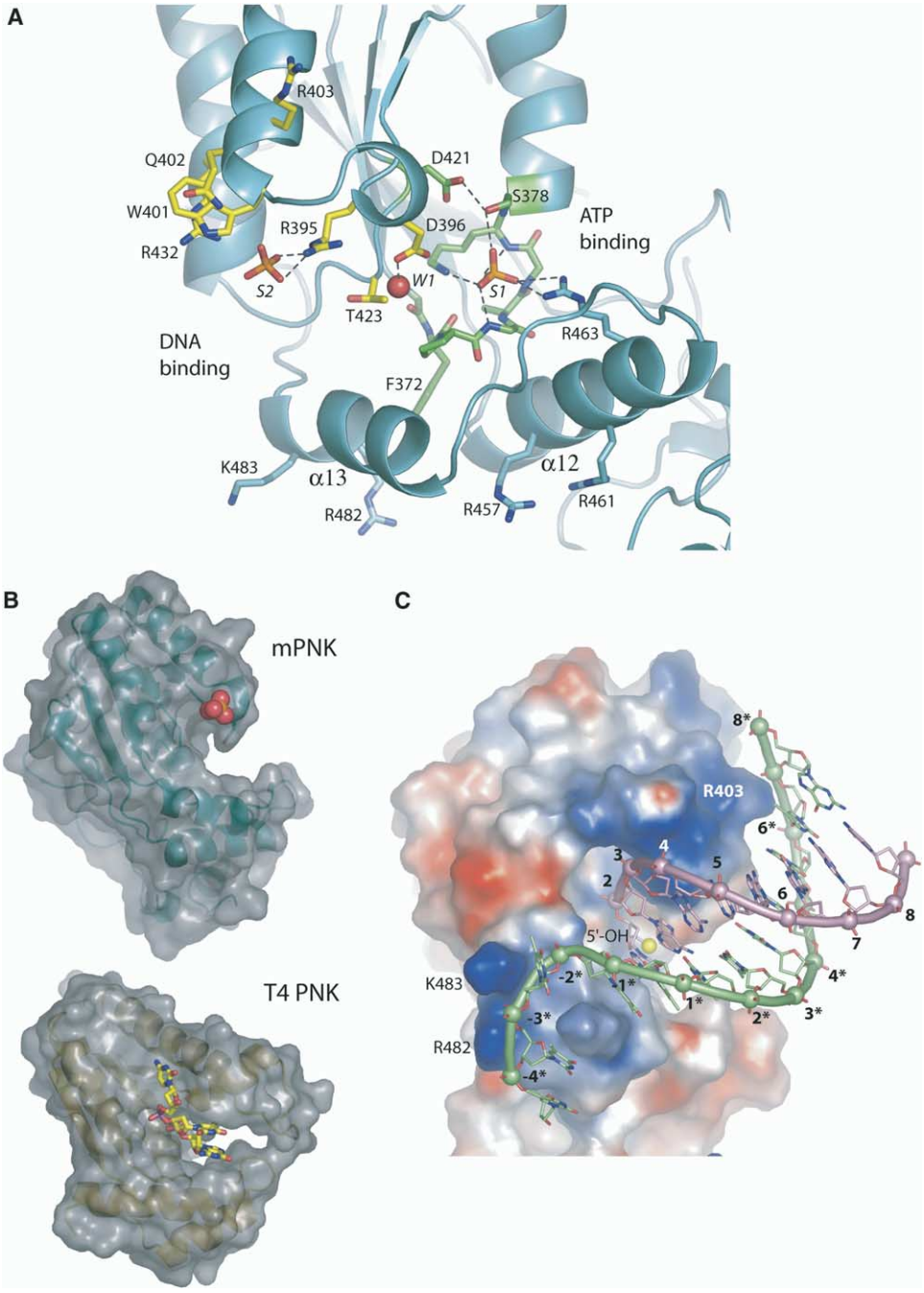

Figure 3. The Kinase Domain

(A) mPNK active site cleft. The Walker A and Walker B motifs, forming the ATP binding site, are shown in green. Sulfate ions bound at the active site are in orange and red. *S1* corresponds to the β-phosphate of ATP, while *S2* is proposed to mimic the binding of a DNA backbone phosphate. The side chains of residues involved in catalysis or proposed to bind the DNA substrate are shown in yellow. Water molecule *W1*(red sphere) is proposed to correspond to the substrate 5′-hydroxyl.

(B) Surface representations of the kinase domains of mPNK and T4 PNK, viewed from the substrate approach side. mPNK is shown with the bound sulfate S2, and T4 PNK is shown bound to oligonucleotide 5′-GTCAC-3′ (PDB ID 1RC8).

(C) Surface representation of mPNK with the modeled DNA substrate, which contains 8 base pairs and a 5′-hydroxyl recessed by 4 nucleotides.

understand the structural basis for the substrate preference of mPNK, we docked the modeled minimal preferred substrate to the kinase domain. Structural studies of T4 kinase-DNA substrate complexes have revealed particularly critical interactions with the 5′-hydroxyl and 3′-phosphate of the 5′ terminal nucleotide (East-

berg et al., 2004). By analogy, we used two reference points, the 5′-hydroxyl and a backbone phosphate, to position the double helical portion of the substrate at the mPNK active site. The 5′-hydroxyl was aligned with the water molecule *W1* (Figure 3A), which corresponds to the 5′-hydroxyl of the oligonucleotide substrate bound to T4 PNK (Eastberg et al., 2004) and forms a hydrogen bond with the catalytic Asp 396. Secondly, a feature in the electron density interpreted as a bound sulfate ion (*S2* in Figure 3A) was assumed to define the location of a backbone phosphate (3′ to the second nucleotide). In the complex with an oligonucleotide substrate, T4 PNK binds the 3′-phosphate of the terminal nucleotide between Thr 86, which corresponds to Thr 423 in mPNK, and Arg 38, which is not conserved in mPNK. The guanidinium of Arg 395 (mPNK) is in approximately the same location as the guanidinium of Arg 38 (T4), but points away from Thr 423 and cannot interact with a phosphate bound in that space. In mPNK, sulfate ion *S2*, which is observed in both molecules in the asymmetric unit and is located between the guanidinium of Arg 395 and the edge of the Trp 401 indole ring (Figure 3A), is a likely candidate for a backbone phosphate mimic. Interestingly, using *W1* and *S2* as points of reference, a modeled minimal preferred DNA substrate can be docked to mPNK such that both the double-stranded portion and the single-stranded overhang contact conserved positively charged residues on the surface of the protein. In the double-stranded portion, the DNA backbone (positions 6-8 on the complementary strand) approaches the protein near a positively charged surface around Arg 403, while 4 nucleotides of the single-stranded 3′ overhang can be positioned near a second positive patch on the other side of the cleft composed of Arg 482 and Lys 483. The model explains the minimal substrate requirements of 8 base pairs of double-stranded DNA (Karimi-Busheri and Weinfeld, 1997) and at least 3 nucleotides of 3′ single-stranded overhang (Figure 2D).

The Phosphatase Domain

Substrate Specificity

To determine the minimal phosphatase substrate requirements, we assessed the phosphatase activity of mPNK on eight 3′-phosphorylated oligonucleotides (Figure 4). The dephosphorylation reaction proceeds equally well on either nicked or gapped double-stranded substrates, as well as a variety of single-stranded substrates of various lengths. Efficient dephosphorylation is maintained in single-stranded DNAs as small as 3 nucleotides. Further deletion to a dinucleotide substrate showed a significant 3-fold decrease in activity, while a 5′-3′-diphospho-mononucleotide resulted in a further decrease. Thus, in contrast to the kinase domain, the phosphatase domain can accept a wide range of substrates, but only requires a very short segment of oligonucleotide chain 5′ to the 3′-phosphate.

Structure of the Phosphatase Domain

The phosphatase domain encompasses residues 146 to 336, and has a fold typical of the haloacid dehalogenase (HAD) superfamily, which also includes β-phosphoglucomutase (β-PGM), phosphoserine phosphatase (PSP), P-type ATPase and RNA polymerase II CTD phosphatase (Aravind et al., 1998; Kamenski et al., 2004) (Figure 5A). All these enzymes hydrolyze a phosphate in a Mg^{2+}-dependent reaction *via* a phosphoaspartate intermediate. Despite a generally low sequence similarity, these proteins are characterized by the conserved motif Dx(D/T)x(T/V), where the first Asp (170 in mPNK) forms the covalent phosphoaspartate intermediate. Additional conserved motifs include a serine or threonine (216), a lysine (259), and a pair of aspartates (288 and 309) (Figure 5B). These conserved residues are involved in catalysis or in binding the Mg^{2+} and the phosphate moiety of the substrate and phosphoaspartate intermediate, and are all located in the HAD α/β domain. Despite structural variations among the α/β domains, the active site configuration is exquisitely preserved (Figures 5A and 5B). The major structural differences arise due to the nonconserved loops or capping structures that protrude above the active site (Figure 5A). While the α/β domain interacts with the phosphate portion, the capping structure recognizes the rest of the substrate, conferring substrate specificity.

The mechanism of phosphate transfer catalyzed by several HAD-family phosphotransferases has been elucidated in a series of structural studies (Cho et al., 2001; Lahiri et al., 2002; Wang et al., 2002b) and the conservation of key catalytic residues in mPNK suggests that essential features of the mechanism are likely to be maintained (Figure 5B). As the phosphorylated substrate binds in the active site, its phosphate group is stabilized by contacts with Lys 259 and Thr 216 as well by the bound Mg^{2+}. Although Mg^{2+} is not observed in the mPNK structure, its position is inferred from the structure of BeF_3-derivatized PSP (Cho et al., 2001). The phosphate is transferred to Oδ1 of Asp 170 in an in-line nucleophilic attack that forms the phosphoaspartate intermediate. Asp 172 may provide general acid assistance by protonating the alcohol leaving group. The phosphoaspartate intermediate has been observed in β-PGM and in PSP (simulated by a BeF_3 adduct) (Cho et al., 2001; Lahiri et al., 2002). In the analogous intermediate envisioned for mPNK, the phosphate interacts with the side chains Lys 269 and Thr 216, with backbone NH of Leu 171 and Asp 172 and with the Mg^{2+}. The metal ion is coordinated by one of the phosphate oxygens, the main chain of Asp 172, the carboxylate groups of Asp 170 and Asp 288, and two water molecules. The carboxylate of Asp 309 helps to stabilize the active site configuration by forming hydrogen bonds with the backbone NH of Leu 171 and Asp 172 and a salt bridge with Lys 259. In T4 PNK and in β-PGM the carboxylate group at this position is contributed by an Asp from a different part of the peptide chain (Asp 277 in T4 PNK). The precise location of this carboxylate is crucial, as mutation of the corresponding Asp (218) in the *S. cerevisiae* homolog, Tpp1, to Glu severely impairs phosphatase activity (Deshpande and Wilson, 2004). Finally, the free enzyme is regenerated in a dephosphorylation reaction, with Asp 172 activating a water molecule for S_N2 attack on the phosphorus atom. This aspartate is conserved in all the phosphatases where the leaving group is an alcohol.

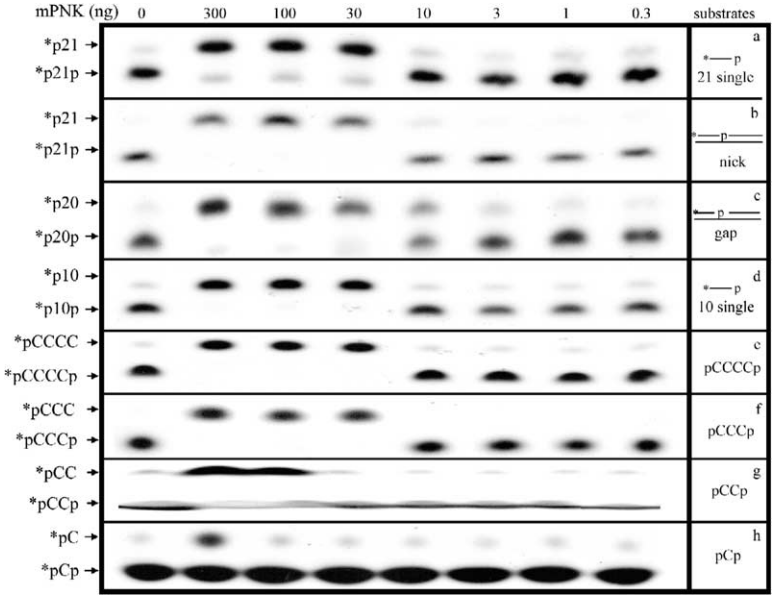

Figure 4. Phosphatase Activity Assay

The amount of 3′ dephosphorylation of various deoxyribonucleotide substrates (shown) is measured at different concentrations of mPNK. The asterisk denotes the position of radiolabeled 5′-phosphate on the substrates. The single-stranded substrates, all bearing a 3′-phosphate, include a (pC_np) series where (n = 1, 2, 3, 4), a 10-mer and a 21-mer (see Supplemental Data for sequences).

A Model for the Phosphatase/Substrate Complex

Substrate specificity varies widely in the HAD phosphotransferase family. To account for our finding that the minimal substrate for the mPNK phosphatase is a single-stranded DNA of at least 3 nucleotides in length, we propose a model of the phosphatase:substrate complex (Figure 5C). In the model, the 3′-phosphate was positioned by overlaying C3′ of the terminal ribose with C1 of glucose-6-phosphate bound to β-PGM (Lahiri et al., 2003), and a nearby sulfate ion located next to Arg 258 in our structure was used as a guide for placing the preceding backbone phosphate. The model was minimized in REFMAC (Winn et al., 2003) to regularize the geometry of the oligonucleotide and to alleviate steric clashes. Arg 223 and Arg 258 form a positive patch at the rim of the active site cleft, and in our substrate model three backbone phosphates can bind at this location. This is consistent with our phosphatase activity assay, which shows that a trinucleotide is a better substrate for mPNK phosphatase than the dinucleotide, which in turn is much better than a mononucleotide (Figure 4, rows f–h). The lack of improvement for substrates longer than three nucleotides agrees with the prediction that only three backbone phosphates make significant contact with this surface patch (Figure 4, rows d and e). The requirement for at least one phosphate at the patch is further supported by the observation of Habraken & Verly that rat PNK does not work on 3′-monodeoxynucleotides, but removes the 3′-phosphate in dinucleotides (Habraken and Verly, 1988). The phosphatase active site cleft appears to be too narrow to admit a double-stranded DNA substrate, but the edges of the cleft are formed by potentially flexible loops that may adapt to different substrates.

Phosphopeptide Recognition by the FHA Domain

The FHA domain possesses the typical FHA β sandwich fold consisting of a 3-stranded and a 4-stranded antiparallel β sheet. The N- and C termini of the domain are located on adjacent β strands and the phosphopeptide is bound by loops on the opposite side of the domain. In the structure of full-length mPNK, the peptide binding loops of the FHA domain were poorly ordered, and it was impossible to predict how this protein would interact with a phosphopeptide. We therefore expressed an isolated FHA domain, which we co-crystallized with an XRCC4-derived phosphopeptide (Ac-YDES(pT)DEESEKK-$CONH_2$ (Koch et al., 2004), Table 1). A number of structures of FHA domains bound to phosphothreonyl peptide ligands are available, and they all display a conserved mode of binding to the phosphothreonine group (Durocher and Jackson, 2002). The FHA domain of mPNK utilizes this standard mode of phosphothreonine recognition, involving the invariant Arg 35 and Ser 47, which is in some cases replaced by Thr. Additional stabilization for the phosphate is provided by Arg 48, which corresponds to Lys or Asn in many other FHA domains (Figures 1B and 6A).

Most FHA domains display additional specificity for side chains 3 residues C-terminal to the phosphothreonine (pT+3, (Durocher and Jackson, 2002)). In contrast, peptide array binding studies showed that the FHA domain of mammalian PNK recognizes peptide positions N-terminal to the phosphothreonine (Koch et al., 2004), and a comparison of the natural PNK FHA binding targets in both XRCC1 and XRCC4 suggests a preference for acidic residues in the vicinity of the phosphothreonine (Loizou et al., 2004; Koch et al., 2004). In our structure of the FHA:XRCC4 peptide complex, the peptide is cradled between two surface loops, β2-β3 and β4-β5. The structure suggests that three positively charged residues from the β2-β3 loop, Arg 44, Lys 45 and Arg 48, may be responsible for the recognition of acidic residues not only N-terminal, but also C-terminal to the phosphothreonine. Arg 48, in addition to its role in phosphate recognition, is also close enough to provide electrostatic recognition of Asp at pT-3. Arg 44 displays different conformations in each of the three inde-

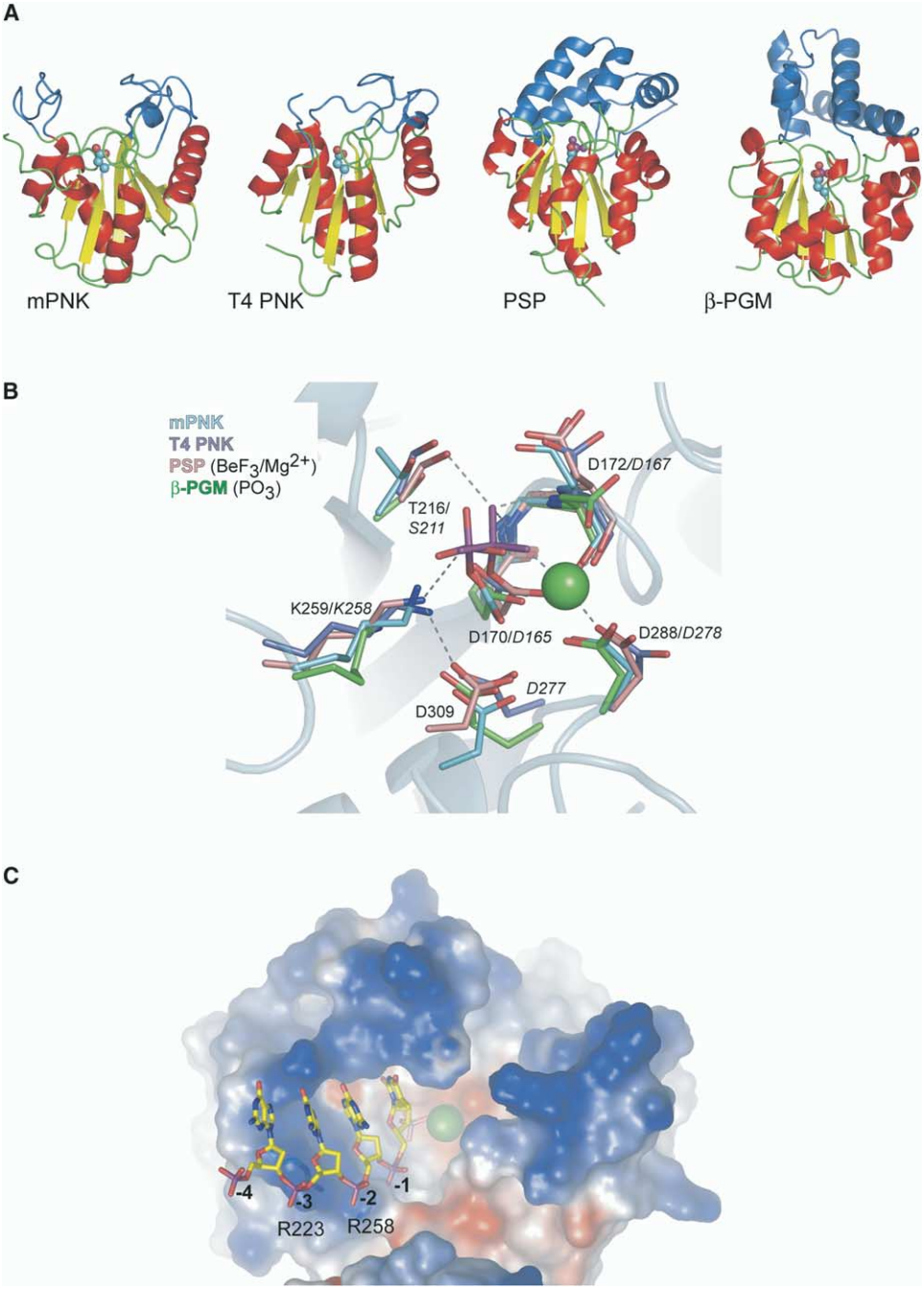

Figure 5. Structure of the Phosphatase Domain

(A) Comparison of mPNK and T4 PNK phosphatase domains, β-PGM and PSP in a common orientation, showing the variation in the active site capping structures (blue). The common α/β domain elements are colored yellow and red, and the catalytic aspartate is shown in ball-and-stick representation.

(B) Structure of the mPNK phosphatase active site (cyan) superimposed on the active sites of phosphorylated β-PGM (green, PDB 1LVH), BeF_3-derivatized PSP (pink, Mg^{2+} in green, PDB 1J97), and T4 PNK (blue, PDB 1LTQ). The hydrogen bonding pattern for PSP is shown. Residue numbering shown is for mPNK and in italics for T4 PNK.

(C) Surface representation of mPNK phosphatase with bound modeled tetranucleotide substrate. The green sphere represents the modeled Mg^{2+}.

pendent complexes in the crystallographic asymmetric unit, occupying a variety of positions between Glu at pT-2 and Asp at the pT+1 position of the peptide (Figure 6B). Lys 45 could potentially provide additional electrostatic selectivity for both Asp at pT+1 and Glu and pT+2. Interestingly, there is no interpretable electron

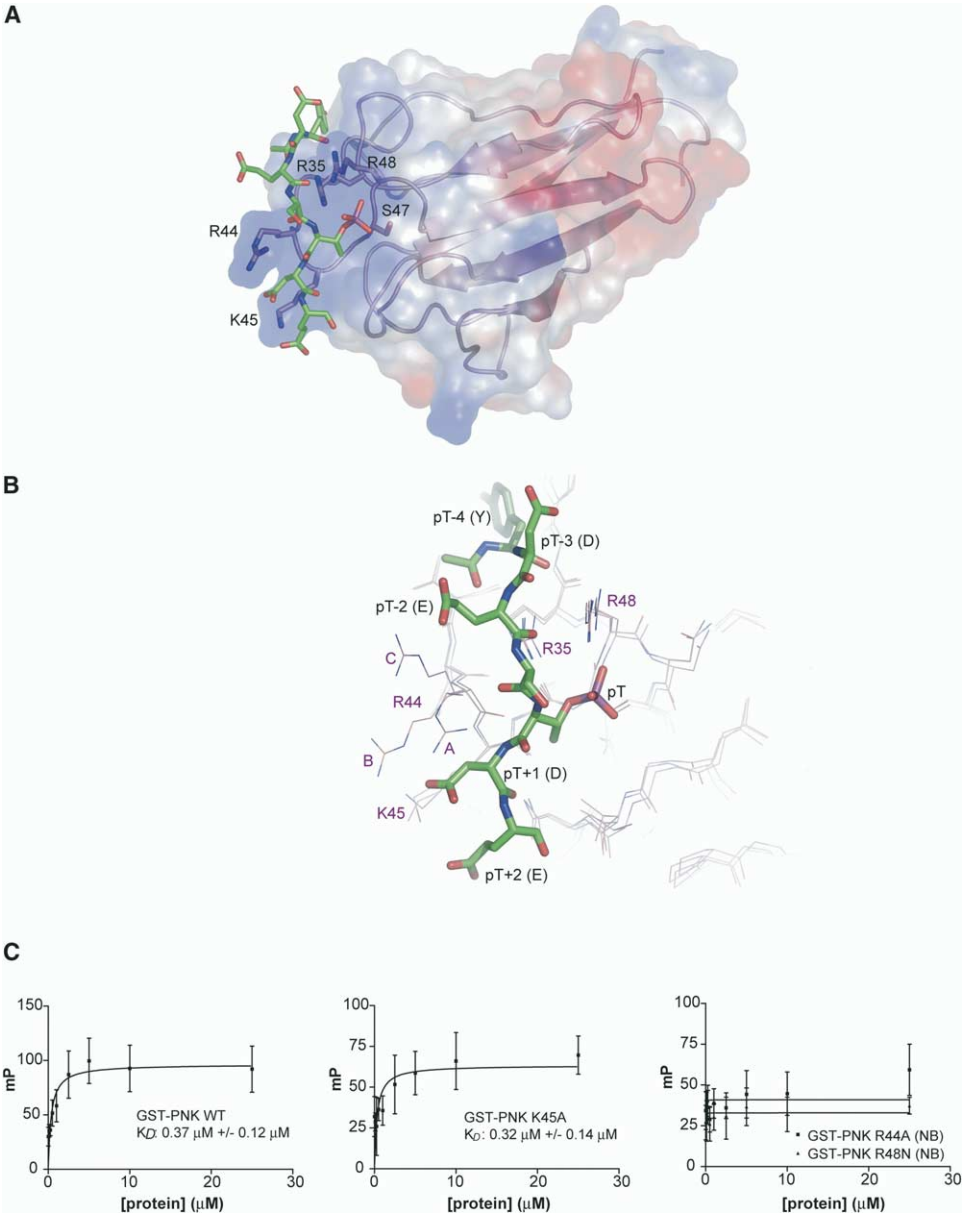

Figure 6. The Basis of Phosphopeptide Recognition by the mPNK FHA Domain

(A) XRCC4 phosphopeptide (carbon atoms in green) bound to the mPNK FHA domain. The electrostatic surface potential indicates that the peptide binding surface is predominantly positively charged.

(B) Three superimposed NCS-related FHA domains (thin lines labeled A, B, and C) are shown with the phosphopeptide from complex (A) (thick lines).

(C) Binding of the GST-tagged mPNK FHA domain (wild-type, R44A, K45A, and R48N) to the fluorescein-labeled peptide GGYDES-pT-DEESKK, measured by fluorescence polarization. Dissociation constants for wild-type and the R48N mutant are shown. NB indicates no binding. Error bars are derived from at least three independent measurements for each data point.

density for regions of the peptide C-terminal to pT+2. This strongly suggests that the PNK FHA does not contact these residues, and is consistent with previous work that has shown that this domain, unlike all other well-studied members of this family, does not exhibit selectivity for the pT+3 residue.

To test the importance of Arg 44, Lys 45 and Arg 48 for phosphopeptide recognition, we mutated each of these residues and used fluorescence polarization to measure their affinities for the XRCC4-derived phosphopeptide, relative to the wild-type domain. Arg 44 and Lys 45 were mutated to alanine, while Arg 48, which makes critical contacts to the phosphothreonine as well as interactions with pT-3, was mutated to Asn, which is found in several FHA domains and has been shown to mediate phosphothreonine recognition. As shown in Figure 6C, the R44A and R48N mutations ablated the ability of the the FHA domain to recognize the

phosphopeptide, confirming an essential role for both of these residues in peptide binding. In marked contrast, mutation of Lys 45 did not affect the binding affinity, indicating that this residue only plays a minor role, if any, in phosphopeptide recognition.

Sequence comparisons of the PNK and aprataxin FHA domains show that all residues involved in phosphopeptide recognition are well-conserved (Figure 1B). It is thus highly likely that the FHA domains of PNK and APTX recognize the same phosphopeptide sequence, in agreement with the observation that these proteins are found in mutually exclusive complexes in vivo (Luo et al., 2004). Curiously, unlike PNK, the pT+3 position appears to be important in the APTX:XRCC1 interaction, since mutation of Glu to Ala at this site in XRCC1 abolished the binding to APTX (Luo et al., 2004). APTX has a single positively charged residue, Lys 75 (Pro 81 in PNK), which we predict to be close to the pT+3 residue and could potentially provide sequence selectivity at this position.

In spite of the striking differences in the mechanism of sequence selectivity between the PNK FHA and other members of the family, the backbone of the peptide in the PNK structure is bound in a similar manner to that found in other FHA-peptide complexes. The PNK FHA contacts the peptide backbone through hydrogen bonding interactions between the conserved Asn 70 and Arg 35, and the peptide backbone at pT+1 and pT-2, respectively. In addition, the pT-4 Tyr is contacted by a stacking interaction with Pro 37.

Integration of mPNK within DNA Repair Pathways

A picture is now emerging of mammalian PNK as an enzyme that has evolved to play specific roles in XRCC1- and XRCC4-mediated DNA repair. The specificity of PNK for these processes is 2-fold. First, the enzyme is initially recruited to the sites of repair through phosphorylation-dependent interactions with XRCC1/4 via its FHA domain. The FHA is flexibly linked to the catalytic domain, and does not appear to interact with, or modulate its activity (not shown). Flexible attachment of the catalytic portion enables either active site to approach the DNA end that requires processing.

The second level of specificity arises from the DNA substrate preferences of the enzyme. Intriguingly, the two catalytic domains have complementary and nonoverlapping minimal substrate requirements. This may have evolved to ensure that the two catalytic activities, while held together in a relatively rigid structure, nevertheless act independently, enabling the enzyme to process a variety of substrates (eg: nicks, gaps of various sizes, and double-strand breaks). While the two active sites are on the same side of the protein, they are far apart (~40 Å), and it seems unlikely that they will be able to bind to the same substrate DNA, especially for small nicks or gaps. In contrast, T4 PNK probably evolved to bind simultaneously the 5′-hydroxyl and the 3′-phosphate of its nicked tRNA substrate (Galburt et al., 2002).

The kinase strongly prefers 5′ recessed ends over blunt or single-stranded ends, consistent with its role in SSBR/BER, in which gapped and nicked substrates are expected. For DSBR, however, blunt and single-stranded 5′ ends will be encountered, which would be poor substrates for mPNK. We suggest that such ends might first be processed by the NHEJ-associated nuclease Artemis (Ma et al., 2002; Pannicke et al., 2004). Artemis has single-strand-specific 5′-3′ exonuclease and endonuclease activities that could remove nucleotides containing 5′-hydroxyl termini from 5′ overhanging ends, leaving a blunt 5′-phosphate terminus that would be suitable for ligation. Artemis also cleaves long 3′ overhangs, resulting in short 4–5 nucleotide 3′ overhangs, the optimal substrate for mPNK.

Finally, we can envision an application for PNK to the treatment of cancer. Radiation therapy and certain chemotherapies work by damaging DNA in tumors. Since various DNA repair processes may attenuate the effect of therapy, it may be beneficial to down-regulate repair. PNK, being essential in both SSBR and DSBR, may be a suitable target for inactivation. Inhibitors may be designed against either catalytic activity, or against the FHA/XRCC(1/4) interaction.

Experimental Procedures

Cloning, Expression, and Purification of mPNK Constructs

Two cDNA clones of mouse PNK (598211 and 2865792) were obtained from the IMAGE consortium and cloned into the EcoRI and XhoI sites of pET19b (Invitrogen). DNA sequencing of the two expression clones showed that clone 598211, derived from mouse spleen, was alternately spliced, lacking exon 4, which encodes part of the phosphatase active site. Clone 2865792, from mouse mammary tumor, coded for the full length PNK, and was used in all subsequent experiments. Mouse PNK was overexpressed in *E. coli* BL21 Gold cells induced by 0.1mM IPTG and grown for 18 hr at 22°C. The protein was purified as described previously for human PNK (Mani et al., 2003). After the last chromatographic step, the buffer was exchanged to 150 mM KCl, 10 mM Tris (pH 8.5), 1mM DTT. Selenomethionyl mPNK was produced as described in (Doublie, 1997) and purified in the same manner as the native protein.

The mPNK FHA domain (residues 1-110) was cloned using the Gateway system (Invitrogen) into the pDEST15 expression vector with an N-terminal GST tag followed by a Prescission protease site. The GST-tagged protein was produced in *E.coli* BL21 Gold99 cells (induced with 0.1mM IPTG and grown overnight at 22°C), and purified on glutathione (GSH) resin (Amersham). The GST tag was cleaved off in solution (20 mM reduced GSH, 50 mM Tris (pH 7.5), 400 mM NaCl, 0.1% β-mercaptoethanol) by Prescission protease (Amersham) at 4°C. The FHA domain was separated from GST by size exclusion chromatography on a Superdex 75 column (Amersham) in 50 mM Tris (pH 7.5), 400 mM NaCl, 1mM DTT.

The R44A, K45A and R48N point mutants of the FHA domain were constructed by PCR mutagenesis and Gateway cloning. The mutants were expressed and purified by the same method as the wild-type mPNK FHA, but without removal of the GST tag.

Kinase Activity Assays

Oligonucleotides (see Supplemental Data for sequences) were purified on a Source Q column (Amersham) in 10 mM NaOH with a gradient of 1 M NaCl, desalted using Sep-Pak cartridges (Waters), and annealed by slow cooling from 95°C in 150 mM NaCl and 10 mM HEPES (pH 7.5). Each 10 μl assay reaction contained the required amounts of DNA and mPNK, 400 pmol ATP, 5 μCi γ^{32}P-ATP, 80 mM succinate (pH 5.5), 10 mM $MgCl_2$, and 1mM DTT. Reactions were assembled on ice, incubated 4 min at 37°C, stopped with 10 μl urea loading buffer, heated 5 min at 95°C and loaded on a 12% denaturing gel (containing 7M urea), run in TBE.

Fluorescence Spectroscopy

Steady-state fluorescence spectra were measured at 5 °C on a Perkin-Elmer LS-55 spectrofluorometer (with 5-nm spectral resolution

for excitation and emission). DNA ligands were added from stock solutions to mPNK (0.15 μM at 5 °C in 50 mM Tris, pH 7.5, 100 mM NaCl, 5 mM $MgCl_2$, and 1 mM DTT). The protein was excited at 295 nm, keeping the total absorption < 0.05, and the fluorescence intensity was monitored at 335 nm as a function of oligonucleotide concentration, correcting fluorescence intensities for dilution. Fluorescence data were analyzed as described previously (Mani et al., 2003).

Phosphatase Activity Assay

The 3′-dephosphorylation of DNA substrates by mPNK was performed as described previously (Karimi-Busheri et al., 1998). The 5′ termini of the oligonucleotides (16 pmol in 10 μl) were radiolabeled by incubation with phosphatase-free T4 PNK (Roche) and 3.3 pmol γ^{32}P-ATP in 50 mM Tris-HCl, 10mM $MgCl_2$, 5 mM DTT, pH 7.5 for 30 s at 37°C. Reactions were stopped by incubation for 10 min at 95°C. Double-stranded substrates were annealed by slow cooling in 50 mM Tris-HCl, 10 mM $MgCl_2$, 5 mM DTT, pH 7.5. The 5′-labeled 3′-phosphorylated constructs (2.5 pmol/10 μl) were used to assess the 3′-phosphatase activity of mPNK in a reaction mixture containing different quantities of mPNK incubated at 37°C for 20 min in the same buffer used for 5′-labeling. The labeled products were resolved by 20% denaturing PAGE for the mono and dinucleotide substrates and by 12% gel for the longer oligonucleotide and duplex substrates.

Fluorescence Polarization

The binding of GST-tagged mPNK FHA domain (wild-type and the R44A, K45A and R48N point mutants) to fluorescein-labeled peptide GGYDES-pT-DEESKK was measured by fluorescence polarization as described previously (Koch et al., 2004).

Crystallization

Mouse PNK and the mPNK FHA domain were crystallized by vapor diffusion in sitting or hanging drops. Native or selenomethionyl mPNK (2.5 mg/mL) in 150 mM KCl, 10 mM Tris (pH 8.5), 0.1 mM EDTA and 1 mM DTT was mixed with an equal volume of the reservoir solution (0.1 M Tris (pH 8.3 - 8.7), 18 - 20% PEG 5000 MME, 0.1 M Li_2SO_4 and 5 mM DTT). Thin plates grew at room temperature within 1-4 days.

The FHA domain was concentrated in the presence of a 1.5 molar excess of XRCC4 peptide (Ac-YDES(pT)DEESEKK-$CONH_2$, Alberta Peptide Institute) to 60 mg/mL and a final peptide to FHA ratio of 1.17:1. The FHA/peptide solution was mixed with an equal volume of reservoir solution (0.1 M Na citrate (pH 5.5), 25% PEG 4000, 0.2 M Li_2SO_4 and 5 mM DTT). A cluster of plates grew at room temperature within 2 weeks.

Structure Determination

See Supplemental materials for details of the diffraction data collection and processing as well as structure determination and refinement. Briefly, the structure of full-length mouse PNK was solved by selenomethionyl MAD. The structure of the FHA:phosphopeptide complex was solved by molecular replacement with the FHA domain from the full-length structure as the search model.

Supplemental Data

Supplemental Data include Supplemental Experimental Procedures and one additional figure and can be found with this article online at http://www.molecule.org/cgi/content/full/17/5/657/DC1/.

Acknowledgments

We would like to thank members of the Glover laboratory, Mesfin Fanta, and staff at the Alberta Synchrotron Institute for helpful discussions and data collection support. This research was funded by grants from the Canadian Institutes for Health Research (J.N.M.G., D.D., and M.W.) and by the National Cancer Institute of Canada (J.N.M.G.) J.N.M.G acknowledges with thanks the award of a Canada Research Chair. N.K.B gratefully acknowledges the Terry Fox Foundation for a postdoctoral fellowship.

Received: December 3, 2004
Revised: January 7, 2005
Accepted: February 2, 2005
Published: March 3, 2005

References

Amitsur, M., Levitz, R., and Kaufmann, G. (1987). Bacteriophage T4 anticodon nuclease, polynucleotide kinase and RNA ligase reprocess the host lysine tRNA. EMBO J. *6*, 2499–2503.

Aravind, L., Galperin, M.Y., and Koonin, E.V. (1998). The catalytic domain of the P-type ATPase has the haloacid dehalogenase fold. Trends Biochem. Sci. *23*, 127–129.

Bosdal, T., and Lillehaug, J.R. (1985). Purification and kinetic properties of polynucleotide kinase from rat testes. Biochim. Biophys. Acta *840*, 280–286.

Caldecott, K.W. (2003). DNA single-strand break repair and spinocerebellar ataxia. Cell *112*, 7–10.

Chappell, C., Hanakahi, L.A., Karimi-Busheri, F., Weinfeld, M., and West, S.C. (2002). Involvement of human polynucleotide kinase in double-strand break repair by non-homologous end joining. EMBO J. *21*, 2827–2832.

Cho, H., Wang, W., Kim, R., Yokota, H., Damo, S., Kim, S.H., Wemmer, D., Kustu, S., and Yan, D. (2001). BeF(3)(-) acts as a phosphate analog in proteins phosphorylated on aspartate: structure of a BeF(3)(-) complex with phosphoserine phosphatase. Proc. Natl. Acad. Sci. USA *98*, 8525–8530.

Clements, P.M., Breslin, C., Deeks, E.D., Byrd, P.J., Ju, L., Bieganowski, P., Brenner, C., Moreira, M.C., Taylor, A.M., and Caldecott, K.W. (2004). The ataxia-oculomotor apraxia 1 gene product has a role distinct from ATM and interacts with the DNA strand break repair proteins XRCC1 and XRCC4. DNA Repair (Amst.) *3*, 1493–1502.

Deshpande, R.A., and Wilson, T.E. (2004). Identification of DNA 3′-phosphatase active site residues and their differential role in DNA binding, Mg2+ coordination, and catalysis. Biochemistry *43*, 8579–8589.

Doublie, S. (1997). Preparation of selenomethionyl proteins for phase determination. Methods Enzymol. *276*, 523–530.

Durocher, D., and Jackson, S.P. (2002). The FHA domain. FEBS Lett. *513*, 58–66.

Eastberg, J.H., Pelletier, J., and Stoddard, B.L. (2004). Recognition of DNA substrates by T4 bacteriophage polynucleotide kinase. Nucleic Acids Res. *32*, 653–660.

Galburt, 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 (Camb.) *10*, 1249–1260.

Gueven, N., Becherel, O.J., Kijas, A.W., Chen, P., Howe, O., Rudolph, J.H., Gatti, R., Date, H., Onodera, O., Taucher-Scholz, G., and Lavin, M.F. (2004). Aprataxin, a novel protein that protects against genotoxic stress. Hum. Mol. Genet. *13*, 1081–1093.

Habraken, Y., and Verly, W.G. (1983). The DNA 3′-phosphatase and 5′-hydroxyl kinase of rat liver chromatin. FEBS Lett. *160*, 46–50.

Habraken, Y., and Verly, W.G. (1988). Further purification and characterization of the DNA 3′-phosphatase from rat-liver chromatin which is also a polynucleotide 5′-hydroxyl kinase. Eur. J. Biochem. *171*, 59–66.

Hazra, T.K., Kow, Y.W., Hatahet, Z., Imhoff, B., Boldogh, I., Mokkapati, S.K., Mitra, S., and Izumi, T. (2002). Identification and characterization of a novel human DNA glycosylase for repair of cytosine-derived lesions. J. Biol. Chem. *277*, 30417–30420.

Izumi, T., Wiederhold, L.R., Roy, G., Roy, R., Jaiswal, A., Bhakat, K.K., Mitra, S., and Hazra, T.K. (2003). Mammalian DNA base excision repair proteins: their interactions and role in repair of oxidative DNA damage. Toxicology *193*, 43–65.

Kamenski, T., Heilmeier, S., Meinhart, A., and Cramer, P. (2004). Structure and mechanism of RNA polymerase II CTD phosphatases. Mol. Cell *15*, 399–407.

Karimi-Busheri, F., Lee, J., Tomkinson, A.E., and Weinfeld, M.

(1998). Repair of DNA strand gaps and nicks containing 3′-phosphate and 5′-hydroxyl termini by purified mammalian enzymes. Nucleic Acids Res. *26*, 4395–4400.

Karimi-Busheri, F., and Weinfeld, M. (1997). Purification and substrate specificity of polydeoxyribonucleotide kinases isolated from calf thymus and rat liver. J. Cell. Biochem. *64*, 258–272.

Kleppe, K., and Lillehaug, J.R. (1979). Polynucleotide kinase. Adv. Enzymol. Relat. Areas Mol. Biol. *48*, 245–275.

Koch, C.A., Agyei, R., Galicia, S., Metalnikov, P., O'Donnell, P., Starostine, A., Weinfeld, M., and Durocher, D. (2004). Xrcc4 physically links DNA end processing by polynucleotide kinase to DNA ligation by DNA ligase IV. EMBO J. *23*, 3874–3885.

Lahiri, S.D., Zhang, G., Dunaway-Mariano, D., and Allen, K.N. (2002). Caught in the act: the structure of phosphorylated beta-phosphoglucomutase from Lactococcus lactis. Biochemistry *41*, 8351–8359.

Lahiri, S.D., Zhang, G., Dunaway-Mariano, D., and Allen, K.N. (2003). The pentacovalent phosphorus intermediate of a phosphoryl transfer reaction. Science *299*, 2067–2071.

Leipe, D.D., Koonin, E.V., and Aravind, L. (2003). Evolution and classification of P-loop kinases and related proteins. J. Mol. Biol. *333*, 781–815.

Lillehaug, J.R., Kleppe, R.K., and Kleppe, K. (1976). Phosphorylation of double-stranded DNAs by T4 polynucleotide kinase. Biochemistry *15*, 1858–1865.

Loizou, J.I., El-Khamisy, S.F., Zlatanou, A., Moore, D.J., Chan, D.W., Qin, J., Sarno, S., Meggio, F., Pinna, L.A., and Caldecott, K.W. (2004). The protein kinase CK2 facilitates repair of chromosomal DNA single-strand breaks. Cell *117*, 17–28.

Luo, H., Chan, D.W., Yang, T., Rodriguez, M., Chen, B.P., Leng, M., Mu, J.J., Chen, D., Songyang, Z., Wang, Y., and Qin, J. (2004). A new XRCC1-containing complex and its role in cellular survival of methyl methanesulfonate treatment. Mol. Cell. Biol. *24*, 8356–8365.

Ma, Y., Pannicke, U., Schwarz, K., and Lieber, M.R. (2002). Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. Cell *108*, 781–794.

Mani, R.S., Karimi-Busheri, F., Cass, C.E., and Weinfeld, M. (2001). Physical properties of human polynucleotide kinase: hydrodynamic and spectroscopic studies. Biochemistry *40*, 12967–12973.

Mani, R.S., Karimi-Busheri, F., Fanta, M., Cass, C.E., and Weinfeld, M. (2003). Spectroscopic studies of DNA and ATP binding to human polynucleotide kinase: evidence for a ternary complex. Biochemistry *42*, 12077–12084.

Meijer, M., Karimi-Busheri, F., Huang, T.Y., Weinfeld, M., and Young, D. (2002). Pnk1, a DNA kinase/phosphatase required for normal response to DNA damage by gamma-radiation or camptothecin in Schizosaccharomyces pombe. J. Biol. Chem. *277*, 4050–4055.

Midgley, C.A., and Murray, N.E. (1985). T4 polynucleotide kinase; cloning of the gene (pseT) and amplification of its product. EMBO J. *4*, 2695–2703.

Mitra, S., Izumi, T., Boldogh, I., Bhakat, K.K., Hill, J.W., and Hazra, T.K. (2002). Choreography of oxidative damage repair in mammalian genomes. Free Radic. Biol. Med. *33*, 15–28.

Moreira, M.C., Barbot, C., Tachi, N., Kozuka, N., Uchida, E., Gibson, T., Mendonca, P., Costa, M., Barros, J., Yanagisawa, T., et al. (2001). The gene mutated in ataxia-ocular apraxia 1 encodes the new HIT/Zn-finger protein aprataxin. Nat. Genet. *29*, 189–193.

Pannicke, U., Ma, Y., Hopfner, K.P., Niewolik, D., Lieber, M.R., and Schwarz, K. (2004). Functional and biochemical dissection of the structure-specific nuclease ARTEMIS. EMBO J. *23*, 1987–1997.

Pheiffer, B.H., and Zimmerman, S.B. (1982). 3′-Phosphatase activity of the DNA kinase from rat liver. Biochem. Biophys. Res. Commun. *109*, 1297–1302.

Rasouli-Nia, A., Karimi-Busheri, F., and Weinfeld, M. (2004). Stable down-regulation of human polynucleotide kinase enhances spontaneous mutation frequency and sensitizes cells to genotoxic agents. Proc. Natl. Acad. Sci. USA *101*, 6905–6910.

Rosenquist, T.A., Zaika, E., Fernandes, A.S., Zharkov, D.O., Miller, H., and Grollman, A.P. (2003). The novel DNA glycosylase, NEIL1, protects mammalian cells from radiation-mediated cell death. DNA Repair (Amst.) *2*, 581–591.

Teraoka, H., Mizuta, K., Sato, F., Shimoyachi, M., and Tsukada, K. (1975). Polynucleotide kinase from rat-liver nuclei. Purification and properties. Eur. J. Biochem. *58*, 297–302.

Wang, J.C. (1996). DNA topoisomerases. Annu. Rev. Biochem. *65*, 635–692.

Wang, L.K., Lima, C.D., and Shuman, S. (2002a). Structure and mechanism of T4 polynucleotide kinase: an RNA repair enzyme. EMBO J. *21*, 3873–3880.

Wang, W., Cho, H.S., Kim, R., Jancarik, J., Yokota, H., Nguyen, H.H., Grigoriev, I.V., Wemmer, D.E., and Kim, S.H. (2002b). Structural characterization of the reaction pathway in phosphoserine phosphatase: crystallographic "snapshots" of intermediate states. J. Mol. Biol. *319*, 421–431.

Whitehouse, C.J., Taylor, R.M., Thistlethwaite, A., Zhang, H., Karimi-Busheri, F., Lasko, D.D., Weinfeld, M., and Caldecott, K.W. (2001). XRCC1 stimulates human polynucleotide kinase activity at damaged DNA termini and accelerates DNA single-strand break repair. Cell *104*, 107–117.

Wiederhold, L., Leppard, J.B., Kedar, P., Karimi-Busheri, F., Rasouli-Nia, A., Weinfeld, M., Tomkinson, A.E., Izumi, T., Prasad, R., Wilson, S.H., et al. (2004). AP endonuclease-independent DNA base excision repair in human cells. Mol. Cell *15*, 209–220.

Winn, M.D., Murshudov, G.N., and Papiz, M.Z. (2003). Macromolecular TLS refinement in REFMAC at moderate resolutions. Methods Enzymol. *374*, 300–321.

Zhou, T., Lee, J.W., Tatavarthi, H., Lupski, J.R., Valerie, K., and Povirk, L.F. (2005). Deficiency in 3′-phosphoglycolate processing in human cells with a hereditary mutation in tyrosyl-DNA phosphodiesterase (TDP1). Nucleic Acids Res. *33*, 289–297.

Accession Numbers

Coordinates and structure factors for mPNK and the FHA:phosphopeptide complex have been deposited in the Protein Data Bank under ID codes 1YJ5 and 1YJM, respectively.