Structure, Vol. 10, 1–20, September, 2002, ©2002 Elsevier Science Ltd. All rights reserved. PII S0969-2126(02)00■

# Structure of a tRNA Repair Enzyme and Molecular Biology Workhorse: T4 Polynucleotide Kinase

**Eric A. Galburt,[1] John Pelletier,[2] Geoffrey Wilson,[2] and Barry L. Stoddard[1,3]**
[1]Fred Hutchinson Cancer Research Center and
The Graduate Program in Biomolecular Structure and Design
University of Washington
1100 Fairview Avenue North, A3-023
Seattle, Washington 98109
[2]New England Biolabs
32 Tozer Road
Beverly, Massachusetts 01915

## Summary

**T4 phage polynucleotide kinase (PNK) was identified over 35 years ago and has become a staple reagent for molecular biologists. The enzyme displays 5′-hydroxyl kinase, 3′-phosphatase, and 2′,3′-cyclic phosphodiesterase activities against a wide range of substrates. These activities modify the ends of nicked tRNA generated by a bacterial response to infection and facilitate repair by T4 RNA ligase. DNA repair enzymes that share conserved motifs with PNK have been identified in eukaryotes. PNK contains two functionally distinct structural domains and forms a homotetramer. The C-terminal phosphatase domain is homologous to the L-2-haloacid dehalogenase family and the N-terminal kinase domain is homologous to adenylate kinase. The active sites have been characterized through structural homology analyses and visualization of bound substrate.**

## Introduction

T4 polynucleotide kinase (PNK) was initially discovered in protein extracts of *Escherichia coli* bacteria infected with T-even phage [1–3]. In the 37 years since, the enzyme has become a standard reagent for the manipulation and study of polynucleotides [4–6]. The product of the T4 early gene *pse*T [7], PNK catalyzes the transfer of the γ-phosphate group from adenosine triphosphate (ATP) or other nucleoside triphosphates [3, 8] to the 5′-hydroxyl of polynucleotides of different lengths, sequences, and types [1]. Acceptable substrates include double- and single-stranded DNA, RNA, and individual 3′-phosphate nucleotide bases. PNK also possesses a 3′-phosphatase activity that is independent of ATP [7, 9] and has been reported to hydrolyze 2′,3′-cyclic phosphodiesters [10]. Each catalytic activity is dependent on magnesium as a cofactor. The *pse*T gene was cloned and sequenced [11], which facilitated both the use of recombinant PNK as a reagent and the biochemical study of PNK itself.

The only directly known biological function of T4 PNK is related to the ability of T4 phage to overcome a suicide defense mechanism used by some strains of bacteria [7, 10]. In response to T4 infection, a latent tRNA anticodon nuclease encoded by *prr* strains of *E. coli* is activated by the Stp DNA restriction inhibitory peptide [12–14] and introduces a break in the bacterial lysine tRNA. This lesion is defined by 2′,3′-cyclic phosphate and 5′-hydroxyl ends and effectively prevents protein synthesis, thus inhibiting phage replication. T4 PNK phosphorylates the 5′-hydroxyl and reprocesses the 3′ end by opening the 2′,3′-cyclic phosphate and then removing the 3′-phosphate. The reprocessed ends are substrates for T4 RNA ligase. Thus, the phage is able to circumvent the tRNA lesion and continue to propagate through the bacterial population.

T4 PNK is the prototypical member of a broad family of 5′-kinase/3′-phosphatase enzymes that mend broken strands in nucleic acids in conjunction with the appropriate RNA or DNA ligase. Human members of this family have recently been implicated in the repair of DNA strand breaks caused by oxidative damage [15–17]. In mammalian cells, the conversion of 5′-hydroxyl/3′-phosphate DNA ends to 5′-phosphate/3′-hydroxyl is a necessary step for the repair of DNA nicks and gaps that may be introduced by ionizing radiation [18], certain alkylating agents, DNase II [19, 20], and the enzymatic removal of dead-end complexes of topoisomerase I generated by camptothecin inhibition [21]. Mammalian PNKs display the bifunctionality of T4 PNK, but differ in specificity. While T4 PNK has a decreased activity against both recessed 5′ ends and nicks in double-stranded DNA [22], mammalian PNKs show no preference for 5′ overhanging termini and are active against gaps and nicks, which further implicates them in DNA repair pathways [23]. Mammalian PNKs have been identified that are DNA specific while others have a preference for RNA [16]. Sequence alignments of *H. sapiens*, *C. elegans*, and *S. pombe* PNK enzymes reveal that T4 and mammalian enzymes share two active site motifs: the ATP binding P loop motif which defines a kinase active site and a short motif (DxDxT) associated with a family of phosphatase domains characterized by l-2-haloacid dehalogenase (HAD) [15]. These conserved motifs suggest that the chemical mechanisms of T4 PNK and mammalian PNKs are similar.

Kinetic studies of T4 PNK have shown that the kinase activity proceeds according to an ordered sequential mechanism [24] and that phosphoryl transfer results in an inversion of configuration at the phosphorus atom [25]. Furthermore, the kinase reaction is reversible. For example, the ATP analog β,γ-imidoadenylyl 5′-triphosphate acts as an inhibitor to the forward reaction, but was able to replace ATP in the reverse reaction to generate β,γ-imidoadenylyl 5′-tetraphosphate [26]. While the minimal phosphorylation substrates are 3′ nucleotide monophosphates, T4 PNK is able to phosphorylate a variety of oligonucleotides with a slight preference for a 5′-terminal guanosine base [27]. Despite the tetrameric form of the enzyme, it does not display kinetic cooperativity [24]. It has also been noted that N-protected deoxy-

[3]Correspondence: bstoddar@fhcrc.org

**Key words:** T4 polynucleotide kinase; phosphatase; tRNA repair; bifunctional enzyme; crystal structure; structural enzymology

oligonucleotides do not interfere with kinase activity [28] and that PNK will phosphorylate several nonnucleosidic moieties when they are 5′ attached to oligonucleotides [29]. The phosphatase activity is known to require magnesium or cobalt, has a pH optimum of 6.0, and is more active against DNA than RNA [9].

The kinase and phosphatase activities of T4 PNK are separable and are loosely localized to the N and C termini of the protein, respectively [30]. Further proteolytic experiments coupled with mutational analyses have refined this separation [31, 32]. A truncated N-terminal PNK consisting of residues 1–181 forms a monomer with no phosphatase and low kinase activity, while a construct missing residues 1–47 forms a dimer with no kinase and residual phosphatase activity. Mutations near the amino terminus (K15A, S16A, D35A, R38A, and R126A) ablate kinase activity and do not affect phosphatase activity or the quaternary structure of the enzyme. This constellation of residues is also present in the active sites of adenylate kinases and similar enzymes. In contrast, mutations nearer to the enzyme C terminus (D165, D167A, R176A, R213A, D254A, and D278A) each inactivate the phosphatase activity without affecting the kinase activity or the quaternary structure of the enzyme. The motif $D_{165}xD_{167}xT_{169}$ found in T4 PNK is shared by a superfamily of phosphotransferases [32].

The structure of T4 PNK represents the second structure of a bifunctional kinase/phosphatase enzyme. The first structure is that of the dimeric 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (PFK/FBP) [33]. This enzyme plays an indirect role in the regulation of glucose metabolism. The N-terminal kinase domain of PFK/FBP also shares sequence (and structural) homology with adenylate kinases, whereas the C-terminal phosphatase domain belongs to the phosphoglycerate mutase and acid phosphatase enzyme families. Both of these families have similar topologies to the haloacid dehalogenase (HAD) family, but have a catalytic histidine instead of an aspartic acid [33]. The presence of two catalytic functions in one polypeptide chain usually indicates a functional linkage. For T4 PNK and the multifunctional eukaryotic polynucleotide kinases, the covalent linkage of catalytic activities surely exists due to the requirement for coupled kinase and phosphatase activities in the repair of RNA and DNA lesions.

We have solved the structure of the intact T4 PNK tetramer and find that the tetramer is created through two N-terminal/N-terminal domain interfaces and two C-terminal/C-terminal domain interfaces. The kinase domain belongs to the adenylate kinase family of enzymes and is observed with a bound ADP molecule in the active site. The phosphatase domain belongs to a family of enzymes characterized by l-2-haloacid dehalogenase and its active site is analyzed through homology modeling. Last, a tRNA binding mode is proposed in which the termini of nicked tRNA are bound in both the kinase and phosphatase active site simultaneously, allowing the enzyme to make both modifications within the same substrate binding event.

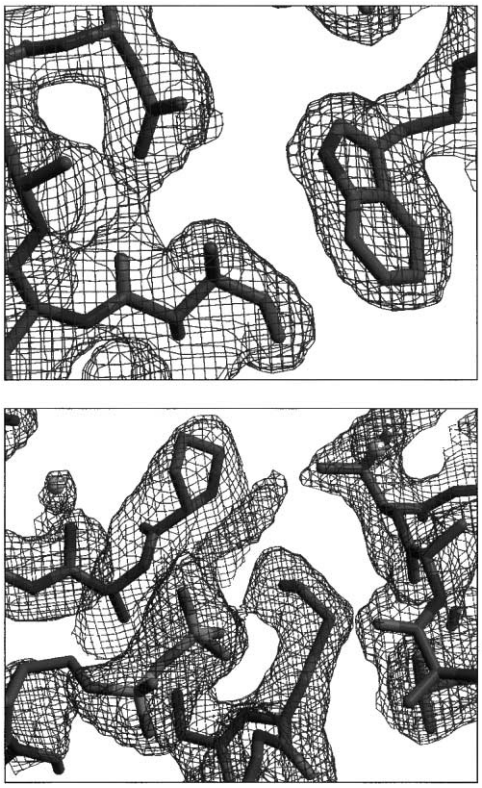
## Results

### Enzyme Fold and Oligomerization

The structure of selenomethionine-substituted T4 PNK with bound nucleotide was determined by the single anomalous dispersion (SAD) method at 2.3 Å resolution. The structure was refined to an $R_{work}/R_{free}$ of 0.236/0.261 with good geometry and an average protein B factor of 36.5 Å$^2$ (Table 1 and side bar). The crystallographic asymmetric unit contains a single 34 kDa subunit of T4 PNK and the I222 crystallographic symmetry operators generate two complete enzyme tetramers in the unit cell. Interpretable main chain electron density was ob-

Table 1. Data Collection and Phasing and Refinement Statistics

| | |
|---|---|
| **Data Collection** | |
| Space group | I222 |
| Unit cell (Å) | a = 78.6 b = 93.7 c = 124.2 |
| Wavelength (Å) | 0.97873 |
| Resolution (highest shell) (Å) | 50–2.3 (2.38–2.30) |
| Total reflections | 191,084 |
| Unique reflections | 38,288 |
| Completeness (%) | 99.1 (97.7) |
| $R_{merge}$ (%) | 7.3 (28.8) |
| Average I/σI | 33.4 (5.7) |
| **Phasing** | |
| Phasing power | 2.56 |
| Figure of merit (before/after DM) | 0.40/0.96 |
| **Refinement** | |
| $R_{cryst}$ (%) | 23.6 |
| $R_{free}$ (%) | 26.1 |
| Total number of atoms | 2532 |
| Number of water molecules | 144 |
| Number of ligand/cryo atoms | 27/4 |
| Rmsd bond length (Å) | 0.007 |
| Rmsd bond angle (°) | 1.4 |
| Mean B value, main chain (Å$^2$) | 35.9 |
| Mean B value, side chain (Å$^2$) | 37.6 |
| Mean B value, solvent (Å$^2$) | 47.3 |



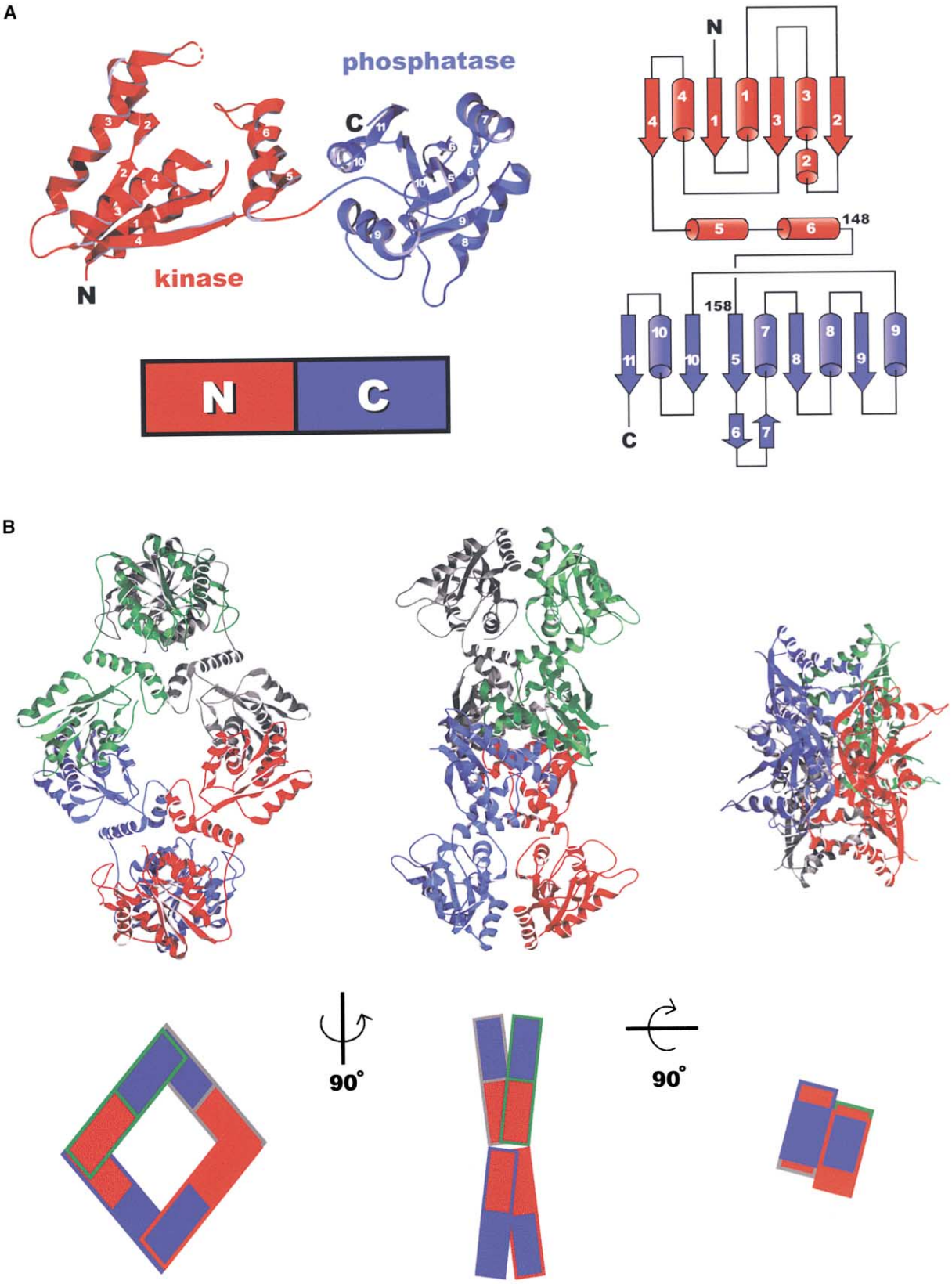



Figure 1. The Structure of T4 Polynucleotide Kinase

A monomer of T4 polynucleotide kinase is shown as a ribbon diagram with the N-terminal kinase domain in red and the C-terminal phosphatase domain in blue (A). The secondary structural elements are numbered according to their occurrence in the chain. A topology diagram shows β strands as arrows and α helices as cylinders. The tetrameric form of the enzyme is shown color coded by monomer (B). Subsequent images are rotated 90° as indicated and reveal the 222 symmetry of the tetramer. The block diagrams under each view of the tetramer show the positions of the N-terminal (red) and C-terminal (blue) domains.

served for all but 3 residues (49–51) of the 301 amino acid residues in the enzyme monomer. Side chain density was observed for all but nine lysines, five glutamic acids, two tyrosines, and one asparagine. These side chains are all solvent exposed and have been truncated in the final model.

The PNK monomer is made up of two distinct N- and C-terminal $\alpha/\beta$ domains comprising residues 1–148 and 158–301 (Figure 1A). These domains correspond to the kinase and phosphatase activities as described below. The tetramer is assembled through two separate interfaces (Figure 1B). One interface is formed exclusively with contacts between the N-terminal domains. This interface buries 964 Å$^2$ of accessible surface area per domain and is formed mostly through the packing of helices 2 and 3 with their symmetry mates. The second interface is primarily formed by contacts between C-terminal domains and buries 1328 Å$^2$ of exposed surface area per domain. This interface is anchored by strand 11, which interacts with its symmetry mate in an antiparallel fashion, creating a pseudocontinuous ten-stranded $\beta$ sheet between two C-terminal domains. Helix 7 packs against its symmetry mate to complete the interface. This interface is further augmented by additional interactions between helix 5 and its symmetry mate from the N-terminal domain. The overall dimensions of the enzyme tetramer are approximately 120 $\times$ 80 $\times$ 50 Å and there is a solvent-exposed channel through the middle of the tetramer 15–20 Å in diameter (Figure 1B).

The distance between kinase and phosphatase active sites within an individual enzyme subunit is approximately 35 Å and is similar to the closest distance between active sites from separate subunits. The sites within an individual subunit point in opposite directions and are physically separated by helices 5, 6, and 10 and strand 11, whereas the closest kinase and phosphatase active sites from separate subunits face each other across a surface valley.

### The Kinase Domain

The N-terminal domain is composed of the first 148 residues and has a core consisting of a four-stranded, parallel $\beta$ sheet with 4-1-3-2 topology (Figure 1A). The fold is completed by four $\alpha$ helices that are found on both sides of the central sheet and two additional helices following the last strand. A structural similarity search using the DALI server [34] with the N-terminal domain of PNK identifies a large structural kinase family that displays homology to PNK. There are 272 structures with a Z score $\geq$ 2.0, 25 structures with Z $\geq$ 5.0, and three structures with Z $>$ 10.0. These three structures, in order of improving Z score, are adenylate kinase (Z = 12.0, rmsd = 2.5 Å), chloramphenicol phosphotransferase (Z = 12.4, rmsd = 2.6 Å), and the N-terminal domain of 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (Z = 14.2, rmsd = 2.5 Å). A superposition of PNK, of 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (PFK/FBP), and adenylate kinase (ADK) reveals interesting similarities and differences between these related structures (Figure 2A). The N-terminal domain of PNK is the smallest of the three domains; PFK/FBP and ADK contain two and one extra strand(s) in their $\beta$ sheets, respectively. ADK contains a functionally important zinc knuckle subdomain inserted between helices 5 and 6 of PNK. In place of this subdomain, PFK/FBP and PNK have short 12 and 7 amino acid loops, respectively.

The kinase domain contains a bound ADP molecule in its active site, which represents the hydrolyzed product of ATP present in the crystallization buffer (Figure 2B). The kinase active site in the nucleotide-bound PNK complex contains a shallow tunnel, formed as a result of contacts between two enzyme surface loops that create a lid over the phosphate tail of the bound ADP. The contacts between these loops include residues E46 and R47 in the helix 2-3 loop and residues T128 and K129 in the helix 5-6 loop. The surface of the active site tunnel consists almost entirely of charged or polar residues, including R126, K129, K15, S16, R138, D35, T86, and N33 (Figure 3, middle). Several hydrophobic residues (P11, V131, V135, and M139) also help to form the active site walls. The bound ADP molecule is located on the side of the tunnel that opens into the channel in the middle of the enzyme tetramer (Figure 3A). Its binding pocket is formed mostly through electrostatic interactions between its diphosphate tail and a series of basic residues (K15, S16, and R126) and backbone amide groups from residue 12 to 17 (Figure 2B). The adenine ring stacks against the R122 side chain. The phosphate tail of the ADP is more deeply buried in the active site and more ordered than its associated ribose and adenosyl groups. The nucleotide binding pocket is well conserved in members of the ADK family and is formed by residues that are a part of the P loop motif. The rmsd for backbone atoms in an alignment of the ADK (residues 10–15) and PNK (residues 12–17) ATP binding sites is 0.2 Å and the ADP molecule bound in the active site of PNK is in the same conformation as AMPPNP bound to ADK (Figure 2B) [35]. This allowed us to easily model the position of the $\gamma$-phosphate into the PNK active site.

Based on the strong active site homology between the kinase domain of PNK and ADK and the structure of ADK with bound adenylate and AMPPNP [35], an adenosine 3′-monophosphate (3′-AMP) substrate molecule can be modeled into the active site on the other side of the tunnel from the bound ADP. This binding site faces the exterior of the enzyme tetramer (Figure 3B) and opens into a broad surface valley capable of accommodating larger polynucleotide substrates such as tRNA. A 3′-AMP model was placed into the PNK kinase active site based on the position of bound adenylate in ADK. The conformation of the bound 3′-AMP model was energy minimized to alleviate any steric clash with PNK while holding the 5′-hydroxyl position fixed. In the resulting model of the ternary complex, the 3′-AMP phosphate makes contacts to the backbone amide of T86 (2.8 Å) and the side chains of R38 (3.0 Å) and K15 (4.4 Å). R34 is 5.2 Å away from the 3′-phosphate oxygen, but has room to rotate toward the putative phosphate binding position. In the proposed binding mode, the adenosine ring fits nicely in the binding pocket but there are no base-specific contacts, the 3′-phosphate points out of the tunnel, and the 5′-hydroxyl points into the tunnel and is positioned correctly for an in-line phosphoryl transfer (3.3 Å from the modeled $\gamma$-phosphorus atom; Figure 2B).

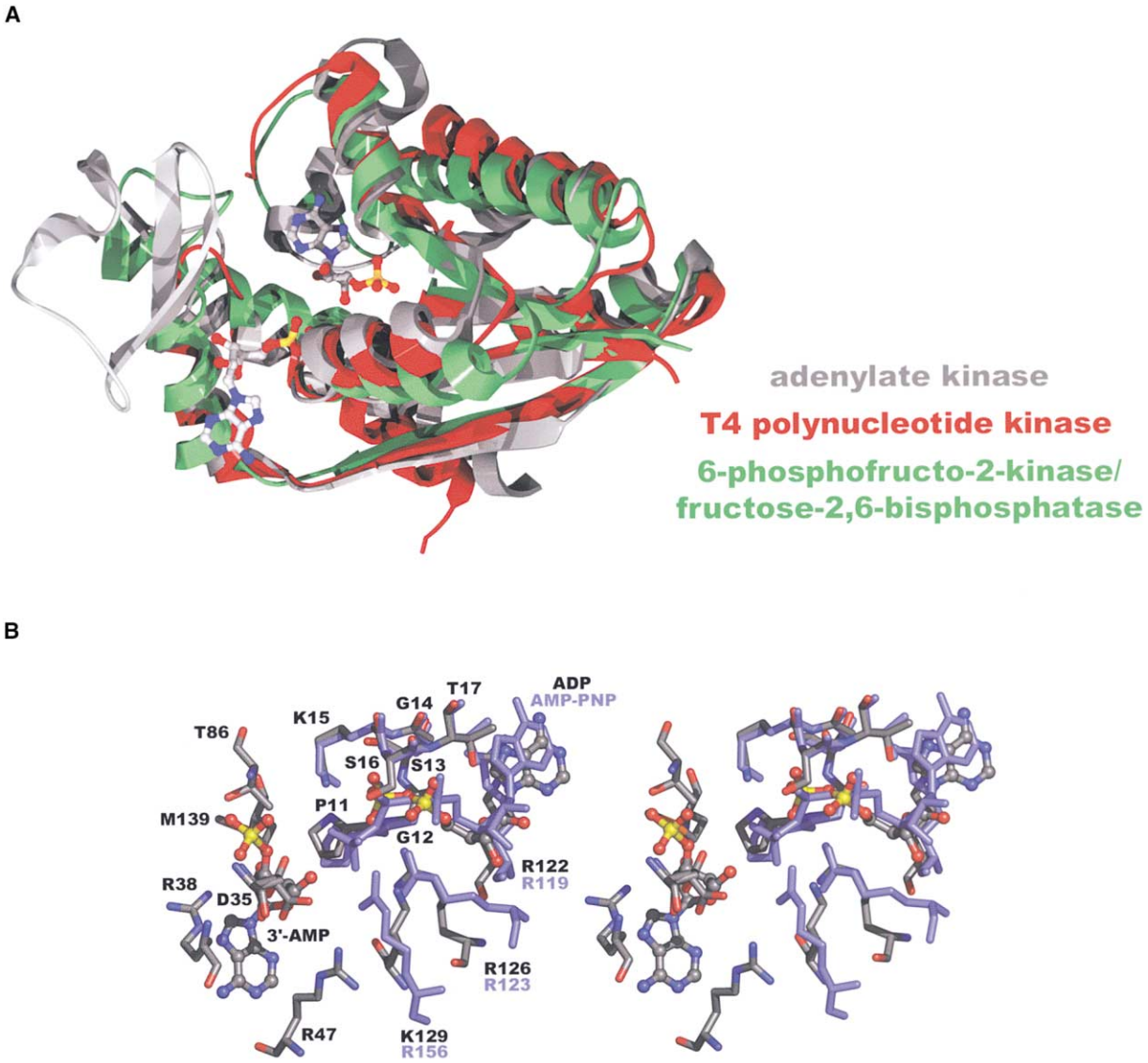



Figure 2. Superposition of the Kinase Domain and the Kinase Active Site

A superposition (DALI [34]) of T4 polynucleotide kinase (red), adenylate kinase (PDB ID code 1adk; gray), and the kinase domain of 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (1bif; green) with both the observed ADP ligand and modeled 3′-AMP ligand of PNK (A). The kinase active site of PNK is shown in a stick representation colored by atom type with residue labels in black (B). Both the observed ADP ligand and the modeled 3′-AMP ligand are displayed in ball and stick representations. On the ATP binding side, the corresponding residues from adenylate kinase along with its bound AMPPNP ligand (1ank) are shown colored in slate.

D35 is found on the 3′-AMP side of the tunnel and is observed in two conformations. In one conformation, it is engaged in a bidentate contact with R47 (3.4 and 4.0 Å) and in the other conformation it forms a salt bridge with R38 (3.6 Å). In either conformation, the D35 side chain is positioned close to the modeled position of the 3′-AMP 5′-hydroxyl (2.9–4.0 Å) and is a likely candidate for a magnesium binding side chain.

### The Phosphatase Domain

The C-terminal domain is composed of the last 143 residues and has a core consisting of a five-stranded, parallel β sheet with 11-10-5-8-9 topology (Figure 1A). The α/β fold is completed by four α helices inserted between consecutive β strands. There are two additional β strands that form a small two-stranded antiparallel sheet between strand 5 and helix 7. A structural similarity search using the DALI server indicates that the C-terminal domain of PNK belongs to the sizeable haloacid dehalogenase (HAD) structural family with 457 known structures with a Z score $\geq$ 2.0, 53 structures with Z $\geq$ 5.0, and 5 structures with Z $>$ 8.0. These five domains are phosphonoacetaldehyde hydrolase (Z = 9.8, rmsd = 2.6 Å), l-2-haloacid dehalogenase (Z = 9.5, rmsd = 3.2 Å), epoxide hydrolase (Z = 8.9, rmsd = 3.3), phosphoserine phosphatase (Z = 8.8, rmsd = 3.1 Å), and hypothetical phosphatase protein hi1679 (Z = 8.5, rmsd = 2.7 Å). As with the N-terminal domain, the

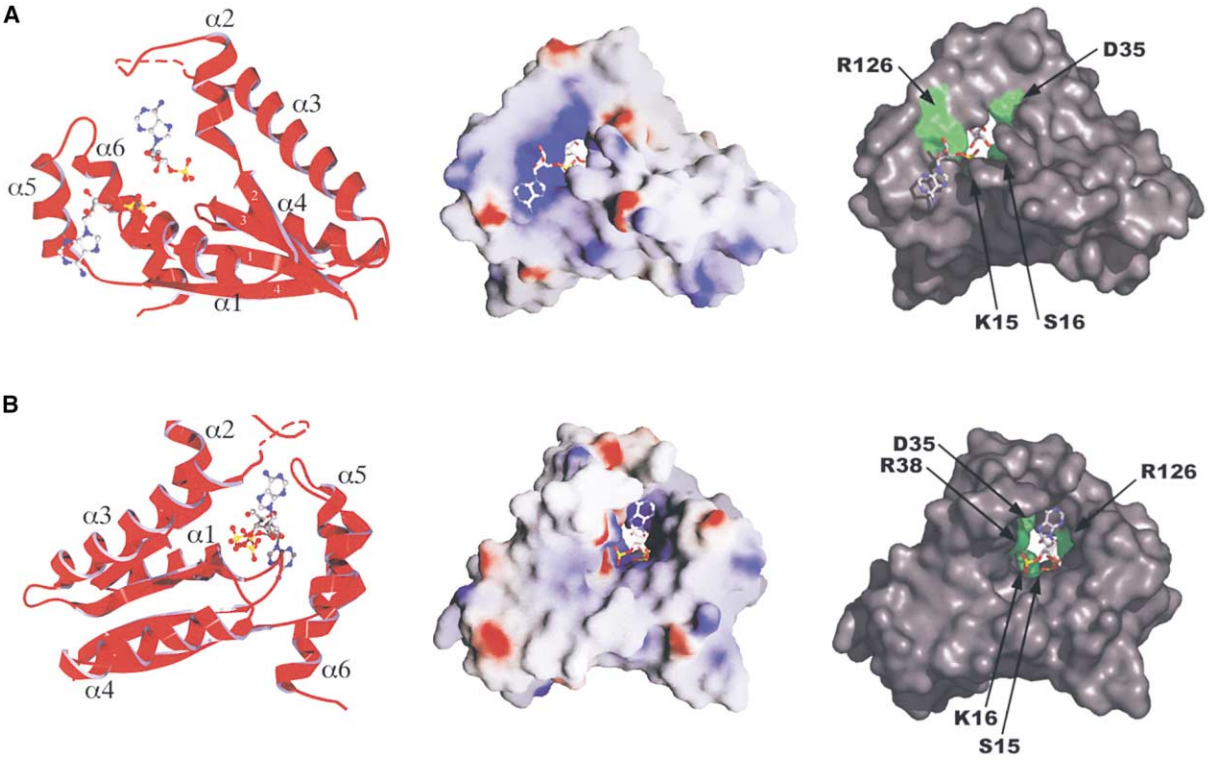



Figure 3. The N-Terminal Kinase Domain

The kinase domain of PNK is shown in two views of three representations each. (A) and (B) show two views that are related by a 180° rotation. The three representations are a ribbon diagram, a molecular surface color coded by electrostatic potential (GRASP), and a molecular surface color coded by mutational results. The ribbon diagram is labeled according to Figure 1 for reference. The GRASP surface shows a gradient of electrostatic potential ranging from −10 kT/e (red) to 15 kT/e (blue). The molecular surface is colored green and labeled where alanine mutations ablate kinase activity. All three illustrations contain the observed ADP ligand and the modeled 3′-AMP ligand in a ball and stick representation.

C-terminal domain of PNK is smaller than its closest structural neighbors. Both l-2-haloacid dehalogenase (HAD) and phosphoserine phosphatase (PSP) have additional helical subdomains that reduce the solvent accessibility of the active site cleft (Figure 4A). The absence of such a subdomain in PNK is likely responsible for the broader specificity of PNK compared to other members of this family. The smaller size of both PNK domains has probably evolved both to broaden the substrate specificity of the enzyme and maximize the information content of the phage genome for effective propagation. The HAD family of enzymes uses a common active site architecture defined by the D-X-D-X-T motif ($D_{165}$VDG$T_{169}$ in PNK). In all cases to date, these enzymes have been shown to proceed via a covalent phospho-aspartic acid intermediate and, with the exception of HAD, are dependent on magnesium for activity [36–38]. The phosphatase active site of PNK is centered on the primary aspartic acid in the conserved motif, D165, which is located on the C-terminal end of the central β strand (strand 5). A representation of the electrostatic potential of the molecular surface reveals that the active site cleft is acidic. Furthermore, residues that have been mapped by mutagenesis to be required for phosphatase activity cluster in the active site of the C-terminal domain (Figure 5, right).

The superposition of the active site of PNK with phosphoserine phosphatase inhibited by beryllium fluoride (a mimic of a phospho-aspartate intermediate; Figure 4B) [39] reveals key functional residues and their likely roles in catalysis. The δO1 of D165 is the nucleophilic atom responsible for attacking the 3′-phosphorus in the $S_N2$ phosphoryl transfer reaction. The ζN from K258 is 2.7 Å away from D165 δO1. D165 δO2 appears to be a metal ligand (2.1 Å) along with D278 (2.7 Å), the backbone carbonyl from D167 (2.1 Å), and a substrate phosphate oxygen (2.0 Å). Two water molecules would likely complete the octahedral coordination of the magnesium ion. The modeled, nonbridging phosphate oxygens are contacted by the backbone amides of G212 (2.7 Å), V166 (3.0 Å), and D167 (2.8 Å), the γOH of S211 (3.2 Å), and the ζN of K258 (2.9 Å). The second aspartic acid in the conserved active site motif (D167) forms a salt bridge with R213 (2.4 and 2.6 Å). D277 is 3.6 Å away from the modeled magnesium position and D254 is 4.5 Å away from a modeled phosphate oxygen atom. These residues correspond spatially to D171 and N170 in phosphoserine phosphatase, respectively, although they

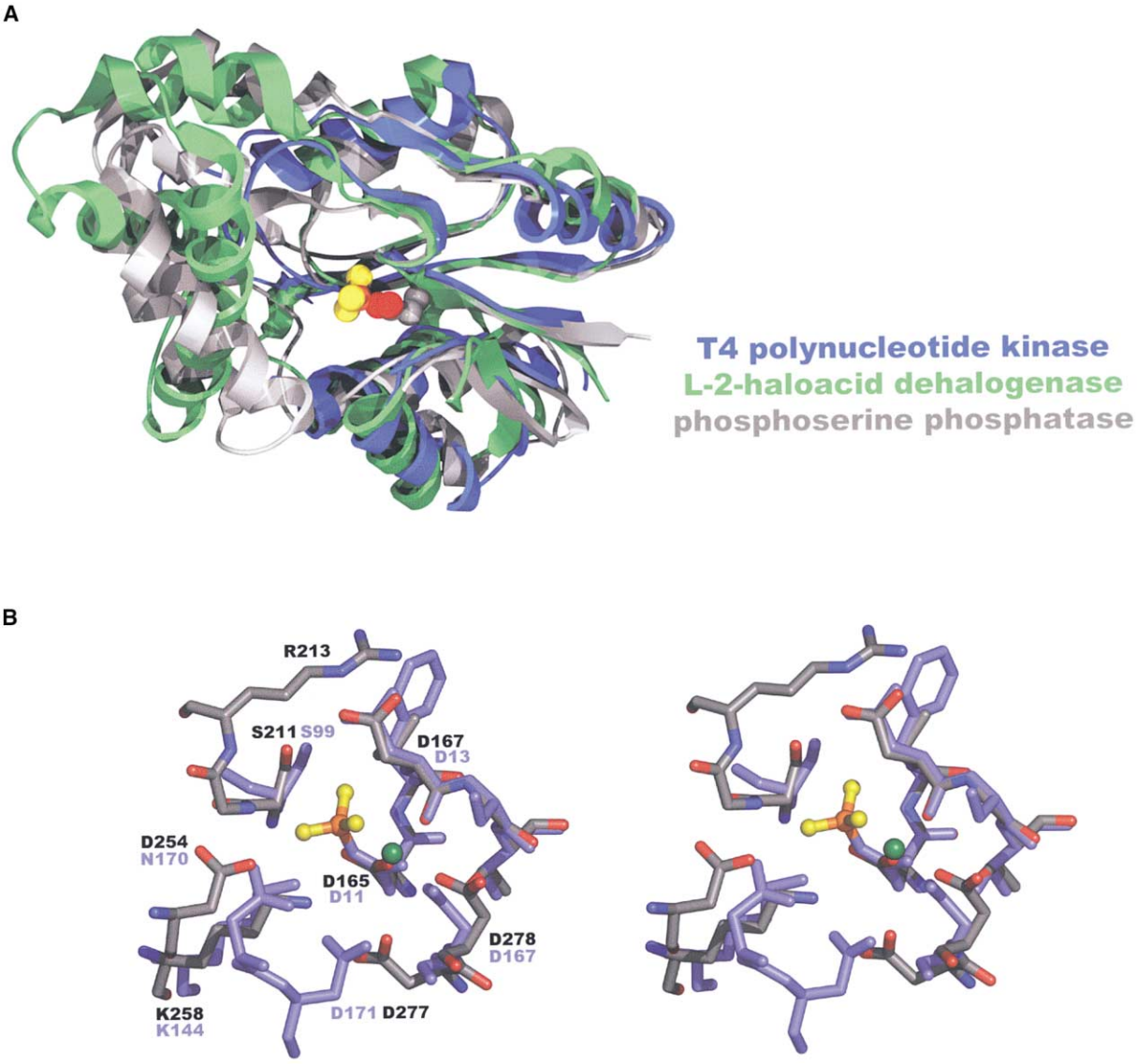



Figure 4. Superposition of the Phosphatase Domain and the Phosphatase Active Site

A superposition (DALI) of the phosphatase domain of T4 polynucleotide kinase (blue), phosphoserine phosphatase (1j97; gray), and l-2-haloacid dehalogenase (1qq5; green) with the beryllium fluoride aspartic acid ester from the phosphoserine phosphatase structure (A). The phosphatase active site of PNK is shown in a stick representation colored by atom type with residue labels in black (B). The beryllium fluoride moiety is displayed in a ball and stick representation (orange, beryllium; yellow, fluoride). Corresponding residues from phosphoserine phosphatase are displayed in slate.

come from different positions within the sequence and overall folds.

## Discussion

### T4 PNK as a Model for Eukaryotic Polynucleotide Kinases

Recently, a human polynucleotide kinase has been described [15–17, 40]. The structure of T4 PNK provides a model for the structure-function relationships of hPNK and other eukaryotic polynucleotide kinases within this enzyme family. Sequence alignments reveal that the eukaryotic enzymes contain both the kinase and phosphatase active site motifs found in T4 PNK [15]. Unlike T4 PNK, these enzymes are active against either DNA or RNA (but not both), and only the DNA kinases possess the additional 3′-phosphatase activity. Mammalian PNKs are believed to play a role in the repair of DNA lesions with 3′-phosphate and 5′-hydroxyl termini. Human polynucleotide kinase [16, 17] and a homolog from yeast [41] have been implicated in the response to DNA damage by oxidation, $\gamma$-radiation, and camptothecin. Other than the active site regions, the enzymes lack significant sequence homology to T4 PNK. The kinase and phosphatase domains are reversed compared to T4 PNK, with the P loop motif found in the C-terminal

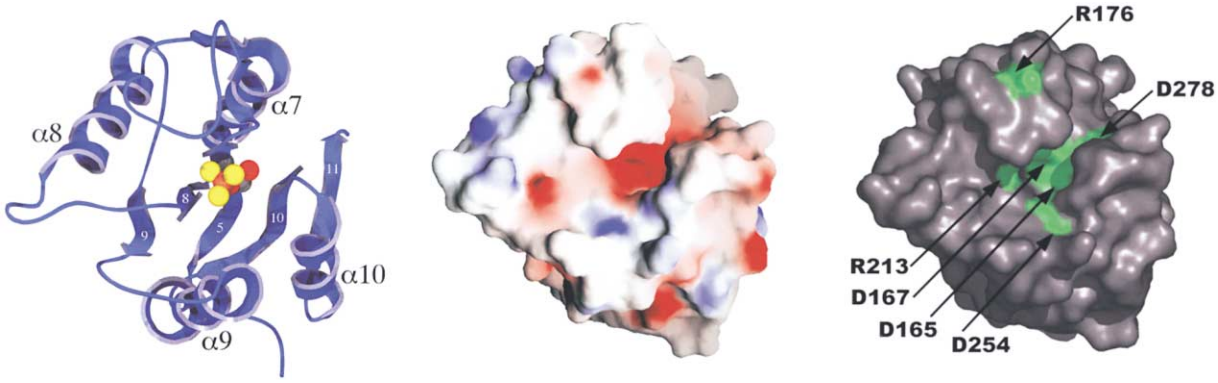



Figure 5. The C-Terminal Phosphatase Domain

The phosphatase domain of PNK is shown in three representations: a ribbon diagram, a molecular surface color coded by electrostatic potential (GRASP), and a molecular surface color coded by mutational results. The ribbon diagram is labeled according to Figure 1 for reference. The GRASP surface shows a gradient of electrostatic potential ranging from −15 kT/e (red) to 15 kT/e (blue). The molecular surface is colored green and labeled where alanine mutations ablate kinase activity. The ribbon diagram contains the modeled phospho-enzyme intermediate in a sphere representation to define the center of the phosphatase active site.

domain and the phosphatase motif found in the N-terminal domain. Furthermore, the eukaryotic enzymes are nearly twice as large (60 kDa) as T4 PNK and exist as monomers in solution [40]. However, it is likely that these enzymes will have active site structures similar to those in T4 PNK, and it will be interesting to compare the structures of the eukaryotic DNA PNKs to those of T4 PNK in terms of substrate binding modes, domain organization, and overall structure.

### Nicked tRNA Ends Can Access Kinase and Phosphatase Active Sites Simultaneously

The structure of PNK indicates that it is possible for nicked tRNA to bind in both the kinase and phosphatase active sites simultaneously (Figure 6). The tetrameric structure of PNK is crucial for this binding mode, as the polynucleotide binding side of the kinase active site from one monomer faces the phosphatase active site of a neighboring monomer. The distance (~35 Å) between these sites appears appropriate for binding the nicked ends of a tRNA anticodon loop. It is formally possible that one PNK tetramer could bind up to four tRNAs simultaneously as symmetry-related tRNA binding modes do not clash with each other. The anticodon stem loop consists of five ribonucleotides that do not base pair with each other or make other nonlocal interactions. The only known in vivo substrate for PNK is lysine tRNA that has been nicked directly 5′ of the anticodon nucleotides. Nicked tRNAs of this type should possess two flexible overhangs extending from the stable, base-paired stem structure. One overhang is two nucleotides long and has a 2′,3′-cyclic phosphate terminus while the other is three nucleotides long, contains the anticodon sequence, and has a 5′-hydroxyl terminus. The ~35 Å distance between active sites can be spanned by these flexible overhangs, suggesting that PNK performs both 3′- and 5′-terminal modifications within one tRNA binding event.

### The Oligomeric State and Catalytic Activity of Truncated Constructs

Much of the published biochemical data on the domain organization of T4 PNK can be interpreted based on the enzyme structure. The oligomerization state and bifunctionality of PNK has been probed previously by proteoly-

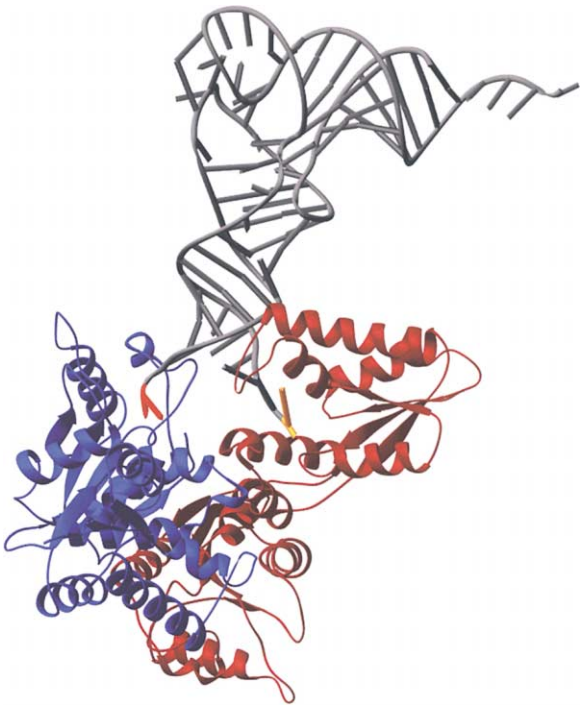


Figure 6. A Model of tRNA Binding

A hypothetical RNA/protein complex in which both PNK active sites are occupied simultaneously is shown with tRNA (gray) and two monomers of PNK (blue and red). The 5′-hydroxyl end of the nicked tRNA (yellow) is bound in the kinase active site of the red monomer while the 3′-phosphate end (orange) is bound in the phosphatase active site of the blue monomer.

sis and truncation experiments. PNK comprised only of residues 1–181 is a monomer with no phosphatase activity and low kinase activity; PNK consisting only of residues 42–301 is a dimer with no kinase activity and residual phosphatase activity [31]. These results indicate that although no allostery has been observed and the active sites are well separated, the tetrameric form of the enzyme is critical for the full activity of both domains. The N-terminal/N-terminal domain interface includes helices 2 and 3, which also form one side of the kinase active site tunnel. More specifically, $\alpha 2$ is the source of several residues that make contacts to the modeled 3′-AMP and have been implicated by mutagenesis experiments to be critical for kinase activity. The $\alpha 2/\alpha 2$ interface consists of two main contacts. Tyrosine 37 stacks against its symmetry mate and D36 makes contacts to both H44 (3.2 Å) and S40 (2.7 Å) of a symmetry-related monomer.

The removal of residues 1–41 (PNK [42–301]) removes a large amount of the N-terminal/N-terminal interface and prevents the tetramerization of the enzyme leading to C-terminal linked dimers in solution. However, the phosphatase activity is also significantly reduced despite its C-terminal location $\sim$30 Å away from the kinase active site. One likely possibility is that the remaining N-terminal domain sterically interferes with the phosphatase active site since it is connected to the C-terminal domain via what appears to be a flexible loop. This explanation is made even more plausible by the basic nature of the N-terminal domain and the acidic nature of the phosphatase active site. The removal of residues 182–301 (PNK [1–181]) eliminates nearly the entire C-terminal domain and thus we would predict an N-terminal linked dimer in solution. The observation of a monomer leads to the conclusion that the N-terminal domain interface is not as robust as the C-terminal one. This hierarchy in the strengths of the domain interfaces correlates well with surface area burial, as the C-terminal interface buries 37% more surface area than the N-terminal interface. As mentioned above, the kinase active site is closely associated with the N-terminal interface via helix 2. The lack of such an interface with the monomeric PNK (1–181) might explain the decrease in kinase activity despite the presence of all the critical active site residues.

### Substrate Binding Modes and Substrate Specificities of ADK Family Members

The substrate profiles of PNK and ADK, as well as additional kinase family members, can be explained by their compared structures. The bound ADP molecule observed in the active site of the kinase domain is bound in the same conformation as other P loop-bound nucleotides. The P loop is a phosphate binding motif with well-conserved sequence and structure [42]. It has the consensus sequence (A/G)-$X_4$-G-K-(G/S/T) and is known to occur in adenylate kinases, ras proteins, elongation factors, ATP synthase $\beta$ subunits, myosin heavy chains, thymidine kinases, and phosphoglycerate kinases. The observed binding mode of ADP confirms that, as in other P loop structures, it is recognized and bound mostly through interactions of the protein with the phosphate oxygens. This mode fits well with the observation that PNK is able to use any nucleotide triphosphate as the phosphoryl donor in the kinase reaction [3, 8]. Similarly, the modeled binding mode of 3′-AMP shows a lack of base-specific contacts, which is reasonable since PNK can phosphorylate a polynucleotide with any 5′-terminal nucleotide [8]. Furthermore, the 3′-phosphate position allows the modeled mononucleotide to be easily extended to a variety of polynucleotides.

The ATP and polynucleotide binding sites are on opposite sides of the active site tunnel (Figure 3). In the context of the tetramer, the tunnel leads from the central channel to the exterior. The ATP binding sites open into the channel while the polynucleotide binding site is on the exterior side of the tetramer. This allows large polynucleotide substrates to bind on the outside of the tetramer without encountering steric clashes from other PNK subunits or from other polynucleotides that might already be bound in the other active sites. Interestingly, molecular surfaces of adenylate kinase do not exhibit a tunnel. In its open form, ADK contains a binding cleft for both ATP and adenylate. However, as the enzyme binds substrate, the latch domain closes down on the substrate, completely cutting it off from solvent [35]. For ADK, the closing of the latch creates a specific binding pocket within the enzyme so that only the desired substrate molecule will bind productively. In contrast, the broad tunnel we observe for PNK allows the enzyme to have a very broad substrate specificity. The tunnel opens up to solvent and as such, polynucleotides of all types and sizes are able to bind and become phosphorylated. With other members of the family, such as PFK/FBP, remnants of a tunnel can be observed, but the tunnel openings are considerably narrower and more restrictive than in PNK. The tunnel/clamp structure has diverged within the PNK/ADK structural family along with the polynucleotide binding site. The selection of a particular tunnel or clamp backbone structure along with the specific residues that line the phosphoryl acceptor binding pocket appears to be an efficient way to generate a large variety of substrate specificities within the context of a single fold.

### Kinase Domain Active Site Structure Correlates with Mutational Analyses

Alanine scanning mutagenesis experiments on PNK have revealed a subset of residues that appear critical to the kinase activity (K15, S16, D35, R38, and R126) [32]. These five mutants all map to the tunnel region (Figure 3, right) and their specific roles in the catalytic mechanism of phosphoryl transfer can be described. K15 is one of the defining residues in the P loop motif and it is thought to be important for the structure of this motif because of its strict conservation [42]. It interacts with the $\beta$-phosphate of the bound ADP and with the modeled position of both the ATP $\gamma$-phosphate and the 3′-AMP 5′-hydroxyl. S16 is also a part of the P loop structure, as both the backbone amide of S16 and the side chain hydroxyl contact the $\beta$-phosphate of the bound ADP. Interestingly, this position in the motif is not as conserved as the preceding lysine. ADK has a glycine in this position and other P loop proteins have a threonine (myosin heavy chains and ATP synthase $\beta$ subunits). Fitting with the apparent role of S16, kinase activity is maintained by PNK with the S16T mutation [32]. R126 also contacts the diphosphate group of ADP,

is conserved in ADK (R123), and is a lysine (K172) in PFK/FBP. Therefore, K15A, S16A, and R126A mutations are deficient in kinase activity because they disrupt the ATP binding site of PNK. D35 is observed in two conformations as described above and is located on the AMP side of the tunnel. In one conformation it forms a salt bridge with R38, which contacts the 3′-phosphate of 3′-AMP in our model, so one possible role for D35 is that it simply helps to position R38 for phosphate binding. However, R38 is also positioned by a salt bridge with E47, and so a D35A mutation might not be expected to ablate kinase activity if this were its main function. D35 is unique in that it is the one acidic residue that is found near the active site of the kinase domain. Since the kinase activity is dependent on magnesium, a more likely function of D35 is metal binding. In its alternate conformation, it is well positioned for this role, as its delta oxygens are 3.7 and 4.0 Å away from the modeled position of the 5′-hydroxyl of 3′-AMP and pointing in the right direction to bind a metal ion that would bridge the gap between the 5′-hydroxyl and the γ-phosphate of ATP.

### Phosphatase Domain Active Site Structure Correlates with Mutational Analyses

Alanine scanning mutagenesis experiments on PNK have also been performed on the C-terminal phosphatase domain and have yielded a set of 6 residues that are critical for activity (D165, D167, R176, R213, D254, and D278) and a set of 2 residues that partially reduce activity (R246 and R279) [32]. D165 and D167 make up the core of the conserved active site motif D-X-D-X-T. Based on the structural alignment of PNK with the beryllium fluoride-bound phosphoserine phosphatase structure, D165 functions both as the nucleophile responsible for attacking the phosphorus atom in an $S_N2$-type reaction and as a magnesium ligand. Main chain atoms of D167 are involved in both phosphate (amide) and metal (carbonyl) binding, but the D167A mutant should still be able to function in that capacity. The side chain of D167 forms a salt bridge with R213 and the 2 residues are widely separated in sequence as D167 is found at the C-terminal end of strand 5 and R213 is located at the C-terminal end of strand 10. Therefore, the D167-R213 salt bridge appears to be important for stabilizing the structure of the active site. Strand 10 is a critical constituent of the active site structure since both S211 and G212 interact with the bound phosphate. Furthermore, since both D167A and R213A are defective in phosphatase activity, this suggests that the salt bridge between them plays a critical role in the formation of the active site. This also explains why the substitutions D167N and D167E are not active [32]. Asparagine is unable to form a strong electrostatic interaction with arginine and a while a glutamate residue can form a salt bridge with arginine, its extra length would perturb the positions of S211 and G212 relative to the rest of the active site. Similarly, the modified salt bridge that would exist between lysine, in an R213K mutant, and aspartic acid must not position the rest of strand 10 appropriately for phosphate binding, as this mutant is also inactive. D254 is positioned away from the phosphate binding site (4.5 Å) and forms a salt bridge with R279 (3.0 Å). Both of these residues may be important for polynucleotide binding as they are positioned in the entrance to the active site cleft and replacement of R279 with alanine partially reduces activity. The partially reduced activity of R246A involves another arginine residue that contacts D254 on the opposite side from R279 (3.0 and 3.2 Å) and is again likely to be important for substrate binding. D278 is located in close proximity to the modeled magnesium ion and would therefore be important for metal binding. The fact that D278E is partially active is consistent with this role. Out of the 6 identified residues, R176 is the furthest away from the phosphate binding center (11.0 Å) and does not contribute in any clear way to the active site.

The 2′,3′-cyclic phosphodiesterase activity must also reside in the phosphatase domain active site, but there is a dearth of data on this activity. Future experiments involving cocrystal structures with the various substrates of both domains along with bound magnesium ions should clarify the relationship of this activity to the 3′-phosphatase activity.

### Biological Implications

The structure of T4 PNK is the first structure of a bifunctional RNA repair enzyme. The enzyme has both 5′-kinase and 3′-phosphatase activities and, in conjunction with T4 RNA ligase, is able to repair lesions introduced in tRNA by a bacterial response to infection. The structure verifies that the kinase and phosphatase activities reside in structurally distinct domains that are members of the adenylate kinase family and the l-2-haloacid dehalogenase family, respectively. Both active sites have been located in the context of the enzyme tetramer and described via a combination of the observation of bound substrates and homology modeling. Truncation and mutational results have been mapped onto the structure and explained. Last, a tRNA binding mode has been proposed in which both 3′ and 5′ modifications occur in the context of a single binding event.

## Experimental Procedures

### Subcloning, Expression, and Purification

Subcloning was required to move the T4 PNK gene into an inducible vector for expression of selenomethionine-derivatized protein. The gene for T4 PNK was amplified from p.PNK plasmid (New England Biolabs; NEB) using the polymerase chain reaction (PCR) with a forward primer containing an NdeI cleavage site at the start codon (5′-GGTAGGTCATATGAAAAAGATTATTTTGACTATTG-3′) and a reverse primer containing a BamHI site just 3′ of the stop codon (5′-CGAAGCCCTCTAAAAATTCCTAGGAT-3′). pAII17 (a pET-based vector, W. Jack, NEB) and the PCR-generated fragment were digested with BamHI and NdeI and then ligated to generate the pET-PNK expression plasmid.

Selenomethionine-containing T4 PNK was expressed in minimal media from the *E. coli* strain BL21(DE3) adapted for growth with methionine pathway inhibition (Doublie, 1997). Cells were grown in minimal media at 37°C to an $OD_{600}$ of 0.8 and the following amino acids were added to inhibit the methionine biosynthetic pathway: 100 mg/L lysine, threonine, phenylalanine; 75 mg/L selenomethionine; 50 mg/L leucine, isoleucine, and valine. Following a 15 min incubation at 37°C, 0.5 mM isopropyl-thio-β-d-galactosidase (IPTG) was added to induce expression and the cultures were left at 37°C for 10–14 hr.

The cell pellets (25 g) were resuspended in 100 ml of phenyl column buffer A (0.2 M NaCl, 20 mM Tris-HCl, 0.1 mM EDTA, 10 mM β-mercaptoethanol [pH 7.5]) at room temperature. The resuspended cells were lysed by sonication with a sonicator cell disruptor (Heat Systems-Ultrasonics) with 2 min intervals at 70% power until the

soluble protein reached a plateau. The resulting lysate was centrifuged for 30 min at 12,000 RPM in a Beckman JA17 rotor. The crude cell supernatant was applied to an 80 ml phenyl Sepharose (lo sub) column which was washed with three column volumes (CV) of phenyl column buffer A. A gradient was then applied from 0% to 50% ethylene glycol (Fisher) in the same buffer, and 18 ml fractions were collected. The peak T4 PNK-containing fractions were determined by running samples on a 10%–20% SDS-PAGE gel stained with Coomassie blue, and were pooled (175 ml final volume). The phenyl Sepharose pool (175 ml) was diluted to a final volume of 1.0 L with heparin column buffer A (0.1 M NaCl, 20 mM Tris-HCl, 0.1 mM EDTA, 10 mM β-mercaptoethanol, 10% glycerol) and was applied to a 7.8 ml AF-Heparin 5PW column. The column was washed with three CV of the same buffer, and an NaCl gradient from 0.1 M to 1.0 M NaCl in ten CV of the same buffer was applied to elute the enzyme. The peak T4 PNK-containing fractions were determined by running samples on a 10%–20% SDS-PAGE gel stained with Coomassie blue, and were pooled (12 ml). The resulting pool was concentrated by dialysis into 0.3 M NaCl, 20 mM Tris-HCl, 0.1 mM EDTA, 1 mM DTT, 50% glycerol (pH 8.0) at 4°C. The concentrated AF-Heparin 5PW pool (~4 ml) was applied to a 318 ml HiPrep Sephacryl S-200 size exclusion column and was eluted with 0.3 M NaCl, 20 mM Tris-HCl, 10 mM β-mercaptoethanol, 0.1 mM EDTA, 10% glycerol (pH 8.0) at 4°C, and 4 ml fractions were collected. The peak T4 PNK-containing fractions were determined by running samples on a 10%–20% SDS-PAGE gel stained with Coomassie blue, and were pooled (28 ml). The HiPrep Sephacryl S-200 pool was dialyzed for 16 hr against 2.0 L of Source Q15 buffer A (0.1 M NaCl, 20 mM Tris-HCl, 10% glycerol, 0.1 mM EDTA, 10 mM β-mercaptoethanol [pH 8.5]) at 4°C and was applied to a 7.8 ml Source Q15 column. The column was washed with three CV of Source Q15 buffer A and was eluted with an NaCl gradient from 0.1 to 1.0 M in the same buffer. The peak T4 PNK-containing fractions were determined by running samples on a 10%–20% SDS-PAGE gel stained with Coomassie blue, and were pooled (34 ml). The final pool was dialyzed against storage buffer (50 mM NaCl, 10 mM Tris-HCl, 0.1 mM EDTA, 1 mM DTT, 50% glycerol, 0.1 mM ATP [pH 7.4]) at room temperature.

### Crystallography

Selenomethionine T4 PNK was dialyzed from storage conditions into 10 mM Tris (pH 7.4), 50 mM KCl, 1 mM DTT, 0.1 mM EDTA, 0.1 mM ATP. It was crystallized in hanging drops (2 μl of protein solution at 1 mg/ml with 2 μl of well solution). The well solutions ranged from 5% to 10% PEG 4000 and 0 to 5 mM MES (pH 6.5). The protein crystals grew within a week and were between 50 and 200 μm on a side. They were initially transferred to a cryo-solution of well solution plus 30% (w/v) sucrose in two steps of increasing sucrose and flash frozen in liquid nitrogen. With this treatment the crystals diffracted in a primitive orthorhombic space group (P222), with split spot profiles and diffracted to 2.8 Å with unit cell dimensions a = 77.6 Å, b = 97.3 Å, c = 121.5 Å. This data did not lead to a solution of the phase problem. Later, crystals were transferred from the crystallization drop into a cryo-solution of well solution plus 30% (v/v) dimethyl sulfoxide (DMSO) in four steps of increasing DMSO concentration and flash frozen. With the new cryo-condition, the crystals diffracted to 2.3 Å in a body-centered orthorhombic space group (I222) with unit cell dimensions a = 78.6 Å, b = 93.7 Å, c = 124.2 Å. A three-wavelength multiple anomalous dispersion (MAD) [43] data set was collected on beamline 5.0.2 at the ALS (Advanced Light Source, Lawrence Berkeley Laboratory) using a four-panel ADSC CCD area detector, but the crystal showed significant decay during data collection. Data were processed and scaled using HKL2000 [44]. Subsequent data analysis was performed using the CNS package [45]. Two selenium sites were located by manual inspection of the anomalous Patterson map and nine more were found through anomalous difference Fourier methods. Because of the crystal decay, phases were obtained with the set of 11 selenium sites using single anomalous dispersion (SAD) [46]. An interpretable electron density map was obtained after density modification with solvent flipping with a solvent content of 50% (CNS), and an initial model was built using XtalView [47]. The model was refined with CNS using the mlf target (maximum likelihood, f's) with 7% of the data excluded for the calculation of the crossvalidating free R [48]. Ninety percent of all the built residues are in the most favorable regions of Ramachandran space and 10% are in the allowed regions [49]. Statistics from phasing and refinement are shown in Table 1 along with typical examples of the experimental electron density map generated with XtalView and Raster 3D [50].

Ribbon diagrams were generated with SwissPDB Viewer [51] and POV-Ray. Molecular surface figures and active site figures were generated with PyMOL [52] and GRASP [53].

### Acknowledgments

We acknowledge the assistance of Roland Strong, Adrian Ferre-D'Amare, and the Stoddard laboratory in the structure determination. Funding was provided by NIH grant GM49857 (B.L.S.) and NIH training grant GM08268 (E.A.G.).

Received: May 22, 2002
Accepted: July 26, 2002

### References

1. Richardson, C.C. (1965). Phosphorylation of nucleic acid by an enzyme from T4 bacteriophage-infected *Escherichia coli*. Proc. Natl. Acad. Sci. USA *54*, 158–165.
2. Novogrodsky, A., and Hurwitz, J. (1966). The enzymatic phosphorylation of ribonucleic acid and deoxyribonucleic acid. I. Phosphorylation at 5′-hydroxyl termini. J. Biol. Chem. *241*, 2923–2932.
3. Novogrodsky, A., Tal, M., Traub, A., and Hurwitz, J. (1966). The enzymatic phosphorylation of ribonucleic acid and deoxyribonucleic acid. II. Further properties of the 5′-hydroxyl polynucleotide kinase. J. Biol. Chem. *241*, 2933–2943.
4. Berkner, K.L., and Folk, W.R. (1980). Polynucleotide kinase exchange as an assay for class II restriction endonucleases. Methods Enzymol. *65*, 28–36.
5. Chaconas, G., and van de Sande, J.H. (1980). 5′-32P labeling of RNA and DNA restriction fragments. Methods Enzymol. *65*, 75–85.
6. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Enzymes used in molecular cloning. In Molecular Cloning: A Laboratory Manual, Volume 1, Second Edition (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
7. Sirotkin, K., Cooley, W., Runnels, J., and Snyder, L.R. (1978). A role in true-late gene expression for the T4 bacteriophage 5′ polynucleotide kinase 3′ phosphatase. J. Mol. Biol. *123*, 221–233.
8. Richardson, C.C. (1981). Bacteriophage T4 polynucleotide kinase. In The Enzymes, P.D. Boyer, ed., Volume 14 (San Diego: Academic Press), pp. 299–314.
9. Cameron, V., and Uhlenbeck, O.C. (1977). 3′-Phosphatase activity in T4 polynucleotide kinase. Biochemistry *16*, 5120–5126.
10. 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.
11. 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.
12. Penner, M., Morad, I., Snyder, L., and Kaufmann, G. (1995). Phage T4-coded Stp: double-edged effector of coupled DNA and tRNA-restriction systems. J. Mol. Biol. *249*, 857–868.
13. Tyndall, C., Meister, J., and Bickle, T.A. (1994). The *Escherichia coli* prr region encodes a functional type IC DNA restriction system closely integrated with an anticodon nuclease gene. J. Mol. Biol. *237*, 266–274.
14. Levitz, R., Chapman, D., Amitsur, M., Green, R., Snyder, L., and Kaufmann, G. (1990). The optional *E. coli* prr locus encodes a latent form of phage T4-induced anticodon nuclease. EMBO J. *9*, 1383–1389.
15. Karimi-Busheri, F., Daly, G., Robins, P., Canas, B., Pappin, D.J., Sgouros, J., Miller, G.G., Fakhrai, H., Davis, E.M., Le Beau, M.M., et al. (1999). Molecular characterization of a human DNA kinase. J. Biol. Chem. *274*, 24187–24194.
16. Jilani, A., Ramotar, D., Slack, C., Ong, C., Yang, X.M., Scherer, S.W., and Lasko, D.D. (1999). Molecular cloning of the human gene, PNKP, encoding a polynucleotide kinase 3′-phosphatase

and evidence for its role in repair of DNA strand breaks caused by oxidative damage. J. Biol. Chem. *274*, 24176–24186.
17. 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.
18. Henner, W.D., Rodriguez, L.O., Hecht, S.M., and Haseltine, W.A. (1983). γ Ray induced deoxyribonucleic acid strand breaks. 3′ Glycolate termini. J. Biol. Chem. *258*, 711–713.
19. Torriglia, A., Perani, P., Brossas, J.Y., Chaudun, E., Treton, J., Courtois, Y., and Counis, M.F. (1998). L-DNase II, a molecule that links proteases and endonucleases in apoptosis, derives from the ubiquitous serpin leukocyte elastase inhibitor. Mol. Cell. Biol. *18*, 3612–3619.
20. Lown, J.W., and McLaughlin, L.W. (1979). Nitrosourea-induced DNA single-strand breaks. Biochem. Pharmacol. *28*, 1631–1638.
21. Yang, S.W., Burgin, A.B., Jr., Huizenga, B.N., Robertson, C.A., Yao, K.C., and Nash, H.A. (1996). A eukaryotic enzyme that can disjoin dead-end covalent complexes between DNA and type I topoisomerases. Proc. Natl. Acad. Sci. USA *93*, 11534–11539.
22. Lillehaug, J.R., Kleppe, R.K., and Kleppe, K. (1976). Phosphorylation of double-stranded DNAs by T4 polynucleotide kinase. Biochemistry *15*, 1858–1865.
23. 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.
24. Lillehaug, J.R., and Kleppe, K. (1975). Kinetics and specificity of T4 polynucleotide kinase. Biochemistry *14*, 1221–1225.
25. Jarvest, R.L., and Lowe, G. (1981). The stereochemical course of phosphoryl transfer catalysed by polynucleotide kinase (bacteriophage-T4-infected *Escherichia coli* B). Biochem. J. *199*, 273–276.
26. Lillehaug, J.R. (1978). Inhibition of T4 polynucleotide kinase by the ATP analog, β, γ-imidoadenylyl 5′-triphosphate. Biochim. Biophys. Acta *525*, 357–363.
27. van Houten, V., Denkers, F., van Dijk, M., van den Brekel, M., and Brakenhoff, R. (1998). Labeling efficiency of oligonucleotides by T4 polynucleotide kinase depends on 5′-nucleotide. Anal. Biochem. *265*, 386–389.
28. van de Sande, J.H., and Bilsker, M. (1973). Phosphorylation of N-protected deoxyoligonucleotides by T4 polynucleotide kinase. Biochemistry *12*, 5056–5062.
29. Fontanel, M.L., Bazin, H., and Teoule, R. (1994). Sterical recognition by T4 polynucleotide kinase of non-nucleosidic moieties 5′-attached to oligonucleotides. Nucleic Acids Res. *22*, 2022–2027.
30. Soltis, D.A., and Uhlenbeck, O.C. (1982). Independent locations of kinase and 3′-phosphatase activities on T4 polynucleotide kinase. J. Biol. Chem. *257*, 11340–11345.
31. Wang, L.K., and Shuman, S. (2001). Domain structure and mutational analysis of T4 polynucleotide kinase. J. Biol. Chem. *276*, 26868–26874.
32. Wang, L.K., and Shuman, S. (2002). Mutational analysis defines the 5′-kinase and 3′-phosphatase active sites of T4 polynucleotide kinase. Nucleic Acids Res. *30*, 1073–1080.
33. Hasemann, C.A., Istvan, E.S., Uyeda, K., and Deisenhofer, J. (1996). The crystal structure of the bifunctional enzyme 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase reveals distinct domain homologies. Structure *4*, 1017–1029.
34. Holm, L., and Sander, C. (1993). Protein structure comparison by alignment of distance matrices. J. Mol. Biol. *233*, 123–138.
35. Berry, M.B., Meador, B., Bilderback, T., Liang, P., Glaser, M., and Phillips, G.N., Jr. (1994). The closed conformation of a highly flexible protein: the structure of *E. coli* adenylate kinase with bound AMP and AMPPNP. Proteins *19*, 183–198.
36. Ridder, I.S., Rozeboom, H.J., Kalk, K.H., and Dijkstra, B.W. (1999). Crystal structures of intermediates in the dehalogenation of haloalkanoates by L-2-haloacid dehalogenase. J. Biol. Chem. *274*, 30672–30678.
37. Ridder, I.S., and Dijkstra, B.W. (1999). Identification of the Mg2+-binding site in the P-type ATPase and phosphatase members of the HAD (haloacid dehalogenase) superfamily by structural similarity to the response regulator protein CheY. Biochem. J. *339*, 223–226.
38. Collet, J.F., Stroobant, V., and Van Schaftingen, E. (1999). Mechanistic studies of phosphoserine phosphatase, an enzyme related to P-type ATPases. J. Biol. Chem. *274*, 33985–33990.
39. 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.
40. 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.
41. 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 γ-radiation or camptothecin in *Schizosaccharomyces pombe*. J. Biol. Chem. *277*, 4050–4055.
42. Saraste, M., Sibbald, P.R., and Wittinghofer, A. (1990). The P-loop—a common motif in ATP- and GTP-binding proteins. Trends Biochem. Sci. *15*, 430–434.
43. Hendrickson, W.A. (1991). Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. Science *254*, 51–58.
44. Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.
45. Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.
46. Brodersen, D.E., de La Fortelle, E., Vonrhein, C., Bricogne, G., Nyborg, J., and Kjeldgaard, M. (2000). Applications of single-wavelength anomalous dispersion at high and atomic resolution. Acta Crystallogr. D Biol. Crystallogr. *56*, 431–441.
47. McRee, D.E. (1999). XtalView/Xfit—a versatile program for manipulating atomic coordinates and electron density. J. Struct. Biol. *125*, 156–165.
48. Kleywegt, G.J., and Brunger, A.T. (1996). Checking your imagination: applications of the free R value. Structure *4*, 897–904.
49. Laskowski, R.J., Macarthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK—a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.
50. Merritt, E.A., and Bacon, D.J. (1997). Raster3D: photorealistic molecular graphics. Methods Enzymol. *277*, 493–505.
51. Guex, N., and Peitsch, M.C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis *18*, 2714–2723.
52. DeLano, W.L. (2002). The PyMOL Molecular Graphics System (San Carlos, CA: DeLano Scientific).
53. Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins *11*, 281–296.
54. 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.

## Accession Numbers

The structure has been deposited in the Protein Data Bank under ID code 1LTQ.

## Note Added in Proof

A paper describing the structure of the isolated kinase domain of T4 PNK was published while this manuscript was in review [54]. The results regarding the overall structure of the kinase domain and active site are quite similar to the results described above.