Uncoupling of 3’ phosphatase and 5’ kinase functions in budding yeast:

Characterization of *S. cerevisiae* DNA 3’ phosphatase (*TPPI*)

John R. Vance and Thomas E. Wilson*

*To whom correspondence should be addressed.

Department of Pathology
University of Michigan Medical School
1301 Catherine Rd.
M4214 Med Sci I, Box 0602
Ann Arbor, Michigan 48109-0602

Phone: (734) 936-1887
FAX: (734) 763-6476
Email: wilsonte@umich.edu

Running Title: *S. cerevisiae* DNA 3’ phosphatase
Summary

Polynucleotide kinase is a bifunctional enzyme containing both DNA 3’ phosphatase and 5’ kinase activities seemingly suited to the coupled repair of single-strand nicks in which the phosphate has remained with the 3’ base. We show that the yeast *Saccharomyces cerevisiae* is able to repair transformed dephosphorylated linear plasmids by nonhomologous end-joining with considerable efficiency independently of the end-processing polymerase Pol4p. Homology searches and biochemical assays did not reveal a 5’ kinase that would account for this repair, however. Instead, open reading frame YMR156C (here named *TPP1*) is shown to encode only a polynucleotide kinase-type 3’ phosphatase. Tpp1p bears extensive similarity to the ancient L-2 haloacid dehalogenase and DDDD phosphohydrolase superfamilies, but is specific for double-stranded DNA. It is present in high levels in cell extracts in a functional form and so does not represent a pseudogene. Moreover, the phosphatase-only nature of this gene is shared by *Saccharomyces mikatae* YMR156C and *Arabidopsis thaliana* K15M2.3. Repair of 3’ phosphate and 5’ hydroxyl lesions is thus uncoupled in budding yeast as compared to metazoans. Repair of transformed dephosphorylated plasmids, and 5’ hydroxyl blocking lesions more generally, likely proceeds by a cycle of base removal and resynthesis.
Introduction

Oxidative damage to DNA can result from endogenously generated reactive oxygen species, or due to exposure to exogenous agents such as ionizing radiation or anticancer agents such as bleomycin and neocarzinostatin (reviewed in 1). Such damage and the enzymes involved in its repair frequently produce fragmentation of the deoxyribose sugar backbone resulting in DNA strand breaks bearing abnormal structures at the 3’ and 5’ termini. These are termed “blocking lesions” because they prevent the reactions necessary to achieve final repair of the damaged strand, namely polymerization and ligation. Since both single and double-strand lesions can occur with potential consequences including replication failure and genomic rearrangement, the resolution of blocking lesions is of major importance in genome maintenance.

Although many chemical forms are possible, important blocking lesions on 5’ termini include hydroxyls and deoxyribose phosphates (dRP)\(^1\). As an example of the redundancies in end processing, dRP moieties can be removed in short patch base excision repair (BER) by the lyase function of DNA polymerase β (2,3) or certain glycosylases (4,5), or in long patch BER by flap excision and resynthesis (6). The extensive 5’ resection that occurs in the first steps of recombination also likely removes blocking lesions at double-strand breaks (7,8). Common blocking lesions on 3’ termini include phosphates, α,β-unsaturated aldehydes resulting from β-elimination reactions (9), and phosphoglycolate moieties that are the primary product of bleomycin action (10). The most potent 3’ processing enzymes are the apurinic-apyrimidinic (AP) endonucleases, which, in addition to cleaving strands at abasic sites, possess 3’ diesterase activities capable of removing most nucleotide fragments (11-13). It is again possible that 3’ lesions might be resolved by a more extensive degradation during recombination, for example by recBCD (8).
Polynucleotide kinase (PNK) is best known due to the utility of the T4 enzyme (14) in molecular cloning, but it was demonstrated to exist in eukaryotes thirty years ago (15). Although not clearly indicated by their name, PNK proteins studied to date bear two distinct catalytic activities, a 5’ kinase and a 3’ phosphatase (16,17). Although the precise biological role of eukaryotic PNK remains to be determined, it is clearly suited to directly reverse the two reciprocal blocking lesions that would result from a strand break with a misplaced phosphate, i.e. a 3’ phosphate and 5’ hydroxyl. Indeed, the preferred substrate of mammalian PNK is a DNA nick (18), and its enzymatic activities are stimulated by interaction with the XRCC1 repair protein (19), strongly suggesting a role in BER/single-strand break repair.

As a continuation of our interest in delineating the mechanisms by which terminal damage is resolved during DNA double-strand break repair we have been attempting to identify and characterize PNK from the yeast *S. cerevisiae*, whose existence we inferred from the successful repair of transformed dephosphorylated linear plasmids. The recent cloning of human PNK (hPNKP, ref. 16,17) has greatly facilitated this by allowing homology searching against the yeast and other sequenced genomes. Surprisingly, we find that *S. cerevisiae*, at least one other *Saccharomyces* yeast, and species as distantly related as *Arabidopsis thaliana*, contain a gene with homology to only the putative 3’ phosphatase portion of hPNKP. The *S. cerevisiae* protein, encoded by open reading frame (ORF) YMR156C, here named *TPP1*, shares many of the biochemical properties of the hPNKP 3’ phosphatase but indeed is not a 5’ kinase. Despite the observed plasmid repair, structural comparisons and enzymatic assays failed to detect an unlinked 5’ kinase. Evolutionary models to explain these results are discussed in the context of alternative pathways for resolution of terminal damage during DNA repair.
Experimental Procedures

Yeast Strains. Wild-type S. mikatae strain #1815, obtained from Dr. Mark Johnston, is described in reference (20). S. pombe strain FY254 (hi, leu1-32, ura4-048, ade6, can1-1) was obtained from Dr. Dennis Thiele. S. cerevisiae strains used in the plasmid assay were the wild-type strain YW389 (MATα, ade2Δ0, his3Δ200, leu2, lys2-801, trp1Δ63, ura3Δ0) and its isogenic derivatives YW513 (dnl4-K282R) and YW514 (pol4-D367E). The chromosomal point mutations in these strains were constructed by the de novo mutation strategy described by Erdeniz et al. (21). Strains were verified by diagnostic PCR, allele sequencing, and by documenting their deficiencies in previously described DNL4 and POL4-dependent NHEJ assays (22,23). S. cerevisiae strains used for biochemical assays were the wild-type strain YW465 and its isogenic derivatives YW573 (tpp1Δ::MET15), YW605 (apn1Δ::HIS3), and YW619 (tpp1Δ::MET15, apn1Δ::HIS3). The deletions in these strains were constructed by PCR-mediated gene replacement and verified by PCR as described (24). Single mutants were constructed directly in YW465, and YW619 by disruption of APNI in YW573.

Plasmid transformation assay. Plasmid pES26 has been previously described (23). Methods of plasmid preparation and transformation were exactly as described (22), but with the following addition. After Bgl II digestion, but before extraction and precipitation, an equal volume of 2X CIP buffer (Roche) was added, followed by nothing (ligation-competent control plasmid) or 0.16 units CIP per µg plasmid (dephosphorylated plasmid) and further incubation at 37 °C for 30 min.

Multiple sequence alignment. BLASTP and Psi-BLAST homology searches were performed via the NCBI web server at http://ncbi.nlm.nih.gov/. Sequences included in the overall alignment were all hPNKP BLASTP hits from the non-redundant GenBank, EST and
STS databases with an E value < 0.01, in addition to Trl1p, T4 PNK and AcNPV-2. EST and STS sequences were assembled into contigs and translated prior to inclusion in the alignment; details are available on request. Accession numbers for the sequences that are ostensibly complete but uncharacterized are (see Table 1 for species/gene abbreviations): Mm, AAF36487; Dm (CG9601), AAF54229; At-1, BAA97052; At-2, CAB81914; Sp (C23C11.04C), CAB11157; SeNPV (ORF54), AAF33584; AcNPV-1 (Ac-HisP), AAA66663; AcNPV-2 (Ac-PNK/PNL), AAA66716. In the case of Caenorhabditis elegans (Ce, accession T21197), we extended the putative ORF F21D5.5 by appending both amino and carboxyl-terminal exon translations that were not originally included. Alignments were performed with MACAW (25).

**S. mikatae YMR156C sequence.** Primers were designed that corresponded to S. mikatae sequences (kindly provided by Dr. Mark Johnston) homologous to S. cerevisiae ORFs YMR154C and YMR157C (5’-TCCAGTTCAAAAGTAGGATTCC and 5’-TAGGTAAGGCCGACATCATC, respectively). These were used in a PCR reaction with S. mikatae genomic DNA, using the HF Advantage PCR kit according to the manufacturer’s instructions (Clontech). The resulting single ~5 kb amplified fragment was sequenced directly by the University of Michigan DNA Sequencing Core by walking from YMR157C through YMR156C and into YMR155W. The entire S. mikatae YMR156C coding sequence (accession # AF326782) was read without ambiguities, including a stop codon read clearly from two independent runs.

**Extraction and purification of proteins from yeast.** Yeast strains expressing GST fusion proteins were isolated from the ORF array described and kindly provided (via Dr. Dennis Thiele) by Dr. Eric Phizicky and colleagues (26). Yeast from the YMR156C well were streaked to single colonies, and anti-GST (Santa Cruz) Western blotting was used to identify isolates.
expressing proteins whose size corresponded to GST-Tpp1p and non-recombinant GST.
Purification was by glass bead lysis and salt extraction followed by batch chromatography on glutathione agarose (Amersham Pharmacia Biotech) as described (26) except that protein expression was induced by adding 0.1 mM CuSO4 for 3 hr prior to harvest. Typical final GST-Tpp1p dialysates derived from 50 ml of yeast culture contained 50 µg/ml fusion protein in 600 µl of 20 mM Tris-HCl (pH 7.5), 2 mM EDTA, 4 mM MgCl2, 1 mM DTT, 50 mM NaCl, and 50% (v/v) glycerol.

Crude whole-cell extracts of *S. cerevisiae*, *S. mikatae* and *S. pombe* were all prepared by glass bead disruption of cells in approximately one cell pellet volume of 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 M NaCl, 10 mM MgCl2, 1 mM DTT, 10% glycerol, 2 µg/ml aprotinin, 1 µg/ml each of leupeptin and pepstatin, and 1mM PMSF followed by centrifugation to remove cellular debris. Final extracts were diluted to 0.5 µg/µl protein.

**Enzyme activity assays.** Oligonucleotides with and without 3’ phosphates were purchased from Operon (see Figs. 5 and 7 for sequences). Oligonucleotides were 5’ end labeled with [γ-32P] ATP using 3’ phosphatase-free polynucleotide kinase (Roche). Final substrates were prepared by annealing labeled oligonucleotides to a 2-fold molar excess of the required unlabeled strands by heating to 90 °C followed by slow cooling. Standard assays of 3’ phosphatase activity contained 50 fmol DNA substrate and 10 fmol GST-Tpp1p or 1 µg crude cellular protein in a reaction volume of 10 µl such that the final buffer was 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 10 mM MgCl2, 1 mM DTT, 50 µg/ml BSA. Kinase assays were similar except used 1 pmol GST-Tpp1p, were 25 mM NaCl, and also included 100 units T4 DNA ligase (NEB), 1 mM ATP, and, where indicated, 5 units T4 PNK (NEB) as an internal control. After
incubation at 30 °C for 10 min, formamide/EDTA loading buffer was added and samples were
electrophoresed on 7 M urea, 12 % polyacrylamide gels followed by autoradiography.
**Results**

**Repair of 5’ hydroxyl blocking lesions during yeast NHEJ.** To begin to determine the impact of 5’ hydroxyl blocking lesions on the repair of double-strand breaks, plasmid pES26 was digested *in vitro* with the restriction enzyme Bgl II and transformed in yeast cells, both with and without pretreatment with calf-intestinal alkaline phosphatase (CIP). In this well-established assay (22,23), recircularization by nonhomologous end-joining (NHEJ, ref. 7,23,27) is required for plasmid stability, and thus for expression of the plasmid URA3 marker gene. Because the Bgl II site resides in an essential region of the ADE2 marker gene also on this plasmid, transformation to Ade+ further requires that repair be precise (imprecise repair yields ade2 colonies that appear red instead of white). Dephosphorylated plasmid transformed into wild-type yeast showed only a 2 to 3-fold decrease in Ura+ colony recovery as compared to ligation-competent plasmid (Fig. 1). Moreover, there was no increase in red colony recovery, indicating that joining remained precise. Transformation by dephosphorylated plasmids was decreased 54-fold in yeast deficient in the ligase required for NHEJ (dnl4-K282R, ref. 23), verifying that the damaged ends are not routed into another pathway. As an additional control, it was verified that CIP-treated plasmids could not be religated *in vitro* by T4 DNA ligase unless further treated with T4 PNK, as indicated by gel electrophoresis and a >700-fold decrease in colony recovery following bacterial transformation (not shown).

As shown in Fig. 1A, there are at least two ways that NHEJ of 5’ hydroxyl lesions might be achieved. In the first, the damaged 5’ nucleotide is nucleolytically removed and subsequently resynthesized on the opposite side of the break. We have previously observed that the yeast PolX family polymerase Pol4p can catalyze such base addition during NHEJ (22). pol4-D367E mutant yeast, which express catalytically inactive Pol4p (22), showed the same pattern as wild-
type, however. Importantly, some polymerization-dependent NHEJ events show only a 2-fold defect in pol4 mutants (22), so it possible that another polymerase can substitute for Pol4p at dephosphorylated lesions as well. Nonetheless, these observations were consistent with the alternative model in which the 5' hydroxyl lesion is directly reversed by a PNK 5' kinase. We sought to identify this in the experiments described below.

**Conservation and modular evolution of 5’ kinase and 3’ phosphatase PNK domains.**

BLAST searches against the non-redundant GenBank, EST and STS databases using hPNKP as the query revealed a set of 22 distinct PNK-like genes from 20 different species and viruses (Table 1). Among these was the *S. cerevisiae* ORF YMR156C, but this weaker hit (E = 0.006) corresponded to only a portion of the human protein and surprisingly lacked an apparent Walker A (i.e. P-loop) motif for ATP binding (28), which would be expected for a 5' kinase. Absent as hits were three genes identified from literature searching that are known or believed to encode 5’ kinases: *TRL1* tRNA ligase from *S. cerevisiae* (29,30), bacteriophage T4 PNK (14), and a putative PNK from *Autographa californica* nucleopolyhedrovirus (31).

To examine the relationship between YMR156C and the other PNK-related genes in more detail, we performed an extensive multiple sequence alignment as shown in Figs. 2-4. Two sequence motifs, the Walker A box and the phosphotransferase motif DXDX(T/V) (32), were used as a means of unambiguously identifying the 5’ kinase and 3’ phosphatase catalytic cores, respectively, as suggested by previous alignments of smaller numbers of bifunctional PNK sequences (16,17). It was first apparent that genes may contain just the 3’ phosphatase domain, just the 5’ kinase domain, or both. This was not limited to *S. cerevisiae*. For example, *Arabidopsis thaliana* contains two hypothetical genes, one encoding a 3’ phosphatase (designated for simplicity as At-1; see Experimental Procedures for accession numbers) and one
a 5’ kinase (At-2). These genes are on different chromosomes, and examination of the genomic sequence surrounding At-1 did not reveal any cryptic 5’ kinase domain exons. Second, in some species there is an apparent redundancy of function. For example, Dictyostelium discoideum has two distinct putative 5’ kinase genes. We note, however, that these genes correspond to two structurally evident 5’ kinase groupings (see below), so it is possible, if not likely, that 5’ kinases perform distinct tasks in the cell. Finally, in genes containing both domains, each of the two possible orders are observed.

Taken together, these observations reveal a tremendous modularity in the evolution of the PNK 3’ phosphatase and 5’ kinase catalytic domains. As a result, it is now possible to assign clear borders to these domains as independent units. The extensive overall conservation can be viewed as five distinct regions of homology to the hPNKP sequence.

Regions 1 and 2 correspond to the 3’ phosphatase domain (Fig. 3), best understood by comparison to the bacterial histidinol phosphatase domain, a homology which was revealed by a more sensitive iterative Psi-BLAST search. This domain catalyzes one step in histidine biosynthesis and is typified by one portion of the bifunctional E. coli hisB (HIS7) protein. Most broadly, hisB-type histidinol phosphatases, like hPNKP (16,17), are part of a large superfamily of proteins containing the L-2 haloacid dehalogenase (HAD) fold (33). These proteins catalyze a wide variety of hydrolytic reactions via a covalent substrate-enzyme intermediate. More restrictively, hisB is a member of the “DDDD” superfamily of phosphohydrolases, so named due to the presence of four invariant aspartates (34). We have labeled the overlapping amino acid motifs that define these domains as DDDD motifs 1-4 in Fig. 3A. Their presence in all putative PNK-related 3’ phosphatases, including YMR156C, very strongly indicates that these proteins form a covalent bond with their substrate via the first aspartate of DDDD motif 1.
Two additional conserved non-DDDD motifs were evident in the putative eukaryotic 3’ phosphatases (called motifs A and B, Fig. 3A). Motif B, the more highly conserved, has a consensus SX_2DX_2FAX_6FXTPEX_2F. It is entirely absent from hisB, but present in YMR156C, and likely serves in determining substrate recognition. Motifs A and B, as well as other characteristic amino acids in the DDDD motifs (Fig. 3), thus appear to provide a signature for identifying 3’ phosphatases. For example, AcNPV-1 has been called Ac-HisP because the presence of hisB motifs suggested that it was a histidinol phosphatase (see GenBank accession AAA66663). This protein in fact bears the more extended homology typical of 3’ phosphatases. Also, our alignment in motif 3, in contrast to that previously described (35), contains the high-scoring 3’ phosphatase sequence RX_5MW in T4 PNK, even though this places a threonine at the first X instead of the nearly invariant HAD-domain lysine.

Regions 3-5 correspond to the 5’ kinase domain (Fig. 4). Region 3, which includes the Walker A motif, is strikingly conserved at non-Walker positions from T4 to human, as previously described (16,17). This allowed T4 PNK, but not any S. cerevisiae kinases, to be returned by the Psi-BLAST search. Region 4 is primarily responsible for delineating what appears to be two subtypes of eukaryotic 5’ kinases. Most notably, genes Dd-2 and At-2 possess the tetrapeptide DRCN, compared to the conserved DNTN sequence in the other eukaryotic kinases. The similarity of the former to the DRNN sequence in yeast Trl1p might suggest that these proteins are involved in metabolism of RNA rather than DNA. Region 5 contains a very highly conserved EGF tripeptide specific to the eukaryotic proteins.

Subsets of PNK-related proteins showed additional homology outside of Regions 1-5 to each other or to non-PNK proteins, further extending the observed modularity. Of particular interest is a weak homology of At-1 to both poly(ADP-ribose) polymerase (PARP) and DNA
ligase III (LIG3). Both of these proteins interact with XRCC1 and assist in single-strand break repair/BER (36), suggesting a similar role for At-1 and this family of proteins more generally. Indeed, hPNKP has recently been shown to interact structurally and functionally with XRCC1 (19).

Saccharomyces mikatae YMR156C lacks a fused 5’ kinase domain. Based on the above, we hypothesized that YMR156C would be functional as a 3’ phosphatase despite lacking the 5’ kinase domain. Nonetheless, it was essential to exclude the alternative possibility that YMR156C is simply a rearrangement artifact of the sequenced S. cerevisiae strain S288C, or perhaps of laboratory strains more generally. We therefore sequenced the orthologous gene from the wild-type yeast species Saccharomyces mikatae, strain #1815, a member of the S. cerevisiae sensu stricto group (20). The S. mikatae locus exhibited synteny with S. cerevisiae chromosome 13 that extended to the centromere–proximal and distal ORFs. The encoded 245 amino acid protein was 75% identical to S. cerevisiae YMR156C and also possessed the six DDDD and 3’ phosphatase motifs described above, with the non-conservative substitutions all clustered in the transitions between these motifs (not shown). Most importantly, the S. mikatae ORF ended immediately after 3’ phosphatase motif B, so this species also clearly lacks a fused kinase domain. It is thus highly likely that YMR156C evolved to support a specific function in budding yeast despite lacking the 5’ kinase domain.

YMR156C encodes a double-stranded DNA 3’ phosphatase. Although YMR156C possesses all of the putative 3’ phosphatase motifs, it is the most diverged among the eukaryotic species as determined by calculation of phylogenetic distances (not shown). As a result, alternative phosphatase-related functions could not be ruled out based on sequence alone. Indeed, prior to this work the only described function for YMR156C was that forced
overexpression reduced cellular sensitivity to ketoconazole, a non-phosphorylated antifungal agent that acts by inhibiting ergosterol biosynthesis (37). In order to directly assess its biochemical properties, YMR156C was overexpressed in yeast as a GST fusion protein and purified. As shown in Fig. 5B (lane 2), the purified GST-YMR156C fraction contained no other bands visible in Coomassie-stained gels as candidate 5’ kinase partners. The protein catalyzed the efficient removal of a 3’ phosphate (Fig. 5C, lane 3) from a nicked oligonucleotide substrate (Figure 5A, similar to ref. 38). Importantly, there was no loss of the 5’ 32P-label, thus demonstrating the 3’ specificity of the phosphatase. As a control, no 3’ phosphatase activity was detected when non-recombinant GST was purified in parallel and similarly tested (lane 4).

As expected from the sequence alignment, repeated examination of the GST-YMR156C fusion protein did not reveal an associated 5’ kinase activity (not shown and Fig. 7, lane 3, below). Thus, we conclude that YMR156C acts exclusively as a 3’ phosphatase and have designated the gene \textit{TPP1} (three prime phosphatase 1). These results further indicate that there is not a 5’ kinase that is tightly associated and co-purifies with Tpp1p. Given the sensitivity of the assays and amount of fusion protein used, it is not unreasonable to expect to have observed a co-purifying activity even in the face of GST-Tpp1p overexpression. This is supported by our observation that an array-based two-hybrid screen (39) did not identify any P-loop-containing Tpp1p-interacting proteins (not shown).

Examination of the conditional dependence of GST-Tpp1p 3’ phosphatase activity revealed the following. Activity was maximal at 100 mM NaCl while salt concentrations above 200 mM were inhibitory (Fig. 5D). The enzyme was active over a broad pH range (6-9, not shown) and displayed an absolute requirement for a metal cofactor. Activity was optimal at 10 mM MgCl$_2$ (Fig. 5E), but the enzyme was similarly active in buffer containing 1 mM MnCl$_2$
Other divalent cations including Ca$^{2+}$, Ni$^{2+}$, Co$^{2+}$, and Zn$^{2+}$ could not support enzyme activity (not shown). Utilization of different DNA substrates revealed that GST-Tpp1p removed 3’ phosphates from nicks and single-nucleotide gaps with equal efficiency (Fig. 6). Although active at blunt ends, the enzyme was inactive on single-stranded oligonucleotides, further demonstrating that it is a structure-specific DNA 3’ phosphatase. Thus, it is highly likely that Tpp1p acts in the repair of damaged DNA. At present, it is unclear how this pattern of biochemical activity might explain the observed suppression of ketoconazole resistance by $TPP1$.

**Removal of 3’ phosphates and 3’ phosphoryl-terminated nucleotides by yeast cell-free extracts.** Crude whole cell extracts were next tested to assess the constitutive levels of Tpp1-dependent and independent 3’ phosphatase activities in yeast cells. It has been previously demonstrated that Apn1, the major AP endonuclease/3’ diesterase in *S. cerevisiae*, can remove 3’ phosphates (13). We therefore anticipated that Apn1 activity would substantially compete with Tpp1 in this experiment and so tested extracts prepared from wild-type as well as isogenic $tpp1$, $apn1$ and $tpp1\;apn1$ yeast. Wild-type extract resulted in the formation of the expected 3’ dephosphorylated 22-mer oligonucleotide product, as well as an unexpected 21-mer product corresponding to removal of the entire 3’ nucleotide at the nick (Fig. 7, lane 3). Again unexpectedly, nucleotide removal was not observed with either of the extracts lacking Apn1 (i.e. $apn1$ and $tpp1\;apn1$, lanes 5 and 6). These data suggest that Apn1 itself is able to remove both a 3’ phosphate and a nucleotide at a nick, a hypothesis that has been validated in a separate study. Of primary interest here, 3’ phosphate removal, evidenced by disappearance of the substrate, proceeded efficiently in both the $tpp1$ and $apn1$ mutant extracts (lanes 4 and 5), but was completely absent in the $tpp1\;apn1$ mutant extract (lane 6). Comparison of the $apn1$ and $tpp1$
apn1 extracts (lanes 5 and 6) specifically demonstrates that TPP1 contributes an abundant 3’ phosphatase activity.

Uncoupling of 3’ phosphatase and 5’ kinase activities in Saccharomyces as compared to Schizosaccharomyces extracts. Finally, crude extracts of S. cerevisiae, S. mikatae, and Schizosaccharomyces pombe were prepared and assayed using nicked oligonucleotide substrates in order to compare the relative levels of 5’ kinase and 3’ phosphatase activities. As predicted from the above results, all extracts possessed a 3’ phosphatase activity, although its level was considerably lower in the S. pombe extract (Fig. 7B). In contrast, but again predicted by the homology searches, only the S. pombe extract showed a detectable 5’ kinase (Fig. 7D, lanes 4-7) in an assay that depends on conversion of the nick to a ligatable form by 5’ phosphorylation (Figure 7C, similar to ref. 38). Importantly, the appearance of the 47-mer product was not the result of displacement polymerization because it was dependent on addition of T4 DNA ligase (not shown). In addition to the S. pombe 5’ kinase, these assay conditions also detected T4 PNK activity when it was spiked into the Saccharomyces extracts (Fig. 7D, lanes 9-12). We conclude that, in contrast to their abundant 3’ phosphatase, Saccharomyces yeast do not constitutively express a detectable DNA 5’ kinase, and likely lack one entirely, despite the initial observations with the plasmid transformation NHEJ assay.
Discussion

Given the potentially disastrous consequences of persistently blocked DNA termini, it is not surprising that multiple mechanisms have evolved to deal with these lesions. Blocking lesions can be resolved by strand degradation mechanisms such that nucleotides must be resynthesized in excess of those that have been directly damaged. Alternatively, direct reversal can occur by removal of only the fragmented nucleotide to leave the next available 3’ hydroxyl or 5’ phosphate, or by simple rephosphorylation of a 5’ hydroxyl. The recently described bifunctional hPNKP appears to have evolved to directly reverse both of the reciprocal lesions created by damage that results in a misplaced phosphate (16,17). The results presented here reinforce that this is not the only mechanism of dealing with such lesions, however, and that coupling is not obligatory.

Ultimately, it must be assumed that the PNK domains maintained through evolution reflect the biology of lesion repair in a given organism (Table 1, Figs. 2-4). Both 5’ kinase and 3’ phosphatase domains are present in metazoans and fission yeast in a linked fashion, indeed suggesting the potential for efficient reversal of coupled 3’ phosphate and 5’ hydroxyl lesions. In at least some plants this gene linkage was lost or never established. Importantly, it is as yet uncertain whether all putative 5’ kinase genes listed in Table 1 will function in DNA metabolism (note especially At-2, the only 5’ kinase gene in A. thaliana), and so these plants may represent 3’ phosphatase-only species from the standpoint of DNA repair.

Budding yeast stand out as the clearest example of a eukaryotic branch that possesses only a 3’ phosphatase. Caution is indicated, however, given that this conclusion is based on essentially negative results, namely that no PNK 5’ kinase homologue is revealed by comparison searching, no 5’ kinase is detected in the same crude extracts that display abundant 3’
phosphatase activity (Fig. 7), and no apparent 5’ kinase has been identified as interacting with Tpp1p in co-purification (Figs. 5 and 7) and two-hybrid approaches (not shown). It is possible that another gene provides a cryptic unlinked 5’ kinase activity. For example, Trl1p possesses coupled 5’ kinase and ligase activities whose biological role is in tRNA processing (29,40). The inviability of trl1 mutants prevents a simple testing of the hypothesis that there is a functional overlap between tRNA processing and DNA repair, but this seems unlikely, especially since Trl1p-dependent DNA 5’ kinase activity was not detected in the crude extracts. An alternative search approach based on pattern matching identified two ORFs that contain the PNK-like P-loop consensus [FYWVLIMC]GXP[GAS]XGKS[TSHY][FYWVLIMC] (Fig. 4) but that have no clear function based on literature or homology searching (YFR007W and YOR262W). These ORFs have no similarity to hPNKP beyond the P-loop, however, and it is difficult to reconcile why evolution would have resulted in a clearly conserved 3’ phosphatase and yet a highly diverged 5’ kinase. Even if these or another ORF are ultimately shown to provide cryptic DNA 5’ kinase activity, our results nonetheless demonstrate that there has been both a biochemical and structural uncoupling of 5’ kinase and 3’ phosphatase functions in Saccharomyces yeast.

It must be emphasized that no eubacteria or archaea proteins were uncovered in sequence comparisons or literature searches as containing either 3’ phosphatase or 5’ kinase domains, despite the large number of complete genome sequences that are available. It is clear, though, that a common DDDD-type phosphatase domain precursor was transmitted to both eukaryotes and bacteria, which in bacteria became histidinol phosphatase (34). In many bacteria this was fused to create a bifunctional enzyme containing the additional activity imidazole glycerolphosphate dehydratase (e.g. E. coli hisB). Interestingly, the yeast histidinol phosphatase His2p evolved from a structurally distinct phosphatase precursor of the “PHP” type that is also
found in DNA polymerases in all branches of life (41). A tremendous modularity is thus apparent in the evolution of phosphatase domains, with the ancient precursor types being variably adapted to fulfill critical cellular functions, and in some cases fused to novel domains. In contrast, no clear bacterial counterpart to the 5′ kinase domain was made evident in the Psi-BLAST search, beyond the ubiquitous Walker A motif. This pattern strongly suggests that the PNK 3′ phosphatase domain is evolutionarily older, while the DNA 5′ kinase arose independently and was fused only after the divergence of budding yeasts (Fig. 8). An alternative hypothesis would be that the 5′ kinase and 3′ phosphatase domains were only uncommonly split and/or lost in some eukaryotic lineages from a common coupled PNK progenitor. Redundant mechanisms for 5′ blocking lesion resolution might have simply allowed the 5′ kinase domain to be lost without detriment as a result of a rare deletion.

As sequences are generated from still more budding yeasts as a result of comparative genomic efforts (42) it should be possible to determine whether the 5′ kinase domain is indeed absent from this lineage. At present, we favor this hypothesis in part because it is less cumbersome than a cycle of fusion and loss. Moreover, the alternative hypothesis presupposes that coupling of 5′ kinase and 3′ phosphatase functions is inherently advantageous in most species. In fact, the fragmentation of bases generates a great diversity of blocking lesions, of which coupled 3′ phosphates and 5′ hydroxyls are only a subset. For example, 3′ phosphate generation by Fpg entails sequential β and δ elimination reactions, leaving 3′ and 5′ phosphates at a single nucleotide gap (5). Also, cleavage of topoisomerase I-DNA covalent complexes by tyrosyl-DNA phosphodiesterase (Tdp1p in S. cerevisiae) is thought to occur after replication fork collapse caused by collision-induced double-strand breaks, and so the resulting 3′ phosphate and 5′ hydroxyl would no longer be physically linked in a single-strand nick (43). We suggest that
the overall pattern of PNK conservation reflects this lack of coupling. The greatest question is whether domain linkage in fission yeast and metazoans underlies a true increase in nick repair efficiency, fulfills a biological role specific to cells that divide by fission, or was more simply a way of ensuring co-inheritance once the domains each occupied a unique niche in nick repair.

In the case of 5’ blocking lesions, higher eukaryotes depend on DNA polymerase β lyase activity to remove 5’ dRP lesions in the rate-limiting step in BER (2,3). In contrast, the only yeast β-like polymerase, Pol4p, is not required for BER (44,45). Instead, polymerization function is provided primarily by DNA polymerase δ (45,46), presumably in conjunction with more extensive 5’ base loss in a long patch repair. Similarly, budding yeast process double-strand breaks almost exclusively by recombination mechanisms that entail highly efficient 5’ degradation, whereas end-preserving NHEJ plays a far greater role in higher eukaryotes (7,27). We argue that this emphasis on more extensive 5’ degradation also applies to 5’ hydroxyl lesions in S. cerevisiae. As a result, the need for a 5’ kinase, just as for the 5’ dRP lyase, does not exist. The observed repair of dephosphorylated plasmids would thus occur by base removal and resynthesis which, based on our results, must occur by a Pol4-independent pathway of end-processing in NHEJ. Because NHEJ entailing simple nucleotide gaps shows a minimal pol4 effect, this is perhaps not surprising (compare Figure 1 and ref. 22).

In the case of 3’ blocking lesions, further studies will be required to determine the specific role(s) that TPP1 has been maintained to fulfill. The observed specificity of this enzyme for 3’ phosphates present in double-stranded DNA strongly suggests a role in genome repair (Fig. 6). More specifically, the competitiveness of Tpp1p and the promiscuous 3’ diesterase Apn1p observed using cell-free extracts seems to foreshadow a complex interplay between multiple pathways of 3’ processing (Fig. 7). Results to be presented elsewhere indicate that
several DNA repair pathways display functional overlap between Tpp1p and Apn1p². In contrast, bacteria apparently require only the universally conserved 3’ diesterases (47). Why the seemingly redundant Tpp1p/Apn1p enzyme combination was established and maintained in budding yeast, and indeed all eukaryotes characterized to date, while the PNK 5’ kinase was not will be important issues to resolve.
Acknowledgements

We thank Elissa Karathanasis for technical assistance on Fig. 1, and Peter Uetz of the Stan Fields laboratory for performing the array-based two-hybrid analysis with YMR156C. This work was supported in part by the Pew Scholars Program in the Biomedical Sciences of the Pew Charitable Trusts (TEW), and by NIH postdoctoral training grant 5T32HL07157 (JRV).
References

1. Friedberg, E. C., Walker, G. C., and Siede, W. (1995) DNA Repair and Mutagenesis, ASM Press, Washington, D.C.

2. Sobol, R. W., Prasad, R., Evenski, A., Baker, A., Yang, X. P., Horton, J. K., and Wilson, S. H. (2000) Nature 405, 807-10

3. Srivastava, D. K., Berg, B. J., Prasad, R., Molina, J. T., Beard, W. A., Tomkinson, A. E., and Wilson, S. H. (1998) J Biol Chem 273, 21203-9

4. Girard, P. M., Guibourt, N., and Boiteux, S. (1997) Nucleic Acids Res 25, 3204-11

5. Bhagwat, M., and Gerlt, J. A. (1996) Biochemistry 35, 659-65

6. Memisoglu, A., and Samson, L. (2000) Mutat Res 451, 39-51

7. Paques, F., and Haber, J. E. (1999) Microbiol Mol Biol Rev 63, 349-404

8. Anderson, D. G., and Kowalczykowski, S. C. (1997) Genes Dev 11, 571-81

9. Pierson, C. E., McCullough, A. K., and Lloyd, R. S. (2000) Mutat Res 459, 43-53

10. Povirk, L. F. (1996) Mutat Res 355, 71-89

11. Ramotar, D. (1997) Biochem Cell Biol 75, 327-36

12. Mol, C. D., Hosfield, D. J., and Tainer, J. A. (2000) Mutat Res 460, 211-29

13. Johnson, A. W., and Demple, B. (1988) J Biol Chem 263, 18017-22

14. Midgley, C. A., and Murray, N. E. (1985) Embo J 4, 2695-703

15. Ichimura, M., and Tsukada, K. (1971) J Biochem (Tokyo) 69, 823-8

16. 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., and Weinfeld, M. (1999) J Biol Chem 274, 24187-94
17. Jilani, A., Ramotar, D., Slack, C., Ong, C., Yang, X. M., Scherer, S. W., and Lasko, D. D. (1999) *J Biol Chem* **274**, 24176-86

18. Karimi-Busheri, F., and Weinfeld, M. (1997) *J Cell Biochem* **64**, 258-72

19. Whitehouse, C. J., Taylor, R. M., Thistlewaite, A., Zhang, H., Karimi-Busheri, F., Lasko, D. D., Weinfeld, M., and Caldecott, K. W. (2001) *Cell* **104**, 107-117

20. Fischer, G., James, S. A., Roberts, I. N., Oliver, S. G., and Louis, E. J. (2000) *Nature* **405**, 451-4

21. Erdeniz, N., Mortensen, U. H., and Rothstein, R. (1997) *Genome Res* **7**, 1174-83

22. Wilson, T. E., and Lieber, M. R. (1999) *J Biol Chem* **274**, 23599-609

23. Wilson, T. E., Grawunder, U., and Lieber, M. R. (1997) *Nature* **388**, 495-8

24. Brachmann, C. B., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P., and Boeke, J. D. (1998) *Yeast* **14**, 115-32

25. Schuler, G. D., Altschul, S. F., and Lipman, D. J. (1991) *Proteins* **9**, 180-90

26. Martzen, M. R., McCraith, S. M., Spinelli, S. L., Torres, F. M., Fields, S., Grayhack, E. J., and Phizicky, E. M. (1999) *Science* **286**, 1153-5

27. Featherstone, C., and Jackson, S. P. (1999) *Curr Biol* **9**, R759-R761

28. Walker, J. E., Saraste, M., Runswick, M. J., and Gay, N. J. (1982) *Embo J* **1**, 945-51

29. Koonin, E. V., and Gorbalenya, A. E. (1990) *FEBS Lett* **268**, 231-4

30. Westaway, S. K., Phizicky, E. M., and Abelson, J. (1988) *J Biol Chem* **263**, 3171-6

31. Ayres, M. D., Howard, S. C., Kuzio, J., Lopez-Ferber, M., and Possee, R. D. (1994) *Virology* **202**, 586-605

32. Collet, J. F., Stroobant, V., Pirard, M., Delpierre, G., and Van Schaftingen, E. (1998) *J Biol Chem* **273**, 14107-12
33. Aravind, L., Galperin, M. Y., and Koonin, E. V. (1998) *Trends Biochem Sci* **23**, 127-9

34. Thaller, M. C., Schippa, S., and Rossolini, G. M. (1998) *Protein Sci* **7**, 1647-52

35. Jilani, A., Slack, C., Matheos, D., Zannis-Hadjopoulos, M., and Lasko, D. D. (1999) *J Cell Biochem* **73**, 188-203

36. Thompson, L. H., and West, M. G. (2000) *Mutar Res* **459**, 1-18

37. Launhardt, H., Hinnen, A., and Munder, T. (1998) *Yeast* **14**, 935-42

38. Karimi-Busheri, F., Lee, J., Tomkinson, A. E., and Weinfeld, M. (1998) *Nucleic Acids Res* **26**, 4395-400

39. Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P., Qureshi-Emili, A., Li, Y., Godwin, B., Conover, D., Kalbfleisch, T., Vijayadamodar, G., Yang, M., Johnston, M., Fields, S., and Rothberg, J. M. (2000) *Nature* **403**, 623-7

40. Abelson, J., Trotta, C. R., and Li, H. (1998) *J Biol Chem* **273**, 12685-8

41. Aravind, L., and Koonin, E. V. (1998) *Nucleic Acids Res* **26**, 3746-52

42. Ozier-Kalogeropoulos, O., Malpertuy, A., Boyer, J., Tekaya, F., and Dujon, B. (1998) *Nucleic Acids Res* **26**, 5511-24

43. Pouliot, J. J., Yao, K. C., Robertson, C. A., and Nash, H. A. (1999) *Science* **286**, 552-5

44. Prasad, R., Widen, S. G., Singhal, R. K., Watkins, J., Prakash, L., and Wilson, S. H. (1993) *Nucleic Acids Res* **21**, 5301-7

45. Halas, A., Baransowska, H., Policinska, Z., and Jachymczyk, W. J. (1997) *Curr Genet* **31**, 292-301

46. Wood, R. D., and Shivji, M. K. (1997) *Carcinogenesis* **18**, 605-10

47. Eisen, J. A., and Hanawalt, P. C. (1999) *Mutar Res* **435**, 171-213
Footnotes

1. The abbreviations used are: dRP, deoxyribose phosphate; BER, base excision repair; AP, apurinic-apyrimidinic; PNK, polynucleotide kinase; ORF, open reading frame; CIP, calf-intestinal alkaline phosphatase; NHEJ, nonhomologous end-joining; HAD, haloacid dehalogenase; PARP, poly(ADP-ribose) polymerase; LIG3, DNA ligase III; GST, glutathione S-transferase.

2. Vance, J.R. and Wilson, T.E. (2001) Competing pathways for repair of 3’ phosphate blocking lesions in the yeast *Saccharomyces cerevisiae*. *In preparation.*
Figure Legends

Fig. 1. Efficient \textit{POL4}-independent NHEJ of transformed plasmids bearing dephosphorylated \textit{Bgl} II ends. (A) The schematic shows potential pathways for processing of dephosphorylated \textit{Bgl} II ends during NHEJ. Direct reversal of the 5’ hydroxyl by PNK or nucleotide removal and resynthesis through the combined action of a nuclease and polymerase (nuc + pol) might each occur either before or after end annealing. (B) After normalization to parallel transformations with uncut plasmid to correct for small differences in plasmid uptake, ligation-competent and dephosphorylated plasmid transformation efficiencies from (wt), \textit{dnl4-K282R} (\textit{dnl4}) and \textit{pol4-D367E} (\textit{pol4}) yeast strains were plotted relative to the wt/non-CIP-treated combination. White circles represent white (\textit{ADE2}) colonies and therefore simple religation of the \textit{Bgl} II ends, while black circles represent red (\textit{ade2}) colonies. Results are from four independent transformation experiments performed on three different days (mean ± standard deviation).

Fig. 2. PNK multiple sequence alignment. This schematic depicts the alignment of a representative subset of PNK sequences. Tall and short boxes represent aligned and unaligned (“unlinked”) regions, respectively, blank regions represent gaps, and diagonal lines indicate that sequences were deleted from the figure to conserve space. Gene/species designations are abbreviated as in Table 1, with a hyphenated number corresponding to the gene number. Sequence Dd-1 is shown twice because the ESTs for this single cDNA could not be assembled into an unbroken contig. AcNPV-2 and T4 PNK are shown twice because the order of the 3’ phosphatase and 5’ kinase domains are inverted relative to the other genes, and so a simple linear alignment is not possible. Homology to non-PNK
sequences are indicated above the relevant regions (PARP, poly(ADP-ribose) polymerase; LIG3, DNA ligase III; YOR258W, yeast hypothetical protein).

**Fig. 3. Structural motifs of the PNK 3’ phosphatase domain.** Selected sequence regions from the complete alignment are shown that correspond to the conserved motifs of the 3’ phosphatase domain. Positions identical in all species from human to bacteriophage T4 are highlighted as yellow-on-red type. Positions identical or similar to a consensus of 50% of the sequences are highlighted as yellow-on-blue and white-on-green type, respectively. “X” in a sequence indicates an ambiguous amino acid or string. Two PNK phosphatase subtypes evident in the alignment are separated by a horizontal line. Motifs discussed in the text are indicated, with the corresponding *E. coli* hisB sequence shown above the alignment where relevant. Asterisks mark amino acid positions in motifs 1-4 that, in addition to motifs A and B, provide strong discrimination between the histidinol and DNA 3’ phosphatase families.

**Fig. 4. Structural motifs of the PNK 5’ kinase domain.** Selected sequence regions from the complete alignment are shown that correspond to the conserved motifs of the 5’ kinase domain. The depiction is the same as in Fig. 3. Additionally, the Walker A box is indicated along with the PNK-type P-loop consensus (h = hydrophobic, x = any amino acid). Asterisks mark amino acid positions that delineate the three apparent 5’ kinase subtypes, which are separated by horizontal lines.

**Fig. 5. *S. cerevisiae* ORF YMR156C (*TPP1*) encodes a DNA 3’ phosphatase.** (A) A nicked oligonucleotide substrate was used to detect 3’ phosphatase activity. Removal of the 3’ phosphate from the 32P-labeled 22-mer (indicated by an asterisk) leads to decreased mobility in a sequencing gel. (B) GST-Tpp1p (i.e. GST-YMR156C, Lane 2) and non-recombinant GST
(Lane 3) proteins were purified from yeast, run on a 10% SDS-polyacrylamide gel, and stained with Coomassie Blue. Relative molecular weights are indicated by standards (M, Lane 1). (C) The 5’ end-labeled substrate (50 fmol) was incubated with 10 fmol GST-Tpp1p (lane 3) or GST protein (lane 4) for 10 min at 30 °C and analyzed on a sequencing gel followed by autoradiography. Lanes 1 and 2 contain the corresponding 22-mer oligonucleotide synthesized without and with a 3’ phosphate, respectively. These marker lanes are repeated in panel (D) and Fig. 7A. (D) Salt dependence of GST-Tpp1p activity, assayed as in (C) except varying the NaCl concentration as shown. (E) and (F) Divalent cation dependence of GST-Tpp1p activity, assayed as in (C) except varying the MgCl2 and MnCl2 concentrations as shown. In (F), triangles indicate decreasing amounts of GST-Tpp1p (10, 4, 2, 1.3 and 1 fmol, left to right).

**Fig. 6. Tpp1p is active on dsDNA but not ssDNA.** Oligonucleotides containing 3’ phosphates in different configurations were tested for their ability to act as substrates for varying amounts of GST-Tpp1p, as in left portion of Fig. 5F. Substrates are schematized above the relevant lanes, and represent 3’ phosphates at a nick, single-nucleotide gap, ssDNA end, extended 5’ overhang, and dsDNA blunt end (left to right, top to bottom). All were approximately equally efficient substrates except ssDNA, which was not dephosphorylated by even the highest amount of GST-Tpp1p.

**Fig. 7. Uncoupling of 3’ phosphatase and 5’ kinase activities in cell-free extracts of Saccharomyces yeast.** (A) Glass-bead extracts from wild-type (wt), tpp1, apn1, and tpp1 apn1 strains were tested for 3’ phosphatase activity as in Fig. 5. The extracts resulted variably in removal of the 3’ phosphate or 3’ terminal nucleotide, yielding 22-mer and 21-mer oligonucleotide products, respectively. (B) GST-Tpp1p and glass-bead extracts from S. cerevisiae, S. mikatae (Sm), S. pombe (Sp) were used in 3’ phosphatase assays as in (A), to
provide an activity comparison to (D). (C) A nicked oligonucleotide substrate was used to detect 5’ kinase activity. Following phosphorylation of the 5’ terminus at the nick, ligation by exogenously added T4 DNA ligase leads to an increase in size of the labeled strand (indicated by an asterisk) from 22 to 47 nucleotides. (D) T4 PNK (5 units) and the same protein samples used in (B) were incubated with the 5’ kinase substrate (50 fmol) in the presence of 100 units of T4 DNA ligase at 30 °C for 10 min and analyzed on a sequencing gel followed by autoradiography. *apn1* mutant *S. cerevisiae* extracts were used to rule out inhibition by the Apn1p nuclease. Also, T4 PNK was included in a duplicate of all reactions to demonstrate that the extracts do not prevent detection of 5’ kinase activity (note that nicked DNA is a poor substrate for T4 PNK).

**Fig. 8. Modular evolution of 3’ phosphatase and histidinol phosphatase domains.**

The diagram illustrates the idealized relationships between the PHP and DDDD-type phosphotransferase superfamilies and eukaryotic and prokaryotic DNA 3’ phosphatases (3’ Pase) and histidinol phosphatases (H’ol Pase). See text for discussion. Dashed lines indicate less favored possibilities. Many other putative relationships are not illustrated, notably the derivation of the 5’ kinase domain from its precursors, as well as its presumed duplication and subsequent fusion with still other domains in the derivation of tRNA ligase, etc. IGP-deHase; imidazole-glycerolphosphate dehydratase.
| Eukaryota | Species  | Gene 1          | Gene 2          |
|----------|----------|-----------------|-----------------|
| Metazoa  | Hs human | 3' Pase-5' kinase |                  |
|          | Mm mouse | 3' Pase-5' kinase |                  |
|          | Gg chicken | 3' Pase EST |                  |
|          | Xl frog | 3' Pase EST |                  |
|          | Dm fruitfly | 3' Pase-5' kinase |                  |
|          | Bm silkworm | 3' Pase EST |                  |
|          | Ov roundworm | 3' Pase EST |                  |
|          | Bma roundworm | " | 3' Pase EST |
|          | Ce | 3' Pase-5' kinase | 5' kinase |
| Viridiplantae | At thale cress | 3' Pase | 5' kinase |
|          | Gm soybean | 3' Pase EST |                  |
|          | Zm corn | 3' Pase EST |                  |
|          | Cr green algae | 3' Pase EST |                  |
| Fungi    | Sp fission yeast | 3' Pase-5' kinase |                  |
|          | Sc baker's yeast | 3' Pase | tRNA lig-5' kinase |
|          | Sm | 3' Pase |                  |
|          | Kl milk yeast | 3' Pase STS |                  |
| Euglenozoa | Tc trypanosome | 3' Pase EST |                  |
| Dictyosteliida | Dd slime mold | 3' Pase-5' kinase EST | 5' kinase EST |
| Viridae  | SeNPV baculovirus | 3' Pase |                  |
|          | AcNPV " | 3' Pase | tRNA lig-5' kinase-3' Pase |
| T4 bacteriophage | 5' kinase-3' Pase |                  |

Table 1. Species distribution of 3' phosphatase and 5' kinase domains identified by homology and literature searches. Species are grouped by taxonomy.

Abbreviations are: Hs, *Homo sapiens*; Mm, *Mus musculus*; Gg, *Gallus gallus*; Xl *Xenopus laevis*; Dm, *Drosophila melanogaster*; Bm, *Bombyx mori*; Ov, *Onchocerca volvulus*; Bma, *Brugia malayi*; Ce, *Caenorhabditis elegans*; At, *Arabidopsis thaliana*; Gm, *Glycine max*; Zm, *Zea mays*; Cr, *Chlamydomonas reinhardtii*; Sp, *Schizosaccharomyces pombe*; Sc, *Saccharomyces cerevisiae*; Sm, *Saccharomyces mikatae*; Kl, *Kluyveromyces lactis*; Tc, *Trypanosoma cruzi*; Dd, *Dictyostelium discoideum*; SeNPV, *Spodoptera exigua* nucleopolyhedrovirus; AcNPV, *Autographa californica* nucleopolyhedrovirus.
domains present in gene 1 and, when relevant, gene 2 for a given species are indicated in
amino-carboxyl terminal order. The Hs gene is hPNKP, Sc gene 1 is YMR156C (TPP1),
Sc gene 2 is TRL1, and the T4 gene is PNK. “EST” and “STS” designate partial sequences,
and therefore conclusions cannot be drawn regarding the domains that are not present. The
Cr gene has a sequence that can be aligned with the consensus Walker A motif, but it
shows substantial divergence relative to all other proteins, and so it is unclear whether it
represents a true 5’ kinase.
Figure 1

A.

B.

Relative Efficiency

CIP:

-  +  -  +  -  +

wt  dnl4  pol4
Figure 5

A

22 nt

POH

5' *AATTACGAATGCCACACCCGCGGCGCGCGCCACCCGACTAGTGCGCG
TAATGCTTACGGGTGTGGCGGCGCGCGGGTGTGGTGATCGACCCG

B

GST-Tpp1p

GST

Kd

158

97

66

56

37

27

1

2

3

D

NaCl₂ (mM)

- 0 100 150 200 250 300 350

22

22p

1 2 3 4

E

MgCl₂ (mM)

- 0 2.5 5 7.5 10 20 50

22

22p
Figure 6
Figure 7

A

wt tpp1 apn1 tpp1 apn1

22p

21p

1 2 3 4 5 6

B

GST-Tpp1p Sc wt apn1 Sm Sp

22p

21p

1 2 3 4 5 6 7

C

47 nt

22 nt

H2O2

5' AATTACGAATGCCACACCGGCACCACCACCACCACACTAGCTGCGG

TAATGCTTACGGGTGTGGCGGCCGCGGGTGGTGATCGACCGG

D

- T4 PNK

GST-Tpp1p Sc wt apn1 Sm Sp

+ T4 PNK

47

22

21

1 2 3 4 5 6 7 8 9 10 11 12

22

21
Uncoupling of 3' phosphatase and 5' kinase functions in budding yeast: characterization of S. cerevisiae DNA 3' phosphatase (TPP1)
John R. Vance and Thomas E. Wilson

J. Biol. Chem. published online January 30, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M011075200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts