Characterization and evolutionary history of an archaeal kinase involved in selenocysteinyl-tRNA formation

R. Lynn Sherrer1, Patrick O’Donoghue1 and Dieter Söll1,2,*

1Department of Molecular Biophysics and Department of Biochemistry and 2Department of Chemistry, Yale University, New Haven, CT 06520-8114, USA

Received November 17, 2007; Revised December 4, 2007; Accepted December 5, 2007

ABSTRACT

Selenocysteine (Sec)-decoding archaea and eukaryotes employ a unique route of Sec-tRNASec synthesis in which O-phosphoseryl-tRNASec kinase (PSTK) phosphorylates Ser-tRNASec to produce the O-phosphoseryl-tRNASec (Sep-tRNASec) substrate that Sep-tRNA:Sec-tRNA synthase (SepSecS) converts to Sec-tRNASec. This study presents a biochemical characterization of Methanocaldococcus jannaschii PSTK, including kinetics of Sep-tRNASec formation (Km for Ser-tRNASec of 40 nM and ATP of 2.6 mM). PSTK binds both Ser-tRNASec and tRNASec with high affinity (Kd values of 53 nM and 39 nM, respectively). The ATPase activity of PSTK may be activated via an induced fit mechanism in which binding of tRNASec specifically stimulates hydrolysis. Albeit with lower activity than ATP, PSTK utilizes GTP, CTP, UTP and dATP as phosphate-donors. Homology with related kinases allowed prediction of the ATP-binding site by in vitro and in vivo analysis of PSTK mutants. A detailed phylogenetic analysis of PSTK in the context of its close relatives from the DxTN kinase family is also presented.

INTRODUCTION

The micronutrient selenium is found in proteins as Sec, a natural amino acid that is cotranslationally inserted into proteins in response to the codon UGA. Sec is formed in a tRNA-dependent transformation of Ser to Sec. In bacteria this is achieved in a single step by selenocysteine synthase, SelA, in the presence of the selenium donor selenophosphate (1). Eukaryotes and archaea add an additional step to Sec-tRNASec formation. Following seryl-tRNA synthetase (SerRS), PSTK specifically phosphorolysates Ser-tRNASec to form Sep-tRNASec (2,3). Recently, the enzyme responsible for Sep-tRNASec to Sec-tRNASec conversion, SepSecS, was uncovered in both mammals and archaea (4,5). While PSTK activity was initially found in rat and rooster liver (6) and in the bovine mammary gland (7), the enzyme was later characterized for bovine liver (8), human (9), mouse (2) and M. jannaschii (3). The enzyme requires both ATP and Mg2+ for activity. PSTK was shown to transfer the γ-phosphate from ATP to Ser-tRNASec, yielding Sep-tRNASec and ADP. A computational search of several archaeal and eukaryotic genomes for a kinase-like gene present only in those organisms containing the Sec insertion machinery identified the pstk gene (2).

Here we present a biochemical characterization of M. jannaschii wild-type PSTK and the identification of the ATP-binding site by in vitro and in vivo analysis of PSTK mutants. A detailed phylogenetic analysis of PSTK in the context of its close relatives from the DxTN kinase family is also presented.

MATERIALS AND METHODS

Materials and reagents

All oligonucleotide synthesis and DNA sequencing was carried out by the Keck Foundation Biotechnology Research Laboratory at Yale University. [γ-32P]ATP (6000 Ci/mmol), L-[U-14C]serine (163 mCi/mmol), and [α-32P]ATP (3000 Ci/mmol) were from GE Healthcare.

Cloning, expression and purification of enzymes

M. jannaschii PSTK (MJ1538) was cloned between the Nde I and Xho I restriction sites in the pET20b vector (Novagen) with a C-terminal His6 tag. PSTK-pET20b was

*To whom correspondence should be addressed. Tel: +1 203 432 6200; Fax: +1 203 432 6202; Email: dieter.soll@yale.edu

© 2008 The Author(s)
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.0/uk/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
transformed into *Escherichia coli* BL21 (DE3) codon plus (Stratagene). A pre-culture was used to inoculate 800 ml of LB broth with 100 μg/ml of ampicillin, 34 μg/ml chloramphenicol, 5052 solution, and phosphate buffer for autoinduction as described previously (10). The cells were grown for 8 h at 37°C and continued at 15°C for 14–16 h. The cells were pelleted and resuspended in 50 mM Tris–HCl (pH 7.0), 500 mM NaCl, 10% glycerol, 0.2 mM PMSF. After sonication and centrifugation, the cell lysates were applied to TALON metal affinity resin (Clontech) and purified according to the manufacturer’s instructions. The eluted enzymes were dialyzed into 25 mM Hepes–KOH (pH 7.5), 500 mM NaCl and 50% glycerol. SDS–PAGE electrophoresis followed by staining with Coomassie blue revealed greater than 95% purity. *Methanococcus maripaludis* SerRS was overexpressed as described above and purified as described previously (11).

**Mutagenesis of PSTK active site**

Point mutations were introduced into amino acid codons for the P-loop residues Gly14, Lys17, Ser18, Thr19, the Walker B motif residue Asp41, and the RxxxR residues Arg116 and Arg120 with the QuikChange site-directed mutagenesis kit (Stratagene) according to the manufacturer’s instructions. After verification by DNA sequencing, the plasmids were transformed into BL21(DE3) codon plus cells (Stratagene). The mutant proteins were overexpressed and purified as described above.

**tRNA purification**

The gene encoding *M. maripaludis* tRNA<sub>Sec</sub> with the preceding sequence of the T7 promoter was cloned into the pUC18 plasmid, and the gene encoding *M. maripaludis* tRNA<sub>Ser</sub> with the preceding sequence of the T7 promoter was cloned into the pUC19 plasmid. Both were expressed in *E. coli* DH5α. Plasmid DNA was purified using the HISpeed Plasmid Maxi kit (Qiagen). The purified plasmid was digested with BstNI for run-off transcription as described previously (12). The transcript was phenol–chloroform extracted, ethanol precipitated, and purified by electrophoresis on a 12% denaturing polyacrylamide gel. The tRNA transcripts were refolded by heating for 5 min at 70°C in buffer containing 10 mM Tris–HCl (pH 7.0), followed by addition of 5 mM MgCl₂ and immediate cooling on ice (11).

**Preparation of labeled tRNA**

Refolded transcript was <sup>32</sup>P-labeled on the 3′ terminus by using the *E. coli* CCA-adding enzyme and [α-<sup>32</sup>P]ATP as described previously with some modification (13). Briefly, 8 μM of tRNA<sub>Sec</sub> or tRNA<sub>Ser</sub> transcript was incubated with the CCA-adding enzyme and 0.5 μCi/μl [α-<sup>32</sup>P]ATP for 45 min at room temperature in buffer containing 50 mM Tris–HCl (pH 8.0), 20 mM MgCl₂, 5mM DTT and 50 μM NaPPi. After phenol/chloroform extraction the sample was passed over a Sephadex G25 Microspin column (Amersham Biosciences) to remove excess ATP.

**Preparation of seryl-tRNA**

Transcript was aminoacylated in 1× PSTK buffer [50 mM Hepes–KOH (pH 7.5), 10 mM MgCl₂, 20 mM KCl, 1 mM DTT] with 1 mM t-Ser (Sigma), 5 mM ATP, 3 μM *M. maripaludis* SerRS and 5 μM <sup>32</sup>P-labeled transcript. The reaction was incubated at 37°C for 1 h followed by phenol/chloroform extraction, ethanol precipitation and resuspension in water. The samples were passed a minimum of two times over Sephadex G25 Microspin columns (Amersham Biosciences) equilibrated with water. Transcript that was used to determine the *Kₘ* for ATP and in the NTP preference studies was passed over two to three G25 columns equilibrated with 25 mM Tris–HCl (pH 7.0), 150 mM NaCl to remove all detectable ATP, followed by one G25 column equilibrated with water. Unlabeled transcripts were aminoacylated in parallel. To check aminoacylation levels, 2 μl aliquots were removed at the start and end of the reactions with labeled tRNA, quenched on ice with 3 μl of 100 mM sodium citrate (pH 5.0) and 0.66 mg/ml nuclease P1 (Sigma), and incubated at room temperature for 35 min (14,15). To separate Ser-AMP from AMP, 1 μl of quenched, digested sample was spotted on glass polyethyleneimine (PEI) cellulose 20 cm × 20 cm thin layer chromatography (TLC) plates (EMD) and developed for 75 min in 100 mM ammonium acetate and 5% acetic acid. The plates were exposed on an imaging plate (FujiFilms) for 16 h, scanned on a Molecular Dynamics Storm 860 PhosphorImager, and quantified using ImageQuant software.

**TLC separation of Ser-AMP, AMP and Sep-AMP using TLC**

Three assay conditions were utilized to confirm the identities of Ser-AMP, AMP and Sep-AMP after TLC analysis. In 1× PSTK buffer, 1 μM <sup>32</sup>P-labeled tRNA<sub>Sec</sub> transcript was incubated with 5 mM ATP, 1 mM L-Ser, 600 nM SerRS and 100 nM PSTK for 45 min at 37°C. The reactions were quenched, digested and analyzed by TLC as stated above. The second reaction conditions included 1× PSTK buffer, 2 μM tRNA<sub>Sec</sub> transcript, 100 μM [<sup>14</sup>C]Ser, 5 mM ATP, 1.2 μM SerRS and 200 nM PSTK for 45 min at 37°C. The reactions were stopped by phenol/chloroform extraction and put over a G25 column to remove unincorporated [<sup>14</sup>C]Ser, and 2 μl aliquots (3000 cpm/μl) were digested in 3 μl of 100 mM sodium citrate (pH 5.0) and 0.66 mg/ml nuclease P1 (Sigma) and analyzed by TLC as above. Lastly, 1 μM Ser-tRNA<sub>Sec</sub> or Ser-tRNA<sub>Ser</sub> was incubated in 1× PSTK buffer with 1.67 μM γ-[<sup>32</sup>P]ATP and 100 nM PSTK for 45 min at 37°C. The reactions were immediately put over a G25 column, and 2 μl aliquots were quenched, digested and analyzed by TLC as stated above.

**Measurement of tRNA binding to PSTK**

The affinity of PSTK for Ser-tRNA<sub>Sec</sub>, tRNA<sub>Sec</sub>, Ser-tRNA<sub>Sec</sub> and tRNA<sub>Ser</sub> was measured using the filter-binding method described previously with slight modifications (16,17). <sup>32</sup>P-labeled transcript and <sup>32</sup>P-labeled and
serylated transcript was prepared as stated above. Dissociation constants for each were determined by incubating increasing concentrations of PSTK on ice in 25 μl of binding buffer [50 mM Hepes–KOH (pH 7.5), 10 mM MgCl2, 20 mM KCl, 1 mM DTT, 5% glycerol] for 15 min after the addition of 5 nM 32P-labeled Ser-tRNA^{Sec}, tRNA^{Sec}, Ser-tRNA^{Ser} or tRNA^{Ser}. Minimal deacylation of Ser-tRNA^{Sec} and Ser-tRNA^{Ser} was detected over this time frame under the described conditions. A 96-well vacuum manifold (Hybri-dot 96; Whatman Biometra) was used to spot aliquots of the binding reaction onto the upper nitrocellulose membrane (MF-Membrane Filter; Millipore) and a lower nylon membrane (Hybond-N+; Amersham). Prior to use the membranes were pre-washed in RNase-free water and soaked for at least 30 min in binding buffer at 4°C. Aliquots of 7 μl from each binding reaction were spotted in triplicate and washed with 200 μl of ice-cold binding buffer. The levels of radiolabeled tRNA on each filter were quantified by PhosphorImager and then used to determine the ratio of RNA bound to RNA total after correction for nonspecific binding (17). The corrected data were plotted as a semilog plot of the fraction of total RNA bound to nitrocellulose versus the log of PSTK concentration, and the resulting plot was fitted to the following equation: 

\[
[RE] = (\frac{[R_{\text{total}}] \times [E]}{[K_d] + [E]})
\]

using Kaleidagraph v. 3.6 (Synergy Software) (R, RNA; E, enzyme).

**Phosphotransferase assay**

These assays were carried out in 1× PSTK buffer at 37°C. Unless otherwise noted, 5 mM ATP and 200–600 nM 32P-labeled Ser-tRNA^{Sec} was added. For kinetic parameter determination, initial velocities were measured in triplicate while varying concentration of one substrate and saturating with the other. When determining kinetic parameters with varying Ser-tRNA^{Sec} concentration, reactions were for 50s with 2 nM enzyme. When determining kinetic parameters with varying ATP concentration, reactions were for 120s with 6 nM enzyme. Reaction mixtures were preincubated at 37°C and started by addition of enzyme. At each time point, 2 μl aliquots were taken and treated as described above for aminoacylation. The previous TLC conditions described were also used to separate Ser-AMP, AMP and Sep-AMP (Figure 2A). Kaleidagraph v. 3.6 (Synergy Software) was used to calculate the kinetic parameters using nonlinear regression plots of the initial velocity versus substrate concentration (Ser-tRNA^{Sec} or ATP).

The phosphotransferase assays performed with wild-type PSTK and the PSTK active site mutant enzymes, G14W, K17A, S18A, T19W, D41A, R116A and R120A were either carried out with 10 nM enzyme for 1 min or 100 nM enzyme for 5 min. Aliquots of 2 μl were taken and treated as described above for aminoacylation. The aminoacylated tRNA substrate remained stable over these time periods.

For phosphoryl donor preference determination, the phosphotransferase assays were carried out for 6 min with 6 nM PSTK and either no NTP or 20 nM ultrapure (>99%) NTP (ATP, GTP, CTP, UTP) (Sigma) or dATP (New England Biolabs). For assays using 20 nM α,β-methyleneadenosine-5'-triphosphate (AMP-CPP) or β,γ-methyleneadenosine-5'-triphosphate (AMP-PCP), 200 nM PSTK was used. Aliquots of 2 μl were taken at time points and quenched, digested and analyzed by TLC as stated previously, except reactions with AMP-CPP and AMP-PCP were digested with 6.6 mg/ml nuclease PI (Sigma). The k phosphorylation (k phosph) values for Ser-tRNA^{Sec} to Sep-tRNA^{Sec} conversion were calculated for each phosphoryl donor.

**ATPase activity measurement**

ATPase activity was determined by measuring the amount of [γ-32P]ATP converted to [γ-32P]ADP as described before with modifications (18). These assays were carried out in a 5 μl reaction volume including 1× PSTK buffer with 130 nM cold ATP, 100 nM [γ-32P]ATP and 200 nM enzyme at 37°C for 45 min. Unless noted otherwise, 1 μM unlabeled Ser-tRNA^{Sec} was included. The reactions were quenched by the addition of 45 μl ice-cold 55 mM EDTA. One microliter of each reaction mixture was spotted on PEI cellulose TLC plates (EMD) and developed in 1 M LiCl for 60–75 min. After separation, the [γ-32P]ATP and [γ-32P]ADP spots were quantified by PhosphorImager using ImageQuant software.

**Steady-state ATPase activity**

The [ATP]-dependent steady state ATPase activity of tRNA^{Sec}-activated PSTK was measured using the NADH-coupled assay (19,20) at 37°C in 1× PSTK buffer. Decrease of absorbance due to NADH oxidation was monitored at 340 nm (ε<sub>ex</sub> = 6220 M<sup>-1</sup> cm<sup>-1</sup>) using a Lambda 20 UV/Vis Spectrometer (Perkin Elmer, Waltham, MA). Reactions were initiated by the addition of 200 nM PSTK to a solution containing 1× PSTK buffer, 1 μM tRNA^{Sec}, 100 μM pyruvate kinase, 10 μM lactate dehydrogenase, 1 mM phosphoenolpyruvate, 0.5 mM NADH and ATP ranging from 0 to 25 mM (MgCl<sub>2</sub> was increased to 25 mM in the reaction with 25 mM ATP). Background ATP hydrolysis in the absence of PSTK was subtracted from all reactions prior to analysis. Background NADH oxidation in the absence of ATP was minimal. Kaleidagraph v. 3.6 (Synergy Software) was used to calculate the kinetic parameters, k<sub>cat</sub> and K<sub>m</sub> using nonlinear regression plots of the initial velocity versus ATP concentration.

**Complementation of an E. coli ΔselA strain**

M. jannaschii PSTK and the PSTK mutant genes, G14W, K17A, S18A, T19W, D41A, R116A and R120A, were each digested from pET20b with Nde I and Blp I and cloned into the pACYCDuet-1 vector (Novagen). M. jannaschii SepSecS was cloned into pET15b (Novagen) (4). The selA deletion strain JS1 was described previously (4). M. jannaschii PSTK and M. jannaschii SepSecS were transformed into the JS1 strain to serve as a positive control. The following combinations transformed into the strain served as negative controls: M. jannaschii PSTK plus pET15b, pAC<sub>Y</sub>C<sub>D</sub>uet-1 plus M. jannaschii
SepSecS and pACYC184 plus pET15b. Each PSTK mutant was transformed into the strain plus \textit{M. jannaschii} SepSecS. Cell growth and the benzyl viologen test was performed as described previously (4).

**Sequence-based structure alignment and modeling**

The structures of T4 polynucleotide kinase (T4 Pnk) (21,22) and gluconate kinase (GntK) (23) were aligned by using Multiseq in VMD 1.8.5 (24). PSTK sequences taken from the National Center for Biotechnology Information (NCBI) nonredundant database were aligned to T4 Pnk and GntK also using Multiseq in VMD 1.8.5. Representative PSTK sequences were chosen from two archaea and four divergent eukaryotes for Figure 6. A model of \textit{M. jannaschii} PSTK was generated using GntK (PDB code 1ltq) as the template with the program Modeller 9v2 and default parameters (25).

**Phylogenetic analysis of PSTK**

Sequences of PSTK, Kti12 and additional closely related DxTN kinases were identified using BLAST and downloaded from NCBI. Alignments were performed using MUSCLE (26), and manually adjusted using the CINEMA alignment editor (27). As detailed previously (28), the phylogenies were determined using a combined maximum parsimony/maximum likelihood method with the programs PAUP* (29), which was used to generate the 1000 most parsimonious trees, and PHYML v.2.4.4 (30), which was used to find the parsimony tree of maximum likelihood and optimize both the branch lengths and topology of that tree. Local bootstrap partitions were computed with MOLPHY (31). Final trees were graphed with the DRAWTREE program in PHYLIP version 3.66 (32). Trees were calculated similarly for the full length PSTK and Kti12 alignment as well as the kinase domain and C-terminal domain alignments, except that gaps were counted as missing data in the parsimony analysis of the DxTN kinase tree.

**RESULTS**

**Establishment of a \textsuperscript{32}P|tRNA/nuclease P1 phosphotransferase assay**

The biochemical characterization of \textit{M. jannaschii} PSTK required a sensitive assay that would accurately monitor conversion of Ser-tRNA\textsuperscript{Sec} to Sep-tRNA\textsuperscript{Sec}. We thus adapted an assay that was successfully used for studies of the tRNA-dependent amidotransferases (33). It is based on digestion of \textsuperscript{32}P-end-labeled aminoacyl-tRNA with nuclease P1 and separation of the resulting aminoacyl-\([\textsuperscript{32}P]\)AMP derivatives by TLC (34).

First, we \textsuperscript{32}P-labeled the terminal 3'-AMP of the tRNA using the exchange reaction of the \textit{E. coli} CCA-adding enzyme in the presence of \([\textsuperscript{32}P]\)ATP. The radioactive tRNA was then serylated by pure \textit{M. maripaludis} SerRS. After conversion of the Ser-tRNA\textsuperscript{Sec} to Sep-tRNA\textsuperscript{Sec} by \textit{M. jannaschii} PSTK, the aa-tRNA products were digested by nuclease P1. The resulting mixture of Sep-\([\textsuperscript{32}P]\)AMP, \([\textsuperscript{32}P]\)AMP and Ser-\([\textsuperscript{32}P]\)AMP was separated on polyethyleneimine (PEI)-cellulose plates by TLC (Figure 1A), giving \(R_f\) values of 0.29, 0.50 and 0.85, respectively. This assay allows direct monitoring of the deacylation of the substrate Ser-tRNA\textsuperscript{Sec} and the product Sep-tRNA\textsuperscript{Sec} that inevitably occurs over the course of the reaction. We should note that in this study we used the transcript of the \textit{M. maripaludis} tRNA\textsuperscript{Sec}, while SerRS serylated both \textit{M. jannaschii} Ser-tRNA\textsuperscript{Sec} and \textit{M. maripaludis} Ser-tRNA\textsuperscript{Sec} to 70–80%, \textit{M. jannaschii} Ser-tRNA\textsuperscript{Sec} deacylated more rapidly than \textit{M. maripaludis} Ser-tRNA\textsuperscript{Sec}.

To verify the identity of the Sep-AMP and Ser-AMP spots the same assay was used with unlabeled tRNA\textsuperscript{Sec} and \([\textsuperscript{14}C]\)Ser (Figure 1B). The \([\textsuperscript{14}C]\) label enabled us to observe the amino acid attached to the 3'-terminal AMP of the tRNA. Thus, SerRS formed \([\textsuperscript{14}C]\)Ser-tRNA\textsuperscript{Sec} (Figure 1B, lane 1) which was then converted to \([\textsuperscript{14}C]\)Sep-tRNA\textsuperscript{Sec} by PSTK (Figure 1B, lane 2). The \([\textsuperscript{14}C]\)Ser-AMP and \([\textsuperscript{14}C]\)Sep-AMP products of nuclease P1 digestion migrated similarly to Ser-\([\textsuperscript{32}P]\)AMP and Sep-\([\textsuperscript{32}P]\)AMP, respectively, on PEI cellulose plates (Figure 1A, B). The \(R_f\) values were 0.24 for \([\textsuperscript{14}C]\)Sep-AMP and 0.84 for \([\textsuperscript{14}C]\)Ser-AMP. As a final test, we confirmed the identity of the
Sep-AMP spot by carrying out the assay with \([\gamma-32P]\)ATP, unlabeled tRNA Sec and Ser. This yielded \([\gamma-32P]\)Sep-AMP after nuclease P1 digestion (Figure 1C) which migrated with an \(R_f\) value of 0.27. While the aforementioned two assays (Figure 1B, 1C) show the migration of Ser-AMP and Sep-AMP, neither of them affords the sensitivity of the \([32P]\)tRNA/nuclease P1 assay nor allows the aminoa- cylation state to be monitored as easily.

**PSTK and its substrates**

With the \([^{32}P]\)tRNA/nuclease P1 assay (Figure 2A), we were able to determine the steady state parameters of PSTK for ATP and Ser-tRNA Sec. The assay shows that PSTK is saturable at elevated substrate concentrations and does not display cooperativity, as the curve is not sigmoidal in a nonlinear regression plot of the initial velocities versus substrate concentrations (Figure 2B). The enzyme has a \(K_m\) for ATP of 2.6 mM and for Ser-tRNA Sec of 40 nM (Table 1). These values are similar to the ones published previously for bovine PSTK, 2 mM and 21 nM, respectively (8).

**Figure 2.** Kinetics of Sep-tRNA Sec formation. (A) Representative phosphorimage of Sep-tRNA Sec formation (Sep-AMP) over time. The reaction was carried out with 2 nM PSTK, 5 mM ATP and 300 nM \([^{32}P]\)-labeled Ser-tRNA Sec. At the indicated time points reaction aliquots were quenched, digested with nuclease P1, and spotted onto PEI-cellulose TLC plates as in Figure 1A. (B) Representative plot of initial velocity versus Ser-tRNA Sec concentration (4.7–600 nM) with 2 nM PSTK. Following quantification of the intensities of Ser-[\(^{32}P\)]AMP, \([^{32}P]\)AMP and Sep-[\(^{32}P\)]AMP using ImageQuant, the concentration of Sep-tRNA Sec formed at each time point was calculated by dividing the intensity of the Sep-[\(^{32}P\)]AMP spot by the total intensity. KaleidaGraph (Synergy Software) was used to calculate kinetic parameters by nonlinear regression. Error bars represent the difference of three experiments.

**Figure 3.** Binding affinity of PSTK for Ser-tRNA Sec. (A) Representative data showing a transition in the quantity of \([^{32}P]\)-labeled Ser-tRNA Sec retained by the nitrocellulose membrane (RNA bound) to the nylon membrane (RNA free) as PSTK concentration decreases from 5 \(\mu\)M to 0.61 nM. Minimal deacylation of Ser-tRNA Sec occurred over this time frame under the described conditions. (B) Semi-log plot of the fraction of RNA bound versus the log of the concentration of PSTK.

PSTK must be able to discriminate Ser-tRNA Sec from Ser-tRNA Ser (9). We therefore determined the binding affinity of *M. jannaschii* PSTK for Ser-tRNA Sec, tRNA Sec, Ser-tRNA Sec and tRNA Ser in a double filter-binding assay (16,17) where \([^{32}P]\)-labeled tRNA transcript was incubated with increasing concentrations of PSTK, followed by binding to nitrocellulose and nylon filters (Figure 3A) (see Materials and Methods section). A semi-log plot of the fraction of Ser-tRNA Sec bound to the nitrocellulose filter versus the log of the concentration of PSTK (Figure 3B) showed that PSTK has a \(K_d\) value of 53.3 nM for Ser-tRNA Sec (Table 2), which is similar to the \(K_m\) of PSTK for Ser-tRNA Sec, 40 nM (Table 1). Interestingly, the \(K_d\) of PSTK for tRNA Sec, 39.4 nM (Table 2), is similar to the \(K_d\) of PSTK for Ser-tRNA Sec, demonstrating that the

### Table 1. Kinetic data for *M. jannaschii* PSTK

<table>
<thead>
<tr>
<th>Substrate</th>
<th>(K_m) ((\mu)M)</th>
<th>(k_{cat}) (s(^{-1}))</th>
<th>(k_{cat}/K_m) (s(^{-1}) (\mu)M(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ser-tRNA Sec a,d</td>
<td>0.04 ± 0.007</td>
<td>0.098 ± 0.005</td>
<td>2.6 ± 0.49</td>
</tr>
<tr>
<td>ATP d</td>
<td>2600 ± 1000</td>
<td>0.069 ± 0.009</td>
<td>2.7 \times 10^{-4} ± 1.1 \times 10^{-4}</td>
</tr>
<tr>
<td>ATP c,d</td>
<td>2430 ± 970</td>
<td>2.9 \times 10^{-4} ± 3.0 \times 10^{-6}</td>
<td>1.1 \times 10^{-8} ± 4.0 \times 10^{-8}</td>
</tr>
</tbody>
</table>

Steady-state kinetics of *M. jannaschii* PSTK phosphotransferase and ATPase activities are shown. See Materials and Methods section for details.

aATP (10 mM) was added in excess with 2 nM enzyme.
bSer-tRNA Sec (400–600 nM) was added in excess with 6 nM enzyme.
ctRNA Sec (1 \(\mu\)M) was added in excess with 200 nM enzyme.
dMeasurements were taken three times. Standard deviations are reported.
PhosphorImager analysis was used to quantify the intensities of the \([\gamma-\text{32P}]\text{ATP}\) and \([\gamma-\text{32P}]\text{ADP}\) spots. The minimal ATPase activity in the absence of tRNA was subtracted. Error bars represent the standard deviation of three separate experiments.

**Figure 4.** ATPase activity of *M. jannaschii* PSTK. Graph of the ratio of \([z-\text{32P}]\text{ATP}\) converted to \([z-\text{32P}]\text{ADP}\) by PSTK (200 nM) in the absence or presence of 1 µM tRNA<sup>Sec</sup>, Ser-tRNA<sup>Sec</sup>, or Ser-tRNA<sup>Ser</sup>. After incubation at 37°C for 45 min, the reactions were quenched with ice-cold EDTA and spotted on PEI-cellulose TLC plates which were developed in 1 M LiCl for 60 to 75 min. PhosphorImager analysis was used to quantify the intensities of the \([z-\text{32P}]\text{ATP}\) and \([z-\text{32P}]\text{ADP}\) spots.

---

**Table 2.** Affinity of *M. jannaschii* PSTK for tRNA<sup>Sec</sup> and tRNA<sup>Ser</sup>

<table>
<thead>
<tr>
<th>Substrate</th>
<th>(K_d) (µM)&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>tRNA&lt;sup&gt;Sec&lt;/sup&gt;</td>
<td>0.039 ± 0.003</td>
</tr>
<tr>
<td>Ser-tRNA&lt;sup&gt;Sec&lt;/sup&gt;</td>
<td>0.053 ± 0.005</td>
</tr>
<tr>
<td>tRNA&lt;sup&gt;Ser&lt;/sup&gt;</td>
<td>1.30 ± 0.14</td>
</tr>
<tr>
<td>Ser-tRNA&lt;sup&gt;Ser&lt;/sup&gt;</td>
<td>1.26 ± 0.30</td>
</tr>
</tbody>
</table>

\(<sup>a</sup>\) Dissociation constants for binding were determined by enzyme titration using *M. maripaludis* tRNA<sup>Sec</sup> and tRNA<sup>Ser</sup> in vitro transcripts. Experiments were conducted in triplicate and repeated twice. Standard deviations are reported.

As expected, PSTK has a much lower affinity for both tRNA<sup>Sec</sup> and Ser-tRNA<sup>Sec</sup> with \(K_d\) values of 1.30 µM and 1.26 µM (Table 2), respectively. Neither is a substrate for phosphorylation by PSTK (Figure 1C and data not shown) (9). PSTK recognizes both tRNA<sup>Ser</sup> and Ser-tRNA<sup>Ser</sup> with similar \(K_d\) values (Table 2), further revealing that the Ser moiety has no effect on binding affinity.

While PSTK was able to transfer the \(\gamma\)-phosphate of \([\gamma-\text{32P}]\text{ATP}\) to the Ser moiety of Ser-tRNA<sup>Sec</sup>, we proceeded to establish whether the reaction was tRNA-dependent. We attempted to phosphorylate free Ser with \([\gamma-\text{32P}]\text{ATP}\) (data not shown), followed by TLC separation as described for Figure 1. There was, however, no detectable conversion of Ser to Sep, suggesting that PSTK does not recognize free Ser as a substrate for phosphate transfer; Ser attached to tRNA<sup>Sec</sup> is its obligate substrate.

**PSTK is a tRNA<sup>Sec</sup>-induced ATPase**

Since PSTK recognizes both tRNA<sup>Sec</sup> and Ser-tRNA<sup>Sec</sup> with equal affinity, we determined if the binding of tRNA<sup>Sec</sup> would stimulate ATPase activity or if Ser-tRNA<sup>Sec</sup> was required. PSTK was assayed for ATPase activity using \([\gamma-\text{32P}]\text{ATP}\) as a substrate. The reaction product, \([\gamma-\text{32P}]\text{ADP}\) was separated from \([\gamma-\text{32P}]\text{ATP}\) by TLC on PEI-cellulose plates and quantified with a PhosphorImager. The ATPase activity was tested in the absence of tRNA or presence of tRNA<sup>Sec</sup>, Ser-tRNA<sup>Sec</sup> or Ser-tRNA<sup>Ser</sup> (Figure 4). There was little ATP hydrolysis when no tRNA or Ser-tRNA<sup>Sec</sup> was present, but addition of Ser-tRNA<sup>Sec</sup> increased the ATPase activity 20-fold over the control with no tRNA. Interestingly, addition of tRNA<sup>Sec</sup> resulted in a 4.3-fold increase in ATPase activity compared to the control with no tRNA. The steady-state ATPase activity in the presence of tRNA<sup>Sec</sup> was measured by an NADH-coupled assay, in which the production of free ADP was followed spectrophotometrically from the oxidation of NADH in a solution containing NADH, phosphoenolpyruvate, Mg<sup>2+</sup>, pyruvate kinase and lactate dehydrogenase (19,20). The \(K_m\) for ATP in the presence of tRNA<sup>Sec</sup>: 2.43 mM, is similar to the \(K_m\) for ATP, 2.6 mM, measured by the phosphotransferase assay in the presence of Ser-tRNA<sup>Sec</sup> (Table 1). The results suggest that PSTK uses an induced fit mechanism (35) where binding of one substrate (tRNA<sup>Sec</sup>) specifically induces hydrolysis of the ATP substrate similar to the mechanism of yeast hexokinase (36).

**NTP substrate preferences of PSTK**

We tested PSTK to determine activity with different nucleotides by measuring the initial velocity \((k_{\text{phospho}})\) of Ser-tRNA<sup>Sec</sup> to Sep-tRNA<sup>Sec</sup> conversion in the presence of 20 mM ATP, ITP, GTP, CTP, UTP, AMP-CPP, AMP-PCP or dATP (Figure 5). The reaction proceeded most quickly in the presence of ATP, but PSTK was able to use dATP and all the rNTPs tested to form Sep-tRNA<sup>Sec</sup>. Even pyrimidine nucleotides were used at greater than 50% activity. Phosphorylation activity was low in the presence of the ATP analog AMP-CPP, suggesting that the bridging oxygen atom between the \(z\) and \(\beta\)-phosphates may be important for NTP recognition. As expected, there was no detectible phosphorylation activity in the presence of AMP-PCP.

**Modeling of the N-terminal kinase domain of PSTK and mutation of putative functional residues**

The N-terminal region of *M. jannaschii* PSTK is homologous to a large superfamily of P-loop kinases (37). Structures complexed with ADP and ATP/Mg<sup>2+</sup> are available for T4 Pnk (21) and GntK (23), respectively. These proteins are homologous to the kinase domain of PSTK (N-terminal 150 amino acids). We generated a structure-based sequence alignment of T4 Pnk, GntK and
all available PSTK sequences, which revealed conserved motifs that are likely involved in ATP-binding and Mg$^{2+}$ coordination (Figure 6).

Both T4 Pnk and GntK have a P-loop motif, i.e. Walker A motif, GxxxxGK(T/S), which is characteristic of nucleoside triphosphate (NTP) phosphoanhydrolases and phosphotransferases (38,39). The P-loop motif functions to position the triphosphate moiety of a bound nucleotide for hydrolysis of the β-γ phosphate bond. PSTK also has a conserved P-loop motif, 11GxP(G/A)xGK(S/T)18 (in the M. jannaschii numbering) (Figure 6). An invariant aspartic acid residue in PSTK (D41) (Figure 6A) likely corresponds to the Walker B motif (DxxG), which is common to P-loop kinases (40). An additional ATP-binding motif (RxxxxR) is conserved as 116RxxG121 in PSTK (Figure 6A) and is also present in structures of chloramphenicol phosphotransferase (41), adenylate kinase (42), T4 Pnk (22) and GntK (23).

Based on the above information and the knowledge that residues Lys15, Ser16, Asp35 and Arg126 are required for 5'-kinase activity of T4 Pnk, we made mutations of key residues likely involved in ATP-binding and Mg$^{2+}$ coordination in M. jannaschii PSTK (Figure 6). The interactions of the conserved residues Gly14 and Thr19 with ATP were unlikely to be sequence-specific, thus these residues were replaced with a bulky amino acid, Trp. Single Ala mutations were made of the residues Lys17 and Ser18 in the P-loop motif, Asp41, the conserved Asp in the Walker B motif, and Arg116 and Arg120 in the RxxxxR motif.

**In vivo activity of PSTK mutants**

A previously used assay (4) was available to give an *in vivo* functional test of PSTK. *E. coli* produces the selenium-dependent formate dehydrogenase H (FDH$_H$) when grown under anaerobic atmosphere, and FDH$_H$ activity allows the cells to reduce benzyl viologen (BV) in the presence of formate (43). Co-expression of *M. jannaschii* PSTK and SepSecS in an *E. coli* selA deletion strain (JS1) in which selenoprotein production is abolished restores selenoprotein biosynthesis (4). Growth of the *E. coli* cells under anaerobic atmosphere (CO$_2$ 90%: N$_2$ 5%: H$_2$ 5%) for 24 h followed by a top agar overlay containing formate and BV resulted in violet-colored cells when BV was reduced due to the activity of the selenoprotein FDH$_H$ (4). We utilized JS1 to test the ability of our *M. jannaschii* PSTK mutants G14W, K17A, S18A, T19W, D41A, R116A and R120A with *M. jannaschii* SepSecS to complement the selA deletion. Complementation did not occur in the presence of PSTK alone, SepSecS alone, or in the presence of empty expression vectors (Figure 7). Wild-type PSTK and the T19W mutant were able to fully complement the JS1 strain, while the R116A mutant was able to complement to a lesser degree, as seen by the lighter violet colonies (Figure 7B). Mutants G14W, K17A, S18A, located within the P-loop, D41A, in the Walker B motif, and R120A, in the RxxxxR motif, were unable to complement the JS1 strain (Figure 7), demonstrating the importance of the mutated amino acids for PSTK activity in *vivo*. These *in vivo* complementation studies confirmed the importance of the P-loop, Walker B motif, and RxxxxR motif residues.

**In vitro phosphotransferase activity of PSTK mutants**

In order to demonstrate that the *in vivo* deficiencies in BV reduction were properties of PSTK, the phosphotransferase assay was used to determine *in vitro* activity of the mutant enzymes. We found that G14, K17, S18, D41 and R120 are essential constituents of the PSTK active site. All of the mutant enzyme preparations were analyzed in parallel to determine their ability to convert Ser-tRNA$_{Sec}$ to Sep-tRNA$_{Sec}$ using the $[^{32}]$P-tRNA/nuclease P1 assay. At 10 nM enzyme, the mutant enzymes displayed reduced phosphotransferase activity (Figure 8A). The T19W enzyme was 2.8-fold less active than wild-type PSTK, while the R116A enzyme was 23.5-fold less active. The remaining mutants displayed little detectable activity. To determine if these mutant enzymes exhibited any catalytic activity, we increased the time of the assay from 1 to 5 min and assayed the enzymes at a higher enzyme concentration. 100 nM (Figure 8B). However, four mutants, G14W, K17A, S18A and D41A, were grossly deficient in phosphotransferase activity at both enzyme concentrations. Sep-tRNA$_{Sec}$ formation was saturated with the T19W enzyme at 100 nM enzyme. R116A was 1.2-fold less active than wild-type, whereas R120A had a more severe effect on catalysis, with 7.9-fold less activity than wild-type enzyme.

**Phylogenetic analysis of PSTK**

PSTK is a protein composed of two domains. The N-terminal P-loop kinase domain belongs to a family of related kinases including both cellular and viral poly-nucleotide kinase, phosphoserine/phosphothreonine kinase as well as bacterial kinases of unknown function.
The second domain, which is presumed to be involved in tRNA<sup>Sec</sup> recognition in PSTK, is found only in one other group of eukaryotic proteins of which the yeast examples are named Kti12 (killer toxin-insensitive) in reference to the gene's ability to confer yeast resistance to the Kluyveromyces lactis toxin zymocin (44). Kti12 is required for the biosynthesis of certain tRNA modifications (45), but its precise function is unknown. A genetic deletion of KTI12 in yeast abolished formation of mcm<sup>5</sup>- and ncm<sup>5</sup>-modified uridines at the wobble position of some tRNAs and caused growth defects (45). Based on its close homology with PSTK, we predict that Kti12 interacts directly with its tRNA substrates. Perhaps, like PSTK, Kti12 is also a tRNA-dependent kinase.

The strong similarity of PSTK and Kti12 throughout the full-length of the molecule along with misannotations in the sequence databases led to a previous misclassification of some Kti12 proteins as PSTKs (4). In addition to presenting a resolved phylogeny of PSTK and Kti12 (Supplementary Data Figures 1 and 2), we also computed the phylogenetic root between these proteins. Such a root helps determine if Kti12 diverged from PSTK at some earlier time. A set of more distantly related (so-called out group) sequences is required to place the root. Since PSTK and Kti12 only display homology with other proteins in their kinase domain, their root must be determined from a phylogeny based only on the kinase domain (Figure 9).

This tree shows a deep separation between the PSTK/Kti12 cluster and the other kinases in their family. The evolution of PSTK and Kti12 is characterized by vertical gene flow, i.e. organismal groupings are essentially in accord with standard taxonomy, and gene loss. Kti12 is present in nearly all eukaryotes, though noticeably absent from Trypanosomatidae and Plasmodium. All non-Sec decoding eukaryotes and archaea lack PSTK. Based on current genomic data, and with our identification of PSTK in Plasmodium, we observed that all organisms that encode SepSecS also encode PSTK (see Supplementary Table 1). The Plasmodium PSTK, however, cannot be recovered with a typical BLAST search of the NCBI databases. Only by using the PlasmoDB (46) BLAST server with the kinase domain of the fly PSTK as a query and searching over the ‘protein’ database of all eukaryotic proteins was the Plasmodium PSTK recovered. The Plasmodium PSTK is shown with the predicted positions of key active site residues and docked ATP highlighted.

Figure 6. Sequence based structure alignment and modeling of the kinase domain. (A) Alignment of the structures of E. coli gluconate kinase (GntK) and Enterobacter phage T4 polynucleotide kinase (T4 Pnk) with representative PSTK sequences from two archaea and four divergent eukaryotes. Amino acids are colored according to sequence similarity (BLOSUM 50) and asterisks indicate key residues that were mutated and tested experimentally. The P-loop motif, GxxxxGK(S/T), the conserved Asp residue of the Walker B motif, and the conserved RxxxR motif that are involved in ATP-binding and Mg<sup>2+</sup> coordination in T4 Pnk and GntK are shown. Residue numbering is shown adjacent to each sequence. The structures of Pnk (B) and GntK (C) are shown, with active site residues and bound substrates highlighted. (D) A homology model of M. jannaschii PSTK is shown with the predicted positions of key active site residues and docked ATP highlighted.
Plasmodium species could a partial hit to a P. knowlesi open reading frame (ORF) be found. Four additional Plasmodium PSTKs were then readily discovered based on homology with the P. knowlesi ORF. A proper alignment of these Plasmodium sequences to other PSTKs revealed conserved PSTK active site residues and the phylogeny resulting from that alignment clearly supports the grouping of the Plasmodium PSTKs with their other eukaryotic counterparts (Figure 9).

The vertical inheritance pattern and deep phylogenetic divide between archaeal and eukaryotic PSTKs (Figure 9) shows that the protein originated before the evolutionary divergence of the domains Archaea and Eukarya. Kti12 is restricted to eukaryotes, yet it appears to have originated from a duplication of the ancestor of the archaeal and eukaryotic PSTKs. The phylogenetic relationship between PSTK and Kti12, which is also supported in phylogenies of the full-length and C-terminal domains (see Supplementary Data Figure 1 and 2), indicates that the eukaryotic protein, Kti12, evolved before the phylogenetic divergence of the archaeal and eukaryotic domains.

**DISCUSSION**

**PSTK recognition of its tRNA substrate**

PSTK not only has high affinity for its substrate, Ser-tRNA<sub>Sec</sub>, but surprisingly also binds unacylated tRNA<sub>Sec</sub> more tightly than does the human SerRS (41). Aminoacyl tRNA-synthetases typically have a higher affinity for their substrates, unacylated tRNAs, than for their products, aminoacyl-tRNAs. Why PSTK has high affinity for unacylated tRNA<sub>Sec</sub> is unclear. Mizutani and Hashimoto estimated the cytoplasmic concentration of tRNA<sub>Sec</sub> to be quite low at 50 nM (8); PSTK could possibly sequester tRNA<sub>Sec</sub> for acylation by SerRS in order to channel the tRNA from one enzyme to the next so that a ‘mis’-acylated tRNA is never available for use in protein translation. Whether these two enzymes and/or SepSecS form a complex is yet an open question. What is apparent, however, from the $K_d$ values of PSTK for tRNA<sub>Sec</sub>, Ser-tRNA<sub>Sec</sub>, tRNA<sub>Ser</sub> and Ser-tRNA<sub>Ser</sub>, is that selection for the tRNA substrate is based on the tRNA alone and not the serine moiety.

**ATP-binding site of PSTK**

PSTK catalyzes transfer of the $\gamma$-phosphate of an NTP to the hydroxyl group of Ser on Ser-tRNA<sub>Sec</sub> to form Sep-tRNA<sub>Sec</sub>. The active site of PSTK employs a P-loop motif, $^{11}GxP(G/A)xGK(S/T)_{18}$, a Walker B motif (D41), and a $^{116}RxxR$ motif for ATPase and
phosphotransferase activities. Based on our homology model of the PSTK kinase domain (Figure 6D), we predicted the roles of several key residues in these motifs: G14, K17, S18, T19, D41, R116, and R120.

In the P-loop motif, Gly18 in GntK (23), which aligns to Gly14 in *M. jannaschii* PSTK, forms a hydrogen bond with the bridging oxygen of the γ-phosphate of ATP in a non-sequence dependent manner. We perturbed this...
essential contact with a G14W mutation that caused the loss of both in vivo and in vitro activities. Conserved Lys and Ser residues in the P-loops of T4 Pnk and GntK are essential for ATP-binding. These residues make specific contacts with the β- and υ-phosphates of ATP (21,22). In GntK, the side-chain hydroxyl on this serine residue (Ser22) also coordinates the Mg²⁺, which is bound to the substrate ATP (23). Similarly, the corresponding residues in M. jannaschii PSTK, Lys17 and Ser18, were required for in vivo activity and separate alanine mutants displayed minimal in vitro activity. The backbone nitrogen of Thr17 in T4 Pnk and Ala23 in GntK interact with the α-phosphate of ATP, and our model of PSTK suggests a similar role for Thr19. Mutation of this residue to Trp resulted in partially active enzyme.

The Walker B motif is less clearly defined than the P-loop, but it must contain a negatively charged residue (typically Asp) that is involved in binding Mg²⁺ (47). GntK (23) and T4 Pnk (22) also have a conserved Asp at a homologous position. Even though this residue is distant from the active site, it contributes to Mg²⁺ binding via a water-mediated hydrogen bond, as observed in GntK (23) and T4 Pnk (Figure 6C). Mutation of the corresponding Asp (D41A) in PSTK drastically reduced activity of the enzyme. These data support our prediction that D41 is the essential feature of the Walker B motif in PSTK, and it likely has a similar role to the homologous Asp residues in T4 Pnk and GntK.

In both Pnk and GntK, the proximal Arg of the RxxxR motif forms a cation–π interaction with the adenylyl's conjugated rings. The GntK structure suggests that the distal Arg interacts with the α-phosphate of ATP, and the T4 Pnk structure suggests that this Arg interacts with the ester linkage between the 5'-hydroxyl of the ribose and the α-phosphate of ADP after ATP hydrolysis. In PSTK, mutation of the distal Arg (R120A) had a more severe effect on catalysis than mutation of Arg116. This indicates that in PSTK the cation–π interaction of Arg116 with the ATP substrate is less important than the specific interaction of Arg120 with the α-phosphate of ATP.

In summary, we experimentally determined that the Gly14, Lys17, Ser18, Asp41, Arg116 and Arg120 residues are important or required for enzyme activity both in vitro and in vivo, suggesting that they make contacts in ATP-binding and Mg²⁺ coordination similar to those observed in the crystal structures of T4 Pnk and GntK.

**NTP preference of PSTK**

Interestingly, the binding site for the phosphate donating NTP exhibits little specificity in PSTK. T4 Pnk is also able to utilize multiple NTPs (48), which is supported by the crystal structure that shows the contacts between the enzyme and ADP are mostly through interactions with the phosphate oxygens (21). Crystal structures of other P-loop proteins show similar contacts (23,49). The broad occurrence of relaxed NTP specificity for P-loop kinases is unknown. Some P-loop NTPases have higher specificity for a particular phosphate donor, e.g. GTP was preferred 100-fold over other NTPs by yeast tRNA ligase (50). In addition to a well-conserved Walker B motif (DxxG), an evolutionary study of P-loop GTPases and GTPase-related proteins found a NKxD motif is linked to GTPase specificity (51). Though we do not know the physiological phosphate donor for PSTK, there is a preference for ATP in vitro. Characterizing PSTK as an NTP generalist may allow further investigation into the molecular basis of NTP specificity.

**Insights into PSTK catalysis**

P-loop-containing enzymes typically either catalyze their reactions by a proton-abstracting acidic residue or the γ-phosphate of the NTP phosphate donor acts as a base in substrate-assisted catalysis (37). Our observation that PSTK has little detectable ATPase activity in the absence of tRNA⁹, and that this activity is greatly stimulated in the presence of Ser-tRNA⁹ was thus not unexpected. The Ser hydroxyl may be the attacking nucleophile, possibly activated by abstraction of a proton. D35 in T4 Pnk was identified as a potential proton-abstracting residue (22), and D41 in M. jannaschii PSTK aligns to this residue (Figure 6). Our D41A enzyme was severely deficient in both phosphotransferase and ATPase activities in vitro and unable to complement the E. coli selA deletion strain in vivo. Whether this was due to a role for D41 in Mg²⁺ coordination as discussed previously or required as a base for proton-abstraction requires further analysis.

**Evolutionary history of PSTK**

PSTK and SepSecS complete the archaeal and eukaryotic pathway for Sec-tRNA⁹ formation. The more detailed phylogeny presented here resolves evolutionary idiosyncrasies observed previously for PSTK (4) and shows that PSTK exactly co-evolved with SepSecS. This pathway was vertically inherited over the course of its evolution which originated before the archaeal-eukaryotic phylogenetic divide, though how long before will remain an unresolved issue.

The phylogenetic analysis also identifies an evolutionarily bona fide PSTK in the Plasmodium species. These parasites are known for their accelerated evolutionary clock (52), and PSTK is no exception. The Plasmodium PSTKs have evolved at a significantly faster rate (note their long branch in Figure 9) than their counterparts in other eukaryotes. Indeed, the Plasmodium PSTKs evolved at such a fast rate that even detecting these PSTKs by sequence homology is rather difficult, explaining why this gene has not been identified previously or properly annotated.

Kti12, which is the only known protein that is homologous to the full-length of PSTK, is an eukaryotic signature gene that evolved before the archaeal-eukaryotic phylogenetic divide. The signature gene notion was developed by Woese and colleagues as they sought to understand how the Archaea are so distinct from Eukarya and Bacteria. They defined the archaeal genomic signature as ‘the set of genes that function uniquely within the archaeal lineage’ (53). Signature genes also exist in Bacteria and Eukarya. Hartman and Fedorov explored the eukaryotic signature in detail (54), but limited their set of signature genes to those without homologs in Archaea.

---

**Notes:**

1. Hartman and Fedorov explored the eukaryotic signature in detail.
2. The Plasmodium PSTKs have evolved at a significantly faster rate.
3. This pathway was vertically inherited over the course of its evolution.
4. The Plasmodium PSTKs evolved at such a fast rate.
5. The signature gene notion was developed by Woese and colleagues.
6. They defined the archaeal genomic signature as ‘the set of genes that function uniquely within the archaeal lineage’.
7. Signature genes also exist in Bacteria and Eukarya.
8. Hartman and Fedorov explored the eukaryotic signature in detail.
or Bacteria. This study identified genes involved in eukaryotic specific cellular structures including the cytoskeleton, nucleus, internal membranes and signaling as well as eukaryotic specific features of protein synthesis and degradation. Although this study produced an admittedly minimal set of eukaryotic signature genes, it is clear that the eukaryotic signature is broader than previously thought including proteins, which are involved in more basic cellular functions, such as tRNA modification in the case of Kti12.

Since Kti12 is restricted to eukaryotic genomes and because it is closely related to PSTK one might speculate that Kti12 is a duplication of the eukaryotic PSTK. The inclusion of related kinases in the phylogenetic analysis of PSTK and Kti12 (Figure 9) allows the determination of a root between PSTK and Kti12, which leads to the conclusion that Kti12 diverged from the archaeal-eukaryotic ancestral PSTK. Two points of fundamental significance for understanding the evolution of the eukaryotic cell type are brought to light in the PSTK-Kti12 phylogeny: (i) many more eukaryotic signature genes exist with roles in a broader range and perhaps in more basic cellular functions than previously recognized and (ii) lineage defining signature genes can evolve before the phylogenetic split of that lineage from its sister lineage. Archaeal and bacterial signature genes have been shown to predate the phylogenetic divide between the domains Archaea and Bacteria (28) and now we document this phenomenon in eukaryotes.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS
We thank Kelly Sheppard, Jeff Sabina, Jing Yuan, Sotiria Palioura, Adrian Olivares, Lennart Randau and Markus Engler for technical advice and many discussions, and Joanne Ho for her enthusiastic support. We thank Enrique De La Cruz for the use of the Perkin Elmer Lambda 20 UV/Vis Spectrometer in his laboratory. R.L.S. holds a Ruth L. Kirschstein National Research Service Award F32 GM075602 from the National Institute of General Medical Sciences, and P. O’D. is a National Science Foundation postdoctoral fellow in Biological Informatics. This work was supported by grants from the National Institute of General Medical Sciences and the Department of Energy (to D.S.). Funding to pay the Open Access publication charges for this article was provided by National Institute of General Medical Sciences grant GM22854 (to D.S.).

Conflict of interest statement. None declared.

REFERENCES