Substrate Recognition and Mechanism Revealed by Ligand-Bound Polyphosphate Kinase 2 Structures - Supporting Information Appendix

Alice E. Parnell\textsuperscript{a,1}, Silja Mordhorst\textsuperscript{b,1}, Florian Kemper\textsuperscript{c}, Mariacarmela Giurrandino\textsuperscript{a}, Josh P. Prince\textsuperscript{a}, Nikola Schwarzer\textsuperscript{d}, Alexandre Hofer\textsuperscript{d}, Daniel Wohlwend\textsuperscript{d}, Henning J. Jessen\textsuperscript{d,e}, Stefan Gerhardt\textsuperscript{c}, Oliver Einsle\textsuperscript{e}, Petra C. F. Oyston\textsuperscript{f,g}, Jennifer N. Andexer\textsuperscript{b,2}, Peter L. Roach\textsuperscript{a,f,2}

\textsuperscript{a}Chemistry, University of Southampton, Southampton, Hampshire, SO17 1BJ, United Kingdom.

\textsuperscript{b}Institute of Pharmaceutical Sciences, Albert-Ludwigs-University Freiburg, Albertstr. 25, 79104 Freiburg, Germany.

\textsuperscript{c}Institute of Biochemistry, Albert-Ludwigs-University Freiburg, Albertstr. 21, 79104 Freiburg, Germany.

\textsuperscript{d}Organic Chemistry Institute, University of Zürich (UZH), Winterthurerstr. 190, 8057 Zürich Switzerland.

\textsuperscript{e}Institute of Organic Chemistry, Albert-Ludwigs-University Freiburg, Albertstr. 21, 79104 Freiburg, Germany.

\textsuperscript{f}Institute for Life Sciences, University of Southampton, Southampton, Hampshire, SO17 1BJ, United Kingdom.

\textsuperscript{g}Biomedical Sciences,Dstl Porton Down, Salisbury, Wiltshire, SP4 0JQ, United Kingdom.

\textsuperscript{1}These authors contributed equally to this work.

\textsuperscript{2}To whom correspondence should be addressed: J.N.A. (jennifer.andexer@pharmazie.uni-freiburg.de) or P.L.R. (plr2@soton.ac.uk)
SI Materials and Methods

Chemicals. Unless otherwise stated, chemicals and reagents were purchased in the highest purity available.

Synthetic procedures. β,γ-Methylene adenosine 5'-triphosphate (AMPPCP) was prepared by a literature procedure (1). Adenosine 5'-pentaphosphate (ASP) was synthesized by a modification of the method of Cremosnik (2). Adenosine 5'-tetraphosphate (with 1.6 equivalents of tetra-N-butyl ammonium as counterion) (50 mg, 51 μmol) was co-evaporated three times with dry toluene (1 mL) and dissolved in dry DMF (1 mL) under an atmosphere of dry nitrogen. The mixture was cooled to 0 °C and (iPr)₂NP(OFm)₂ (bis((9H-fluoren-9-yl)methyl) diisopropylphosphoramidite, 40 mg, 77 μmol) and 5-ethylthio-1H-tetrazole (13 mg, 102 μmol) were added. The mixture was stirred at 0 °C for 10 min, then the coupling product was oxidized with mCPBA (meta-chloroperoxybenzoic acid, 77%, 23 mg, 102 μmol) at 0 °C for 5 min. Piperidine (50 μL) was added and the mixture was stirred for 30 min. The product was precipitated by addition of diethyl ether. The suspension was centrifuged and the pellets washed with diethyl ether and dried under high vacuum yielding the crude product. Purification by strong anion exchange chromatography on Q-Sepharose Fast Flow using increasing concentration of an ammonium bicarbonate eluent (product eluted at 0.3-0.5 M NH₄HCO₃) afforded pure ASP (as the ammonium salt, 12 mg, 16 μmol, 31%) as a colorless solid.

1H NMR (400 MHz, D₂O) δ 8.57 (s, 1H), 8.30 (s, 1H), 6.18 (d, J = 6.1 Hz, 1H), 4.83 (s, 1H), 4.61 (s, 1H), 4.44 (s, 1H), 4.37 – 4.20 (m, 2H). 31P[1H] NMR (162 MHz, D₂O) δ -8.5 (br, 1P), -10.2 (d, J = 17.2 Hz, 1P), -20.6 (br, 2P), -21.00 – -21.4 (m, 1P). 31P NMR (162 MHz, D₂O) δ -8.5 (br, 1P), -10.2 (d, J = 16.5 Hz, 1P), -20.6 (br, 2P), -21.0 – -21.4 (m, 1P). 13C[1H] NMR (126 MHz, D₂O) δ 155.3, 152.4, 149.1, 140.1, 118.7, 86.8, 84.1 (d, J = 9.0 Hz), 74.3, 70.5, 65.4 (d, J = 5.2 Hz). HRMS (ESI) m/z: calc.: 332.45693, found: 332.45745 [M-H]⁻.

Cloning, heterologous expression and protein purification. FtPPK2 was produced and purified as described previously (3). F. tularensis ppk2 site directed mutants were generated with a modified QuikChange protocol (4) using the primers listed in Table S3. After DpnI digestion, chemically competent XL1-Blue cells were transformed with the PCR mixture, transferred onto plates containing ampicillin (100 μg/mL) and incubated overnight (37 °C). Selected colonies were then cultured overnight in 2YT media (10 mL) with ampicillin. Plasmid DNA isolated from these cultures was sequenced to confirm the introduction of the required mutations. For protein production, E. coli BL21
Rosetta plyS (DE3) cells were transformed with purified mutated plasmids and the variant proteins isolated as described for the wild type (WT) enzyme (3) (Fig. S14).

Inserts for *mrppk2* WT and Asn121Asp were obtained as synthetic codon-optimized DNA from Invitrogen. These were amplified by PCR using the primers *MrPPK2*-Ndel and *MrPPK2*-HindIII (Table S3), digested with Ndel and HindIII, and ligated (T4-Quick Ligase) into similarly digested expression vector pET28a(+) (Novagene). After transformation in chemically competent *E. coli* DH5α cells, the plasmids (pET28a-MrPPK-wt and pET28a-MrPPK2-Asn121Asp) were isolated and verified by sequencing.

Proteins were heterologously produced in *E. coli* BL21(DE3). LB medium (5 mL) containing kanamycin (50 μg/mL) was inoculated with a single colony of a fresh transformation and incubated at 37 °C overnight. This culture was used as a 1% inoculum in LB (500 mL) with kanamycin (50 μg/mL) and cells were cultured at 37 °C and 180 rpm up to an optical density (600 nm) of 0.5. At this point, protein production was induced by addition of IPTG (final concentration 0.2 mM) and the culture was maintained at 20 °C and 180 rpm for a further 24 h. Cells were harvested by centrifugation and stored at -20 °C until use.

For *MrPPK2* purification (Fig. S1), cells were resuspended in lysis buffer (10 mL, 40 mM Tris-HCl, pH 8.0, 100 mM NaCl, 10% glycerol) and sonicated (Branson-Sonifier 250 R, Duty Cycle 60%, Intensity 60%, 3 x 30 s, 30 s rest) on ice. The lysate was cleared by centrifugation (3-30K Sigma, rotor 12158-H, 14811 x g, 4 °C, 45 min). The supernatant was applied to a Ni²⁺-NTA column (4 mL, Sepharose beads, Qiagen). The column was washed stepwise with 5 mL lysis buffer containing 5, 10, 20 and 50 mM imidazol, the proteins were eluted with imidazol concentrations from 100-500 mM and concentrated (Sartorius Vivaspin 20, cut off 30 kDa). For desalting, the proteins were applied to a PD-10 desalting column (GE Healthcare) and eluted with lysis buffer. The proteins were frozen in liquid nitrogen and stored at -20 °C.

**Enzyme activity assays.** The activity of *FtPPK2* was measured in reaction mixtures (1 mL) that contained final concentrations of 50 mM Tris/ HCl (pH 8.0), 25 mM NaCl, 20% glycerol, 10 mM MgCl₂, 80 mM NH₄SO₄, 1 mM ADP, 0.1 mM polyP (average length 25 units, Merck chemicals), 100 nM of enzyme for variant enzymes or 10 nM of WT enzyme) and were initiated by the addition of polyP. Reactions were incubated at 30 °C and, after 30, 45, 60 and 120 min (or 30, 60, 90 and 120 s for the WT enzyme) aliquots (100 μL) were withdrawn and the reaction quenched by the addition of EDTA (to 50 mM), followed by heating (95 °C, 5 min). Negative control assays had no polyP included. Precipitated protein was removed by centrifugation (13,200 x g, 5 min) and the supernatant was analyzed by ion pairing HPLC (Table S5). The time course of product formation was fitted to a linear
function to give the initial rates of reaction using GraphPad Prism which were then used to calculate specific activity and $k_{cat}$ values. This assay method was modified to test AMP and GMP as substrates for $FtPPK2$, but no turnover to form either di- or triphosphate nucleotides was observed.

The activity of $MrPPK2$ was measured in reaction mixtures (500 $\mu$L) that contained 50 mM Tris-HCl, pH 8.0, 20 mM MgCl$_2$, 2 mM nucleotide and 5 mM polyP (Acros organics, calculated as single phosphate residues). The reaction was maintained at 37 $^\circ$C and initiated by the addition of the enzyme (final concentration 1 $\mu$M) and stopped by removing the enzyme using spin-filters (Sartorius Vivaspin 500, cut off 10 kDa). The flow-through was analyzed by HPLC (Table S5). Kinetic parameters were determined for WT $MrPPK2$ and $MrPPK2$ Asn121Asp variant using 0 – 5 mM nucleotides and the reaction was initiated by addition of enzyme at the following concentrations: phosphorylation of ADP to ATP, 1 $\mu$M $MrPPK2$ (WT or Asn121Asp); phosphorylation of AMP to ADP, 10 nM $MrPPK2$ (WT or Asn121Asp). At selected time points, aliquots (100 $\mu$L) were withdrawn and the reaction was quenched by addition of acetonitrile (for the phosphorylation of ADP to ATP) or by addition of EDTA (for the phosphorylation of AMP to ADP). After centrifugation (Eppendorf Minispin, rotor F45-12-11, 12044 x g, 5 min), the supernatant was analyzed by HPLC (Table S5).

As an additional control, rates of reaction for both WT enzymes and the variants $FtPPK2$ Asp117Asn and $MrPPK2$ Asn121Asp were analyzed using the method initially developed for the other enzyme, yielding consistent results.

**Isothermal titration calorimetry (ITC).** All experiments used a MicroCal iTC200 (MicroCal, Inc.) maintained at 20 $^\circ$C and stirring at 750 rpm. Experiments were carried out in triplicate using ITC experimental buffer (50 mM HEPES, pH 8.0, 10 mM MgCl$_2$, 0.3 M NaCl, 20% glycerol, 0.15 mM $\beta$-mercaptoethanol). $FtPPK2$ variants were exchanged into this buffer by dialysis and/or size exclusion chromatography. Titrations began with an initial injection of 0.4 $\mu$L, followed by 19 identical injections of 2 $\mu$L. Data were fitted to a bimolecular binding model using Microcal Origin software. As a control and to ensure consistency of results, a titration with WT enzyme was carried out alongside experiments with $FtPPK2$ variant enzymes.

**Crystallization and structure determination.** The crystals of $FtPPK2$ complexes were all grown by hanging drop vapor diffusion at 20 $^\circ$C. The precipitant well solution for the complex of $FtPPK2$ with polyP (PDB ID 5LL0) was developed from an original condition from the Morpheus screen (5) and consisted of 5% glycerol/PEG 4000, 0.1 M MES/imidazole pH 6.5, 0.2 M Morpheus alcohols and crystals appeared after 2 days. The $FtPPK2$ solution (15 mg/mL) contained 5 mM MgCl$_2$ and 1 mM polyP. The precipitant well solution for $FtPPK2$ complexed with polyP and AMPPCPPPP (PDB ID 5LLB)
consisted of 20% PEG1500 and pH 8 PCTP (propionate, cacodylate, bis-tris propane) buffer. For crystallization experiments, the protein solution contained 10 mM MgCl₂, 0.5 mM polyP and 5 mM AMPPCH₂P and crystals appeared after 5 days. The precipitant well solution for FtPPK2 D117N variant complexed with polyP (PDB ID 5LLF) contained 32.5% MPD/PEG1000/PEG3350, 0.1 M Bicine/Tris, pH 8.5, 0.09 M NPS (nitrate, phosphate, sulfate mix) and the protein solution contained 0.5 mM polyP and 2.5 mM AMPPCH₂P and crystals appeared after 5 days. During harvesting, crystals were dipped into a solution of mother liquor adjusted to 50% glycerol and flash frozen in liquid nitrogen. Data sets were collected under cryogenic conditions (-173 °C) at the Diamond Light Source (Didcot, UK) on beamlines i24 (5LLB), i04 (5LLF) and i04-1 (5LL0). Data were processed with xia2 (6) and the structures solved by molecular replacement with either BALBES (7, 8) or PHASER (9). Models were built with COOT (10) and refined with Phenix.refine (11); ligand files were generated with ELBOW (12).

MrPPK2 (WT and variant) was crystallized by the sitting drop vapor diffusion method. In the optimized crystallization conditions, the drops consisted of 0.3 µL protein solution with a concentration of 10 mg/mL and 0.3 µL of reservoir solution, which contained 100 mM Tris/HCl pH 8.5, 300 mM lithium sulfate and 27% PEG 3350. Crystals appeared within 24 h and were transferred into a cryo-solution which differed by containing an increased PEG content of 30%. For soaking, this solution additionally contained 10 mM of MgCl₂, 10 mM polyP and 10 mM of one of the nucleotides (AMP, ADP, ATP, or ASP). The crystals were incubated for up to one hour before being flash-frozen in liquid nitrogen. X-ray data was collected at the beamlines X06SA (PXI) and X06DA (PXIII) of the Swiss Light Source (Switzerland). The datasets were processed using XDS (13) and aimless (14, 15) from the ccp4-suite (16). The structure of the unliganded protein was determined by molecular replacement with MolRep (17) using a PPK2 protomer from Arthrobacter aurescens (3RHF) as initial search model. The resulting model was refined with Coot (10) and autoBuster (18). The datasets from the ligand-soaked and variant crystals were solved using molecular replacement with the unliganded protein structure and then refined as described above.

Analysis of oligomerization states by means of Right Angle Static Light Scattering (RALS).

Oligomerization states were analyzed by size exclusion chromatography (SEC) combined with right angle light scattering (RALS) using a Viscotek TDA 305 (Triple Detector Array) with a GPCmax sample delivery tool. Instrument constants were determined with defined concentrations of BSA. MrPPK2 and FtPPK2 alone or mixed with substrates were analyzed in aliquots of 100 µl each at a sample concentration of 4 mg/ml on a Superdex 200 10/300 GL column (GE Healthcare, Germany) at a column temperature of 28 °C and a flow rate of 0.5 mL/min. MrPPK2 was analyzed in the presence of 40 mM Tris/HCl (pH 8.0), 100 mM NaCl, and 10% (v/v) glycerol, while FtPPK2 was analyzed in 50 mM Tris/HCl
(pH 8.0), 500 mM NaCl, and 20% (v/v) glycerol. UV<sub>280</sub>, refractive index (RI), and RALS were monitored. For determination of molecular weights the RI and RALS signals were used in the Rayleigh equation (1) as implemented in the Viscotek software OmniSEC:

\[
\frac{Kc}{R_\theta} = \frac{1}{M_w P_\theta} + 2A_2 \cdot c
\]

with \( c \) = concentration of the macromolecule

\( R_\theta \) = Rayleigh quotient of the macromolecule at incident light angle \( \theta \)

\( M_w \) = weight-average of molecular mass:

\[
\bar{M}_w = \frac{\sum M_i c_i}{c_i}
\]

\( P_\theta \) = particle scattering function \( R_\theta / R_0 \) at angle \( \theta \)

\( K \) = optical constant

\( A_2 \) = second virial coefficient

\( K \) in eq. (1) is defined as

\[
K = \frac{2\pi^2 n_0^2}{N_A \lambda_0^2} \left( \frac{dn}{dc} \right)^2
\]

with \( n_0 \) = refractive index of solvent

\( n \) = refractive index of scatterer

\( N_A \) = Avogadro constant

\( \lambda_0 \) = wavelength of incident LASER light

\( P_\theta \) in eq. (1) is defined as

\[
\frac{1}{P_\theta} = 1 + \frac{16\pi^2 n_0^2 \lambda_0^2}{3x_0^2 \sin^2 \left( \frac{\theta}{2} \right)}
\]

with \( R_0 \) being the molecule’s radius of gyration

Samples that appear to contain more than one component in the 280 nm UV trace were fitted to a sum of Gaussian and Lorentzian functions using the program MagicPlot (Magicplot Systems, LLC, Saint Petersburg, Russia). The plots are shown in Fig. S12 and results are reported in Table S4.

**Accession codes.** Protein sequences were obtained from the Uniprot Database under accession codes Q5NEQ5 (Q5NEQ5_FRATT) for FtPPK2 and M9XB82 (M9XB82_MEIRD) for MrPPK2 (Fig. 2). Coordinates have been deposited in the Protein Data Bank under accession codes 5LL0, 5LLB, and 5LLF for FtPPK2 structures and 5LC9, 5LCD, 5LDB, 5LD1, SMAQ, 5O6K, and 5O6M for MrPPK2 structures (Table S1).
**Fig. S1.** Alignment of selected PPK2 sequences. Sequences were aligned with Multalin (19) and the figure prepared with Jalview (20). Individual residues are color-coded as follows: phosphate binding in orange, metal binding in pink, nucleotide-binding in red, other residues from *Ft* and *Mr* PPK2s are highlighted in green and blue, respectively. The secondary structure elements are labeled below the alignment, with α-helices in red and β-sheets in green. The Walker motifs, the lid loop and the base-binding loop are indicated by the letters A, B, L and C, respectively, below the alignment. For other sequence positions, increasingly conserved residues are colored deeper shades of blue. The sequence names and residue numbers of the PPK2 classes I, II and III are grouped on green, pink and blue backgrounds respectively. The sequences were selected from the UniProt or NCBI databases and are listed with organism name and accession codes: Class I: Bups_1, *Burkholderia pseudomallei*, Q63KT2; Vich_1, *Vibrio cholera*, C3LSY1; Mytu_1, *Mycobacterium tuberculosis*, Q05877; FtSc_1, *Francisella tularensis*, Q5NEQ5; Syme_1, *Sinorhizobium meliloti*, Q92SA6. Class II: Rhru_2, *Rhodospirillum rubrum*, Q2RXW7; Psae_2, *Pseudomonas aeruginosa*, Q9HYF1; Acjo_2, *Acinetobacter johnsonii*, Q83XD3; Stau_2, *Staphylococcus epidermidis*, A0A0H2VI9; Baan_2, *Bacillus anthracis*, NP_844244.1. Class III: Arau_3, *Arthrobacter aurescens*, YP_948522.1; Bups_3, *Burkholderia pseudomallei*, Q63YJ7; Meru_3, *Meiothermus ruber*, M9XB82; Labr_3, *Lactobacillus brevis*, YP_795494.1; Nopu_3, *Nostoc punctiforme*, B2J4S4.
**Fig. S2.** Comparing *Ft* and *Mr* PPK2 structures. Mg⁡²⁺ (large spheres) and ligand phosphorus atoms colored by structure, blue (*MrPPK2*) and green (*FtPPK2*). (A) Overlay of *FtPPK2:AMPPCPPP:PolyP₆* (green) and *MrPPK2:ADP:PPᵢ* (blue) complexes. The root mean square deviation (RMSD) over 1304 atoms was 1.619 Å. (B) Overlay of *FtPPK2:PolyP₉* complex (polyP colored green/yellow) and *MrPPK2 Asn121Asp* variant complexed with ATP (blue) and phosphate ions (orange/red). The four phosphate ions positioned in close proximity to the polyP binding channel are labeled with a P. (C) Magnesium ion (green) coordination and distances in the *FtPPK2:AMPPCPPP:PolyP₆* complex, protomer C. Distances are shown in Å.
**Fig. S3.** Nucleotide binding in PPK2 structures. (A) Simulated annealing composite omit map ($2F_o-F_c$) around the AMPPCP PPP nucleotide in complex with *FtPPK2*, contoured at 1.5$\sigma$. (B) Simulated annealing composite omit map ($2F_o-F_c$) around the ADP and pyrophosphate (PP)i in complex with *MrPPK2*, contoured at 1.0$\sigma$. (C) Structure of *MrPPK2* with two ATP molecules bound. Two protomers A and B are shown, each with two bound ATP molecules labeled ATP1 and ATP2. The two ATP molecules associated with each protomer are labeled ATP1 and ATP2 followed by a subscript letter corresponding to the protomer (A or B). (D) Detail of ATP interactions with *MrPPK2*. ATP1 (yellow) is bound in active site of protomer A (blue) similarly to other nucleotides (e.g., AMP and ADP); ATP2 (green) binds in a different conformation that is not observed in other complexes, with the phosphates bound to the Mg$^{2+}$ (green sphere), but the nucleoside moiety protruding out of the active site and into the space between protomer A and protomer B (pink).
Fig. S4. Structural comparison of Ft and Mr ligand bound structures (PDBs: 5LLB and 5MAQ respectively) with unliganded PPK2s. Structures are colored as follows: Ft, pale green; Mr, pale blue; S. meliloti (Sm), magenta; P. aeruginosa (Pa), pink; A. aurescens (Aa), yellow. The lid loops and magnesium ions of Ft and Mr are bright green and bright blue, respectively. (A) Overlay of class I enzymes from Ft and S. meliloti (PDB 3CZQ). (B) Overlay Ft (class I) and P. aeruginosa, (class II, PDB 3CZP). (C) Overlay of Mr (class III) and P. aeruginosa, (class II, PDB 3CZP). (D) Overlay of Mr (class III) and A. aurescens, (class III, PDB 3RHF.).
Fig. S5. Kinetic analysis of MrPPK2 wild type and sequence variants.
Fig. S6. Time course of AMP conversion into ADP and ATP with MrPPK2. 1 μM MrPPK2 was incubated with 1 mM AMP and 5 mM PolyP (calculated as single phosphate residues) at 37 °C as described in the Experimental Methods. At selected time points, aliquots were withdrawn and the reaction was quenched by addition of EDTA, centrifuged, and the supernatant was analyzed by HPLC. Nucleotide concentrations were determined based on HPLC calibration curves for the corresponding substances (24). Since no internal standard was added, the initial concentration of AMP (1 mM) was used as a relative reference; the sum of all nucleotide concentrations amounts to 1 mM (= 100%) at each selected time point. All reactions were performed and analyzed in triplicates.
A. MrPPK2

![Graph A](image1.png)

<table>
<thead>
<tr>
<th>Peak</th>
<th>Ret Vol</th>
<th>Mw</th>
<th>Mn</th>
<th>Mw/Mn</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13.930</td>
<td>68.989</td>
<td>99.120</td>
<td>1.003</td>
</tr>
<tr>
<td></td>
<td></td>
<td>99.153</td>
<td>99.185</td>
<td></td>
</tr>
</tbody>
</table>

B. MrPPK2 plus polyP

![Graph B](image2.png)

<table>
<thead>
<tr>
<th>Peak</th>
<th>Ret Vol</th>
<th>Mw</th>
<th>Mn</th>
<th>Mw/Mn</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10.115</td>
<td>509.276</td>
<td>528.607</td>
<td>1.002</td>
</tr>
<tr>
<td>2</td>
<td>10.705</td>
<td>405.890</td>
<td>382.419</td>
<td>1.002</td>
</tr>
<tr>
<td>3</td>
<td>11.713</td>
<td>259.791</td>
<td>251.477</td>
<td>1.001</td>
</tr>
<tr>
<td>4</td>
<td>13.308</td>
<td>150.579</td>
<td>150.237</td>
<td>1.000</td>
</tr>
</tbody>
</table>
**Fig. S7.** Size exclusion chromatography (SEC) and right angle light scattering (RALS) analysis of MrPPK2. For each panel the following data is presented: top, run parameters; middle: chromatogram including Refractive Index (RI) and RALS plots with integral limits and according baselines for integration; bottom: plot of the molecular weight distribution across the peak and the according experimentally determined parameters. Mp is the molecular weight at the peak maximum, Mn is the number average, Mw the weight average, and Mz the z-average of the molecular mass. Mw/Mn is the polydispersity of the peak. Panels: (A), MrPPK2, (B), MrPPK2 pre-incubated with 10 mM polyP.
A. FtPPK2

B. FtPPK2 plus polyP

<table>
<thead>
<tr>
<th>Peak</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mp</td>
<td>66.741</td>
<td>40.416</td>
</tr>
<tr>
<td>Mn</td>
<td>63.317</td>
<td>38.133</td>
</tr>
<tr>
<td>Mw</td>
<td>63.979</td>
<td>38.203</td>
</tr>
<tr>
<td>Mz</td>
<td>64.635</td>
<td>38.271</td>
</tr>
<tr>
<td>Mw/Mn</td>
<td>1.010</td>
<td>1.002</td>
</tr>
</tbody>
</table>
Fig. S8. Size exclusion chromatography (SEC) and right angle light scattering (RALS) analysis of FtPPK2. For each panel the following data is presented: top, run parameters; middle: chromatogram including Refractive Index (RI) and RALS plots with integral limits and according baselines for integration; bottom: plot of the molecular weight distribution across the peak and the according experimentally determined parameters. Mp is the molecular weight at the peak maximum, Mn is the number average, Mw the weight average, and Mz the z-average of the molecular mass. Mw/Mn is the polydispersity of the peak. Panels: (A), FtPPK2, (B), FtPPK2 pre-incubated with 10 mM polyP.
Fig. S9. Analysis of size exclusion chromatography. The experimental UV absorption traces are shown in black and were fitted to a sum function (made up of up to 4 Gaussian or Lorentzian functions) is shown in red. Curves for the individual functions are shown in green, blue, orange and pink respectively.
**Fig. S10.** Model for the oligomerization of PPK2s in the crystalline state and in solution. This includes the effects of polyP (curved red line) for the *Ft* (A) and *Mr* (B) enzymes.
Fig. S11. Kinetic analysis of FtPPK2 wild type and sequence variants.
Fig. S12. Mechanistic hypothesis for reaction with AMPPCP and polyP. After binding at the enzyme active site 1, nucleophilic attack by polyP leads to a trigonal bipyramidal intermediate 2. The CH₂ containing phosphonate analogue is a poor leaving group (unlike the natural leaving group, ADP), but loss of water leads to compound 3, in which the AMPPCP has become covalently linked to the polyP. Subsequent hydrolysis of the polyP chain (which may occur during the crystallization and may not directly relate to physiological catalysis) leads to the bound AMPPCPPPP 4. This figure summarizes one of several putative mechanisms that might account for the observed structure. Other alternatives include transfer of pyrophosphate by a direct substitution mechanism, or sequential transfer of two individual phosphate groups.
**Fig. 13.** Comparing *FtPPK2* and thymidylate kinase (TK). *FtPPK2* (green) complexed with polyP (P6) and AMPPCP (phosphorus atoms, green) is overlaid with human TK complexed with an ATP analogue, AppNHp and a thymidine analogue, 3-fluoro-3-deoxythymidine monophosphate (FLTMP) (phosphorus atoms, orange, PDB ID: 1NMY). Spheres indicate the positions of the magnesium ions. The highly similar P-loop structures, with Lys66/19 are shown in green and brown for *FtPPK2* and human TK respectively.
**Fig. S14.** SDS-PAGE analysis of PPK2s and PPK2 variants.
### SI Tables

**Table S1. X-ray data collection and refinement statistics.** For clarity, the table is presented in two parts.

<table>
<thead>
<tr>
<th>Protein*</th>
<th>Ligands</th>
<th>Mr</th>
<th>AMP</th>
<th>Mr</th>
<th>ADP</th>
<th>Mr</th>
<th>ATP</th>
<th>Mr</th>
<th>ADP</th>
<th>PP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PDB ID</strong></td>
<td>5LC9</td>
<td>5LCD</td>
<td>5LDB</td>
<td>5LD1</td>
<td>5MAQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Space group</strong></td>
<td>P4-2_{2}</td>
<td>P4-2_{2}</td>
<td>P4-2_{2}</td>
<td>P4-2_{2}</td>
<td>P4-2_{2}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cell dimensions</strong></td>
<td>164.78, 164.78, 95.0</td>
<td>164.4, 164.4, 95.0</td>
<td>165.6, 165.6, 95.2</td>
<td>168.0, 168.0, 95.0</td>
<td>166.2, 166.2, 95.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \alpha, \beta, \gamma ) (°)</td>
<td>90, 90</td>
<td>90, 90</td>
<td>90, 90</td>
<td>90, 90</td>
<td>90, 90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>45.7 - 1.90</td>
<td>58.14 - 2.66</td>
<td>45.93 - 2.30</td>
<td>48.25 - 2.09</td>
<td>117.5 - 2.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( R_{	ext{sym}} ) or ( R_{	ext{merge}} )</td>
<td>0.244 (2.809)</td>
<td>0.322 (2.045)</td>
<td>0.475 (4.413)</td>
<td>0.260 (1.586)</td>
<td>0.367 (1.509)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( R_{	ext{free}} )</td>
<td>0.252 (2.997)</td>
<td>0.329 (2.160)</td>
<td>0.487 (4.975)</td>
<td>0.274 (1.675)</td>
<td>0.396 (1.623)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( I / \sigma I )</td>
<td>19.7 (2.1)</td>
<td>17.3 (2.2)</td>
<td>10.3 (1.1)</td>
<td>11.7 (2.4)</td>
<td>8.6 (2.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC1/2</td>
<td>0.999 (0.893)</td>
<td>0.998 (0.819)</td>
<td>0.997 (0.439)</td>
<td>0.998 (0.746)</td>
<td>0.992 (0.332)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC*</td>
<td>1.000 (0.942)</td>
<td>0.999 (0.914)</td>
<td>0.999 (0.777)</td>
<td>1.000 (0.917)</td>
<td>0.998 (0.917)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>100 (100)</td>
<td>99.0 (99.0)</td>
<td>100 (100)</td>
<td>100 (100)</td>
<td>100 (100)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Redundancy</td>
<td>52.5 (42.4)</td>
<td>44.0 (29.2)</td>
<td>27.0 (28.2)</td>
<td>18.4 (19.2)</td>
<td>13.4 (14.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Refinement**

<p>| No. reflections | 102181(6697) | 37701(2875) | 59125(4537) | 80333(11585) | 49000 (7027) |
| ( R_{	ext{work}} ) / ( R_{	ext{free}} ) | 0.172/0.193 | 0.199/0.247 | 0.177/0.210 | 0.163/0.202 | 0.191/0.233 |
| No. atoms | Protein | 8687 | 8712 | 8707 | 8705 | 8641 |
| | Ligand/ion | 40 | 65 | 139 | 287 | 148 |
| | Water | 762 | 100 | 486 | 874 | 134 |
| B-factors | Protein | 36.8 | 48.5 | 48.8 | 32.3 | 42.1 |
| | Ligand/ion | 56.1 | 79.5 | 50.4 | 27.5 | 55.8 |
| | Water | 40.0 | 32.4 | 40.2 | 35.5 | 22.8 |
| R.m.s. deviations | Bond lengths (Å) | 0.010 | 0.008 | 0.010 | 0.010 | 0.01 |
| | Bond angles (°) | 0.91 | 0.95 | 1.00 | 0.97 | 1.04 |</p>
<table>
<thead>
<tr>
<th>Protein&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Ligands</th>
<th>Ft</th>
<th>Ft</th>
<th>Ft, D117N</th>
<th>Mr, N121D</th>
<th>Mr, N121D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PolyPs</td>
<td>PolyPs</td>
<td>PolyP&lt;sub&gt;23,11,6&lt;/sub&gt;</td>
<td>ATP; P&lt;sub&gt;i&lt;/sub&gt;</td>
<td>PolyP&lt;sub&gt;9&lt;/sub&gt;, D117N</td>
<td>PolyP&lt;sub&gt;23,11,6&lt;/sub&gt;, D117N</td>
</tr>
<tr>
<td>PDB ID</td>
<td>5LL0</td>
<td>5LLB</td>
<td>5LLF</td>
<td>5O6K</td>
<td>5O6M</td>
<td>5O6M</td>
</tr>
<tr>
<td>Space group</td>
<td>P 1 2 1</td>
<td>P 1 2 1</td>
<td>C 2 2 2 1</td>
<td>P4&lt;sub&gt;1&lt;/sub&gt;2&lt;sub&gt;1&lt;/sub&gt;2&lt;sub&gt;1&lt;/sub&gt;</td>
<td>P4&lt;sub&gt;1&lt;/sub&gt;2&lt;sub&gt;1&lt;/sub&gt;2&lt;sub&gt;1&lt;/sub&gt;</td>
<td></td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>59.22, 145.62, 70.43</td>
<td>58.7, 144.91, 70.59</td>
<td>73.43, 165.80, 255.12</td>
<td>166.66, 166.66, 94.90</td>
<td>164.62, 164.62, 95.12</td>
<td></td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90, 113, 90</td>
<td>90, 113, 90</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
<td>90, 90, 90</td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>72.81-1.96</td>
<td>72.46-1.92</td>
<td>44.16-2.20</td>
<td>47.91-2.90</td>
<td>47.53-2.30</td>
<td></td>
</tr>
<tr>
<td>R&lt;sub&gt;sym&lt;/sub&gt; or R&lt;sub&gt;merge&lt;/sub&gt;</td>
<td>(2.01-1.96)</td>
<td>(1.989 - 1.92)</td>
<td>(2.24-2.20)</td>
<td>(3.06-2.90)</td>
<td>(2.36-2.30)</td>
<td></td>
</tr>
<tr>
<td>R&lt;sub&gt;max&lt;/sub&gt;</td>
<td>0.060 (0.420)</td>
<td>0.061 (0.555)</td>
<td>0.130 (2.21)</td>
<td>0.317 (1.79)</td>
<td>0.133 (1.66)</td>
<td></td>
</tr>
<tr>
<td>I / σI</td>
<td>11.1 (2.5)</td>
<td>13.6 (2.1)</td>
<td>8.06 (0.83)</td>
<td>7.5 (1.4)</td>
<td>22.1 (2.9)</td>
<td></td>
</tr>
<tr>
<td>CC1/2</td>
<td>0.998 (0.847)</td>
<td>0.998 (0.785)</td>
<td>0.995 (0.241)</td>
<td>0.997 (0.771)</td>
<td>1.000 (0.928)</td>
<td></td>
</tr>
<tr>
<td>CC*</td>
<td>0.999 (0.958)</td>
<td>0.999 (0.938)</td>
<td>0.999 (0.624)</td>
<td>1.000 (0.921)</td>
<td>1.000 (0.977)</td>
<td></td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>97.1 (90.2)</td>
<td>99.3 (99.5)</td>
<td>99.5 (99.0)</td>
<td>99.9 (99.9)</td>
<td>100.0 (100.0)</td>
<td></td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.3 (3.4)</td>
<td>3.1 (3)</td>
<td>5.3 (5.2)</td>
<td>14.0 (14.3)</td>
<td>26.4 (27.6)</td>
<td></td>
</tr>
</tbody>
</table>

**Refinement**

<table>
<thead>
<tr>
<th>No. reflections</th>
<th>75983 (5217)</th>
<th>81970 (8194)</th>
<th>79080 (5789)</th>
<th>30049 (4310)</th>
<th>1545304 (122783)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R&lt;sub&gt;work&lt;/sub&gt; / R&lt;sub&gt;free&lt;/sub&gt;</td>
<td>0.198/0.239</td>
<td>0.179/0.231</td>
<td>0.236/0.242</td>
<td>0.259/0.297</td>
<td>0.197/0.238</td>
</tr>
<tr>
<td>No. atoms</td>
<td>Protein: 8068</td>
<td>8047</td>
<td>8882</td>
<td>8666</td>
<td>8845</td>
</tr>
<tr>
<td></td>
<td>Ligand/ion: 148</td>
<td>264</td>
<td>167</td>
<td>50</td>
<td>266</td>
</tr>
<tr>
<td></td>
<td>Water: 612</td>
<td>672</td>
<td>245</td>
<td>6</td>
<td>294</td>
</tr>
<tr>
<td>B-factors</td>
<td>Protein: 46.03</td>
<td>58.41</td>
<td>52.04</td>
<td>19.38</td>
<td>56.27</td>
</tr>
<tr>
<td></td>
<td>Ligand/ion: 60.76</td>
<td>79.22</td>
<td>63.97</td>
<td>28.22</td>
<td>67.09</td>
</tr>
<tr>
<td></td>
<td>Water: 42.49</td>
<td>55.52</td>
<td>51.06</td>
<td>6.69</td>
<td>46.00</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td>Bond lengths (Å)</td>
<td>0.003</td>
<td>0.012</td>
<td>0.008</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Bond angles (°)</td>
<td>0.79</td>
<td>1.03</td>
<td>1.07</td>
<td>0.569</td>
</tr>
</tbody>
</table>

*Values in parentheses are for highest-resolution shell. Each data set was collected from a single crystal.

**Footnotes**

<sup>a</sup>Protein sources: Ft, Francisella tularensis; Mr, Meiothermus ruber; proteins have wild type sequences except where noted. Ligands: PolyP, polyphosphate of 25 residues average length was used for crystallization. The subscripts (P<sub>n</sub>) indicate the lengths of polyphosphate chains modeled and refined in each structure.
Table S2. Structural similarity of PPK2 structures calculated using the DALI server (21). Values are RMSD (Å) and the number of residues aligned in brackets.

| Protein (PDB ID, class)
<table>
<thead>
<tr>
<th>Protein (PDB ID, class)</th>
<th>Ft + AMPPCP + polyP (5LLB, I)</th>
<th>Mr + ADP + PPi (5MDQ, III)</th>
<th>P. aeruginosa (3CZP, II)</th>
<th>S. meliloti (3CZQ, I)</th>
<th>A. aurescens (3RHF, III)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ft + AMPPCP + polyP (5LLB, I)</td>
<td>-</td>
<td>1.5 (239)</td>
<td>1.7 (241)</td>
<td>0.9 (249)</td>
<td>1.9 (268)</td>
</tr>
<tr>
<td>Mr + ADP + PPi (5MDQ, III)</td>
<td>1.5 (239)</td>
<td>-</td>
<td>1.7 (241)</td>
<td>1.7 (249)</td>
<td>1.3 (268)</td>
</tr>
<tr>
<td>P. aeruginosa (3CZP, II)</td>
<td>1.7 (241)</td>
<td>1.7 (241)</td>
<td>-</td>
<td>1.7 (249)</td>
<td>1.7 (268)</td>
</tr>
<tr>
<td>S. meliloti (3CZQ, I)</td>
<td>0.9 (249)</td>
<td>1.7 (249)</td>
<td>1.7 (249)</td>
<td>-</td>
<td>1.9 (268)</td>
</tr>
<tr>
<td>A. aurescens (3RHF, III)</td>
<td>1.9 (268)</td>
<td>1.3 (268)</td>
<td>1.7 (268)</td>
<td>1.9 (268)</td>
<td>-</td>
</tr>
</tbody>
</table>

Notes.

a Classes refer to the substrate specificity of PPK2 classes shown in Fig. 1A.
b Residues were trimmed to the core PPK2 catalytic domains as follows: P. aeruginosa (3CZP) residues 255-496; S. meliloti (3CZQ), residues 50-297; A. aurescens (3RHF), residues 4-272.
c Abbreviations: P. aeruginosa, Pseudomonas aeruginosa; S. meliloti, Sinorhizobium meliloti; A. aurescens, Arthrobacter aurescens.
**Table S3.** Oligonucleotides used in this study. Restriction sites or mutagenesis sites are in bold and underlined.

<table>
<thead>
<tr>
<th>Name</th>
<th>Oligo sequence*</th>
</tr>
</thead>
<tbody>
<tr>
<td>MrPPK2-NdeI</td>
<td>5’- TATATATATACATATGAATAATACCGCG -3’</td>
</tr>
<tr>
<td>MrPPK2-HindIII</td>
<td>5’- TATATATAAGCTTTATTCTCAATAACAATTTTTTCGC -3’</td>
</tr>
<tr>
<td>FtPPK2-D117N</td>
<td>5’- AATCGTACTTTTTAATAGGCTTGTTGA -3’</td>
</tr>
<tr>
<td>FtPPK2-D62A</td>
<td>5’- GAAGGTAAGAGGCGAGCTGGTA -3’</td>
</tr>
<tr>
<td>FtPPK2-K66A</td>
<td>5’- TGCAGCTGGTGCGGGTGAACX-3’</td>
</tr>
<tr>
<td>FtPPK2-D192A</td>
<td>5’- AGCCCTATAGCCTAAAGCATCCT -3’</td>
</tr>
<tr>
<td>FtPPK2-R118A</td>
<td>5’- CGTACTTTTTTGATCGTCTTGTTAAT -3’</td>
</tr>
<tr>
<td>FtPPK2-R178A</td>
<td>5’- GATTTGCTGCTGGCGGAAAGTCATCC -3’</td>
</tr>
</tbody>
</table>

*Oligonucleotides were sourced from commercial suppliers.
**Table S4.** Summary of RALS analysis of PPK2 oligomerization.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Oligomer</th>
<th>Calculated MW (kDa)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Without ligands</th>
<th>With 10 mM PolyP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Observed MW (kDa)</td>
<td>Amount of Protein&lt;sup&gt;b&lt;/sup&gt; (%)</td>
</tr>
<tr>
<td><em>FtPPK2</em></td>
<td>Monomer</td>
<td>32.9</td>
<td>39.7</td>
<td>100</td>
</tr>
<tr>
<td>WT</td>
<td>Dimer</td>
<td>65.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>MrPPK2</em></td>
<td>Monomer</td>
<td>33.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT</td>
<td>Dimer</td>
<td>67.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trimer</td>
<td>101.1</td>
<td>99.2</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Tetramer</td>
<td>134.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Octamer</td>
<td>269.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dodecamer</td>
<td>404.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hexadecamer</td>
<td>539.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Includes His-tag (calculated with Expasy protparam tool [https://web.expasy.org/protparam/]).

<sup>b</sup>Amount of protein (as a percentage of sample) represent the percentage of total area fitted in Fig. S9.
**Table S5.** Conditions for HPLC Analysis for PPK2 Kinetic Assays.

<table>
<thead>
<tr>
<th></th>
<th>F\textit{PPK2}</th>
<th>M\textit{PPK2}</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Solid phase</strong></td>
<td>Gemini C18</td>
<td>MultoHigh 100 RP-18</td>
</tr>
<tr>
<td></td>
<td>(150 x 4.6 mm, 5 μ)</td>
<td>(250 x 4 mm, 5 μ)</td>
</tr>
<tr>
<td><strong>Mobile phase</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A (aqueous)</td>
<td>0.1 M triethylamine, pH 7*</td>
<td>40 mM sodium acetate, pH 4.2*</td>
</tr>
<tr>
<td>B (organic)</td>
<td>80% MeOH, 0.1 M triethylamine, pH 7*</td>
<td>acetonitrile</td>
</tr>
<tr>
<td><strong>Detection wavelength</strong></td>
<td>260 nm</td>
<td>254 nm</td>
</tr>
<tr>
<td><strong>Flow rate</strong></td>
<td>0.8 mL/min</td>
<td>0.5 mL/min</td>
</tr>
<tr>
<td><strong>Equilibration</strong></td>
<td>95% A</td>
<td>98% A</td>
</tr>
<tr>
<td><strong>Sample injected</strong></td>
<td>100 μL</td>
<td>10 μL</td>
</tr>
<tr>
<td><strong>Gradient</strong></td>
<td>24 min 95% A</td>
<td>10 min 98% A</td>
</tr>
<tr>
<td></td>
<td>1 min 95% → 90% A</td>
<td>6 min 98% → 70% A</td>
</tr>
<tr>
<td></td>
<td>10 min 90% A</td>
<td>2 min 70% A</td>
</tr>
<tr>
<td></td>
<td>1 min 90% → 0% A</td>
<td>2 min 70% → 98% A</td>
</tr>
<tr>
<td></td>
<td>5 min 100% B</td>
<td>10 min 98% A</td>
</tr>
<tr>
<td></td>
<td>(re-equilibration)</td>
<td>(re-equilibration)</td>
</tr>
<tr>
<td></td>
<td>1 min 0% → 95% A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 min 95% A</td>
<td>(re-equilibration)</td>
</tr>
</tbody>
</table>

*adjusted with acetic acid.
References