The 2.4 Å structure of *Zymomonas mobilis* pyruvate kinase: implications for stability and regulation

Kathleen M. Meneely[1], Jeffrey S. McFarlane[2,ǂ], Collette L. Wright[2], Kathryn Vela[1], Liskin Swint-Kruse[3], Aron W. Fenton[3], Audrey L. Lamb[1]*

[1]Department of Chemistry, University of Texas at San Antonio, San Antonio, TX 78249

[2]Department of Molecular Biosciences, University of Kansas, Lawrence, KS 66045

[3]Department of Biochemistry and Molecular Biology, University of Kansas Medical Center, Kansas City, KS 66160

[ǂ]Current address: Department of Chemistry and Biochemistry, Fort Lewis College, Durango, CO 81301

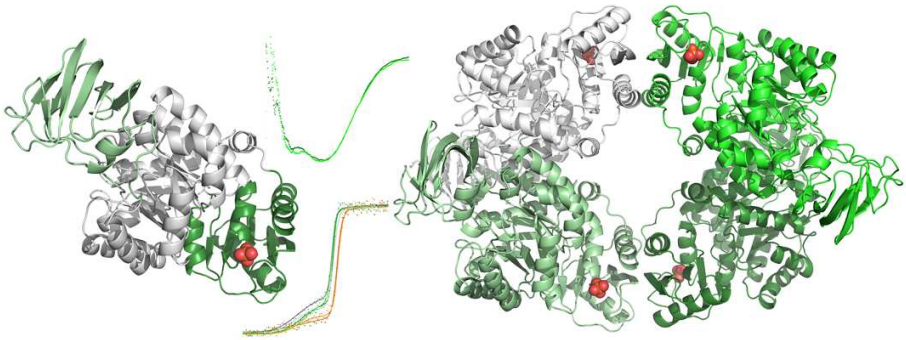
* author to whom correspondence should be addressed, audrey.lamb@utsa.edu

Keywords: pyruvate kinase, allosterism,

HIGHLIGHTS

- The non-allosteric pyruvate kinase from *Zymomonas mobilis* is dimeric in solution
- Dimeric ZmPYK crystallizes as a tetramer using the standard interfaces from higher organisms
- ZmPYK may not possess sufficient protein stability for activity to be tuned by allosteric effectors

GRAPHICAL ABSTRACT

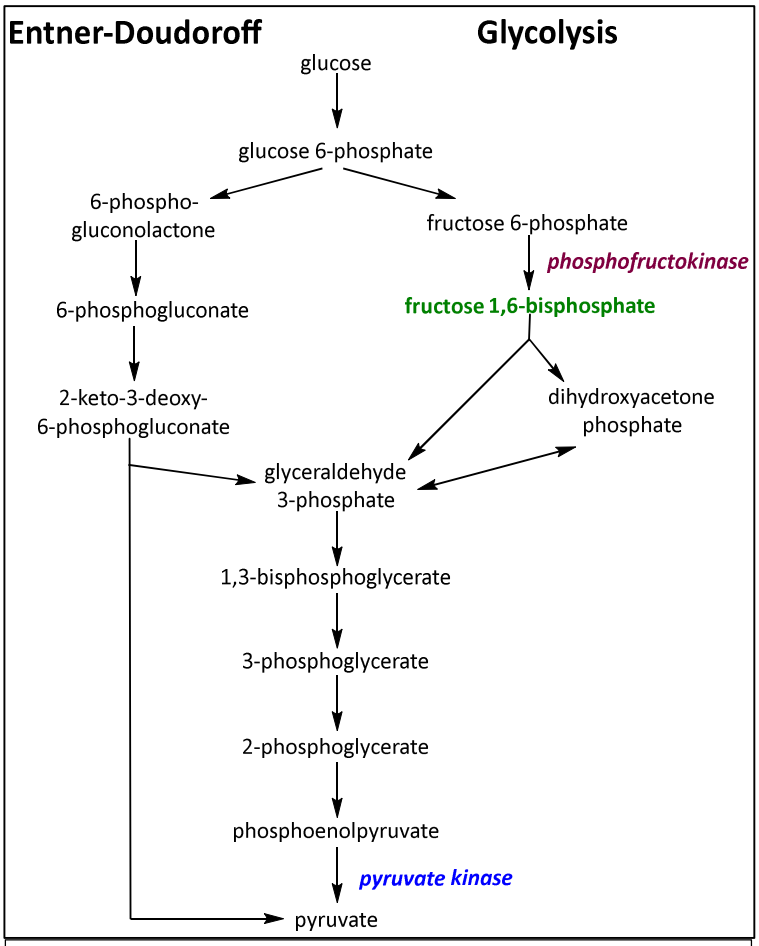


ABSTRACT

Human liver pyruvate kinase (hlPYK) catalyzes the final step in glycolysis, the formation of pyruvate (PYR) and ATP from phosphoenolpyruvate (PEP) and ADP. Fructose 1,6-bisphosphate (FBP), a pathway intermediate of glycolysis, serves as an allosteric activator of hlPYK. *Zymomonas mobilis* pyruvate kinase (ZmPYK) performs the final step of the Entner-Doudoroff pathway, which is similar to glycolysis in that energy is harvested from glucose and pyruvate is generated. The Entner-Doudoroff pathway does not have FBP as a pathway intermediate, and ZmPYK is not allosterically activated. In this work, we solved the 2.4 Å X-ray crystallographic structure of ZmPYK. The protein is dimeric in solution as determined by gel filtration chromatography, but crystalizes as a tetramer. The buried surface area of the ZmPYK tetramerization interface is significantly smaller than that of hlPYK, and yet tetramerization using the standard interfaces from higher organisms provides an accessible low energy crystallization pathway. Interestingly, the ZmPYK structure showed a phosphate ion in the analogous location to the 6-phosphate binding site of FBP in hlPYK. Circular Dichroism (CD) was used to measure melting temperatures of hlPYK and ZmPYK in the absence and presence of substrates and effectors. The only significant difference was an additional phase of small amplitude for the ZmPYK melting curves. We conclude that the phosphate ion plays neither a structural or allosteric role in ZmPYK under the conditions tested. We hypothesize that ZmPYK does not have sufficient protein stability for activity to be tuned by allosteric effectors as described for rheostat positions in the allosteric homologues.

INTRODUCTION

Pyruvate kinase (PYK) is the final enzyme of glycolysis, converting phosphoenolpyruvate and ADP to pyruvate and ATP. The human liver isoform (hlPYK) is tetrameric, and allosterically regulated by fructose-1,6-bisphosphate (FBP), an earlier pathway intermediate of glycolysis (**Scheme 1**). Binding of FBP (yellow in **Figure 1**) to a site distant from the active site alters the catalytic rate (citrate is bound in the active site in this structure and shown in brown) [1]. hlPYK has long been used as a model system for the study of allosterism [1-22]; however, a predominance of the structures determined for hlPYK were determined at low pH values, which render the enzyme unresponsive to allosteric controls[15]. Furthermore, structures have not been determined at all points of the thermodynamic cycle (apo, substrate-bound, modulator-bound, modulator and substrate-bound), making it difficult to ascertain the network of residue interactions that activate catalysis upon binding of FBP.



Scheme 1: Enter-Doudoroff Pathway and Glycolysis

Zymomonas mobilis is a Gram negative facultative anaerobe that was originally

isolated as a contaminant from beer and cider [23]. *Z. mobilis* uses the Entner-Doudoroff pathway, in place of glycolysis, to convert glucose to glucose 6-phosphate (hexokinase) which is oxidized to 6-phosphogluconolactone by glucose 6-phosphate dehydrogenase (**Scheme 1**) [23]. The lactone is ring opened to make 6-phosphogluconate (by 6-phosphogluconlactonase) and dehydrated to make 2-keto-3-deoxy-6-phosphogluconate (by 6-phosphogluconate dehydratase), which is cleaved to generate glyceraldehyde 3-phosphate and pyruvate (by 2-keto-3-deoxy-6-phosphogluconate aldolase). The energetic result of the Enter-Doudoroff pathway is 1 net ATP produced per glucose molecule metabolized, as opposed to 2 ATP per glucose for glycolysis. Since glyceraldehyde 3-phosphate is metabolized using enzymes homologous to those of glycolysis, *Z. mobilis* possesses a pyruvate kinase. Importantly, the Enter-Doudoroff pathway does not contain a phosphofructokinase homologue, and fructose 1,6-bisphosphate is not a pathway intermediate. Genome analysis has confirmed that *Z. mobilis* lacks the phosphofructokinase gene [24], thus it is not surprising that *Z. mobilis* pyruvate kinase (ZmPYK) is not allosterically regulated by fructose-1,6-bisphosphate as is the case for hlPYK [25, 26].

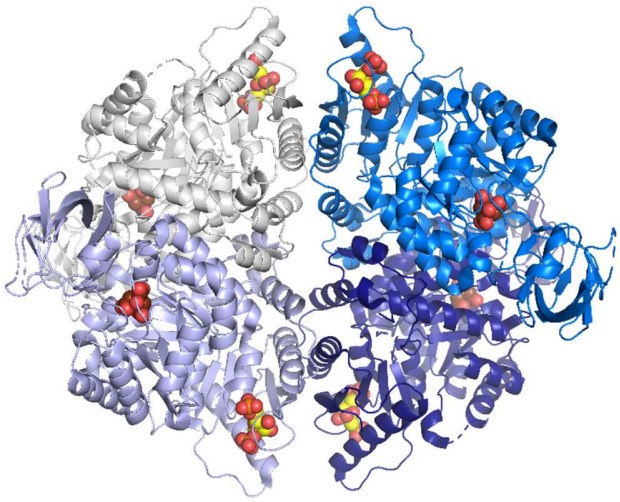


Figure 1. Tetrameric structure of hlPYK (PDB: 4IP7; monomers shown in different colors) with citrate (brown) bound in the active site and fructose 1,6-bisphosphate (yellow) bound in the allosteric site. In this image, the dimers are made up two monomers of a lighter shade (gray and lavender) or 2 monomers of a vibrant shade (azure and navy blue). These two sets of dimers form the tetramers, with tetramerization interfaces between the gray and azure monomers and the lavender and navy monomers.

Here, we determined the structure of the non-allosteric ZmPYK to 2.4 Å resolution. To our surprise, the protein crystallized as a tetramer, despite eluting from a gel filtration column at a molecular weight consistent with a dimer. The structure contained a phosphate ion in the site comparable to where the 6-phosphate of FBP binds in hlPYK. We used circular dichroism (CD) spectroscopy of phosphate-bound and phosphate-free samples to determine that the phosphate in the crystal structure is neither structurally essential nor an allosteric regulator under these conditions.

MATERIALS AND METHODS

Preparation of overexpression plasmids. The gene sequence for the *Zymomonas mobilis* pyruvate kinase was acquired from NCBI using accession number WP_011240112.1. The sequence was codon optimized for expression in *E. coli* by GenScript and ligated into pET-15TV such that the produced protein contains a N-terminal hexahistidine affinity tag that can be cleaved with TEV protease. The plasmid was transformed into *E. coli* BL21 (DE3) cells (Invitrogen) for protein overexpression. The expression plasmid for tobacco etch virus (TEV) protease was a kind gift from Roberto DeGuzman (University of Kansas). This construct produces TEV protease containing an N-terminal hexahistidine tag.

Overexpression and purification of ZmPYK with phosphate. *E. coli* BL21 (DE3) transformed with the ZmPYK overexpression vector were grown overnight in 50-100 mL LB Miller broth containing 200 µg/ml ampicillin in a shaking incubator (185 rpm) at 37 °C. Baffled flasks containing 1 L of LB Miller broth and 200 µg/ml ampicillin were inoculated with 10-35 ml of

overnight culture and grown to an OD_{600} of 0.6-0.87 at 37 °C. The cultures were induced with 200 μM of isopropyl β-D-1-thiogalactopyranoside (IPTG; final concentration) and grown 3 hours at 37 °C. The cells were harvested by centrifugation at 6,000 x g for 10 minutes at 4 °C. The pellet was resuspended in 25 mM Tris pH 8, 500 mM sodium chloride, 50 mM imidazole when the protein was used for circular dichroism or biochemical assays. For protein intended for crystallization, the purification buffer used was 25 mM potassium phosphate pH 8. The cells were lysed using a French Press apparatus, with four passages at 15,000 psi. The lysate was centrifuged for 30 minutes at 11,000 x g at 4 °C, and loaded onto a nickel-chelating sepharose fast flow column (Cytiva) equilibrated with the same buffer. ZmPYK was eluted using a linear gradient of the purification buffer with 500 mM imidazole. ZmPYK elutes at 100 mM imidazole. Pooled fractions containing ZmPYK were dialyzed into 50 mM Tris pH 8, 150 mM sodium chloride, 0.2 mM Tris(2-carboxyethyl)phosphine (TCEP) for CD studies or 50 mM potassium phosphate pH 8, 50 mM sodium citrate, 50 mM sodium chloride, 10% (v/v) glycerol and 0.2 mM TCEP for crystallization studies. The hexahistidine tag was removed by TEV (tobacco etch virus) protease cleavage. Briefly, nickel fractions were dialyzed into 50 mM potassium phosphate pH 8, 50 mM sodium citrate, 50 mM sodium chloride, 10% glycerol and 0.2 mM TEV protease was added at a molar ratio of 1:50 (TEV protease:protein) and incubated at 4 °C overnight. TCEP was then dialyzed out of the buffer before returning the protein over the nickel-chelating fast flow sepharose column. Unbound ZmPYK was collected in the wash, concentrated and loaded onto a GE Superdex 200 size exclusion chromatography column equilibrated with 50 mM MES (2-(N-morpholino)ethanesulfonic acid) pH 6.8, 100 mM potassium chloride, 1 mM TCEP and 10% (v/v) glycerol at 4 °C. ZmPYK was concentrated using an Amicon stirred cell to

a final concentration of 14-27 mg/ml as measured by Bradford assay prior to freezing aliquots at -80 °C. This preparation yields 28 mg per liter of culture.

Overexpression and purification of ZmPYK without phosphate. Phosphate free ZmPYK was expressed as described above. The purification was similar to that described about, except a 25 mM Tris 8 buffer system was used in place of the phosphate buffer system, and an additional phosphate removal step was added after TEV protease cleavage. The TEV protease was transferred to a phosphate-free buffer (25 mM Tris 8.0, 150 mM NaCl) using a PD-10 column (Cytiva) prior to use. ZmPYK and TEV protease were dialyzed together overnight in 25 mM Tris 8.0, 150 mM NaCl, 0.2 mM TCEP. The TCEP was removed from the ZmPYK sample by dialysis after TEV cleavage (1 hr against 25 mM Tris 8.0, 150 mM NaCl). To remove phosphate from ZmPYK, human purine nucleoside phosphorylase (PNP, 10 mg), 5 mM adenosine, and 1 mM magnesium chloride were incubated with ZmPYK overnight at 4 °C before returning the protein to the nickel-chelating sepharose fast flow column and completing the purification using a Tris buffer system. ZmPYK scrubbed of phosphate was concentrated using an Amicon stirred cell to a final concentration of 19 mg/mL as measured by Bradford assay prior to freezing aliquots at -80 °C. This preparation yields 5.2 mg per liter of culture.

Overexpression and purification of hlPYK. The overexpression and purification of hlPYK was completed as previously described [15].

Overexpression and purification of Tobacco Etch Virus (TEV) protease. The plasmid producing TEV protease was transformed into *E. coli* BL21 (DE3) cells. Baffled flasks containing 1 L of

LB Miller broth containing 200 µg/ml ampicillin was inoculated with 25 ml of overnight culture and grown to an OD_{600} of 1.1 at 37 °C. The culture was induced with 1 mM IPTG and the culture was grown at 37 °C for 3 hours. The cells were harvested at 6,000 x g for 10 min at 4 °C. The pellet was resuspended in 50 mM sodium phosphate pH 8, 100 mM sodium chloride, 50 mM imidazole, 10% (v/v) glycerol. The cells were lysed by three passages through a French Press apparatus. The lysate was centrifuged for 30 minutes at 11,500 x g at 4 °C, and loaded onto a nickel-chelating sepharose fast flow column equilibrated with the same buffer. Protein was eluted using a step gradient containing 300 mM imidazole. Fractions containing TEV protease were pooled and dialyzed against 50 mM sodium phosphate pH 8, 200 mM sodium chloride, 10 mM DTT, 2 mM ethylenediaminetetraacetic acid (EDTA), 10% (v/v) glycerol. This preparation yields 50 mg per liter of culture. Aliquots of 4.4 mg/ml as determined by Bradford assay were stored at -80 °C.

Overexpression and purification of purine nucleoside phosphorylase (PNP). The human purine nucleoside phosphorylase enzyme (PNP) construct was purchased from Addgene (catalog number 64076) containing an N-terminal hexahistadine tag and was transformed into *E. coli* BL21 (DE3) cells and expressed in LB Miller broth overnight at 37 °C along with 200 µg/ml ampicillin in a shaking incubator. Baffled flasks containing 1 L of LB Miller media containing 200 µg/ml ampicillin was inoculated with 25 ml of overnight culture and grown to an OD_{600} of 0.6 at 37 °C. The culture was induced with 100 µM IPTG and the culture was grown at 37 °C for 16 hours. The cells were centrifuged at 6,000 x g for 10 min at 4 °C. The pellet was resuspended in 25 mM Tris pH 8, 500 mM sodium chloride, 50 mM imidazole. The cells were lysed by three passages through a French Press apparatus and centrifuged for 30 minutes at 11,500 x g at 4 °C.

The lysate was loaded onto a nickel-chelating sepharose fast flow column equilibrated with the same buffer. Protein was eluted using a step gradient containing 300 mM imidazole. Fractions containing PNP were pooled and dialyzed against 50 mM HEPES pH 8, 100 mM sodium chloride. The final concentration was determined by Bradford assay as 2 mg/ml and stored at -80 °C.

Determination of phosphate content. The removal of phosphate from ZmPYK samples was verified using a colorimetric assay in which inorganic phosphate and 2-amino-6-mercapto-7-methylpurine riboside (MesGR) are substrates for PNP, producing 2-amino-6-mercapto-7-methylpurine (MesG) monitored at A_{360} on a Shimadzu UV-2600 spectrophotomer. 100 µM PNP, 533 µM ZmPYK, 200 µM MesGR in 50 mM Tris-HCl, 1 mM magnesium chloride, pH 7.5 were incubated for 1 hr at room temperature. The absorbance was converted to concentration using a phosphate standard curve at concentrations of 0 to 20 µM.

Crystallization of ZmPYK. Crystals were grown in hanging drops composed of 1.5 µL of 29 mg/mL ZmPYK (purified with phosphate) in 50 mM MES pH 6.8, 100 mM potassium chloride, 1 mM TCEP and 10% (v/v) glycerol and 1.5 µL of well solution composed of 65 mM potassium thiocyanate and 29.5% (w/v) PEG 3350. Crystals grew as rods and cubes within 5 days at 24 °C. Crystals were transferred into well solution supplemented with 15% (v/v) glycerol as a cryoprotectant using a cryoloop prior to flash cooling in liquid nitrogen.

X-ray structure determination of ZmPYK. Diffraction data were collected remotely using Blu-Ice [27] on beamline 12-2 at the Stanford Synchrotron Radiation Lightsource (SSRL, Menlo Park,

CA). A wavelength of 0.97946 Å and detector distance of 250 mm were used to collect 240° of data with 0.15° oscillation with 0.2 s exposures at a temperature of 100 K. Statistics for data collection and refinement are listed in **Table 1**. Data were processed to 2.40 Å using AutoPROC [28]. A phasing solution was determined by molecular replacement in Phaser [29] using pyruvate kinase from *Geobacillus stearothermophilus* (PDB: 2E28, monomer A) as a model with a resulting LLG of 1,491 and TFZ of 22.5. Following the initial build performed in Phenix Autobuild [30] the R_{work} = 27.2 and the R_{free} = 31.8 with 3225 out of 3800 residues placed. Rounds of model building and refinement were completed in Coot [31] and Phenix Refine [30]. Density for the phosphate was evident following molecular replacement, but the phosphate ligand was modeled only after refinement was complete for the ZmPYK protein chains. Waters were placed by Phenix Refine, corrected manually and verified following a round of refinement.

Table 1. Data collection and refinement statistics
Values in parentheses are for the highest resolution shell.

	ZmPYK
Data collection	
Space group	$P2_12_12$
Unit cell (Å, °)	a=207.9, b=256.1, c=82.7, α, β, γ=90°
Resolution range (Å)	161.4 – 2.40 (2.442 – 2.40)
Completeness (%)	98.7 (99.0)
Total reflections	2,319,686 (106,871)
Unique reflections	170,665 (8,504)
I / σ	10.8 (2.1)
R_{merge}[a]	0.234 (1.312)
R_{pim}[b]	0.065 (0.382)
Multiplicity	13.6 (12.6)
Refinement	
Resolution range (Å)	66.9 – 2.40 (2.43 – 2.40)
No. of reflections	159,016 (4,602)
R_{work}	18.5 (24.2)
R_{free}[c]	24.6 (34.5)
No. non-hydrogen atoms	28,119
Protein atoms	27,334
Ligand/ions	70

Waters	715
Ramachandran favored (%)	97.05
Ramachandran allowed (%)	2.58
Ramachandran outliers (%)	0.37
Wilson *B*	28.7
Average *B* (Å^2)	36.2
Protein	36.3
Ligand/ion	36.6
R.m.s. deviations	
Bond lengths (Å)	0.013
Bond angles (°)	1.27

[a] $R_{\text{merge}} = \sum_{hkl}|I_{hkl} - \langle I\rangle_{hkl}|/\sum_{hkl} I_{hkl}$ where $I_{hkl}$is the intensity of reflection *hkl* and $\langle I\rangle$ is the mean intensity of related reflections.

[b] $R_{pim} = \sum_{hkl}\sqrt{1/n-1}\,|I_{hkl} - \langle I\rangle_{hkl}|/\sum_{hkl} I_{hkl}$ where n is the multiplicity of related reflections.

[c] $R = \sum\left|F_o - |F_c|\right| / \sum|F_o|$ where F_o = to the observed structure factors and F_c = structure factors calculated from the model. 5% of the reflections were initially reserved to create an R_{free} test set used during each subsequent round of refinement.

ZmPYK crystallographic model analysis. The asymmetric unit of the ZmPYK crystal contained 8 monomers. A model for the B domain (residues 77-175) that caps the active site could be constructed for 6 of 8 monomeric chains. A summary of the model components can be found in **Table 2**. The model contains 715 water molecules. Ramachandran analysis was performed by MolProbity [32] showing good geometry overall with 0.37% outliers. Threonine 280, an outlier in all 8 chains, is associated with the active site and is a known outlier in pyruvate kinases [15]. Structural alignments were performed using PDBeFold [33]. Structural figures were generated in PyMOLv.3.7 (Schrodinger). The coordinates and structure factors have been archived in the Protein DataBank with accession code 7UEH.

Table 2 Model Components

Chain	Ordered Residues
A	3-475
B	4-475

C	4-170, 173-461, 465-475
D	4-82, 89-154, 157-475
E	4-81, 172-475
F	4-79, 164-463, 466-475
G	3-475
H	3-475

Circular dichroism spectroscopy. Circular dichroism (CD) spectra were collected on a Jasco J-1100 CD Spectropolarimeter using 250 µL solutions of protein in a 1 mm cell. Proteins were diluted to 1 mg/ml in 50 mM MES pH 6.8, 100 mM KCl, 10% glycerol, 1 mM TCEP with ligand concentrations at 2 mM. Each trace represents the average of three scans at 50 nm/min with a 1.00 nm bandwidth and a digital integration time of 4 seconds. Spectral scan data were collected from 185 nm to 260 nm at 0.1 nm intervals. The melting temperature of each protein was measured with a bandwidth of 1.00 nm, monitoring at 222 nm. Data were collected from 4 °C to 80 °C with equilibration time of 60 seconds between each data point. The midpoints of the unfolding transitions were determined by fitting either single or double sigmoidal curves in Kaliedagraph v5.0.3. The reported melting temperature represents the average of three melting experiments. CD spectra of 1 mg/ml protein were collected after melting, and again after cooling the melted sample. For both ZmPYK and hlPYK, the protein did not have refolding capacity (data not shown).

RESULTS

Crystal structure of ZmPYK. ZmPYK was heterologously produced in *E. coli* with an N-terminal histidine tag, and purification proceeded using nickel-affinity and gel

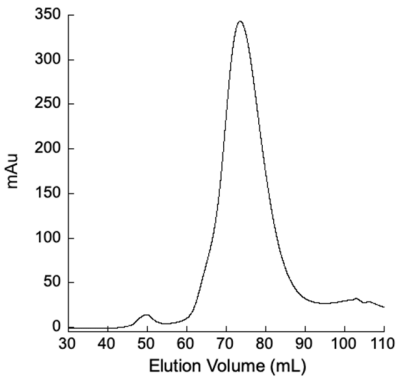


Figure 2. ZmPYK oligomer in solution. ZmPYK eluted from a Superdex 200 size exclusion column as a 101 kDa complex, equivalent to a dimer.

filtration chromatography. This process included a step for the removal of the histidine tag using TEV protease. ZmPYK eluted from the gel filtration column with an elution volume that equates to a molecular weight of 101 kDa, which is equivalent to a dimer of the 51.4 kDa protein (**Figure 2**), agreeing with previous reports that ZmPYK is dimeric and not tetrameric like hlPYK [25, 26]. The peak has a shoulder at lower elution volume (~65 mL), potentially indicating an equilibrium with a larger molecular weight oligomer. ZmPYK crystallized using PEG 3350 as a precipitant, and the structure was determined by molecular replacement using a monomer of the *G. stearothermophilus* pyruvate kinase (PDB: 2E28) as a model (**Figure 3**). The resultant structure has 8 molecules in the asymmetric unit.

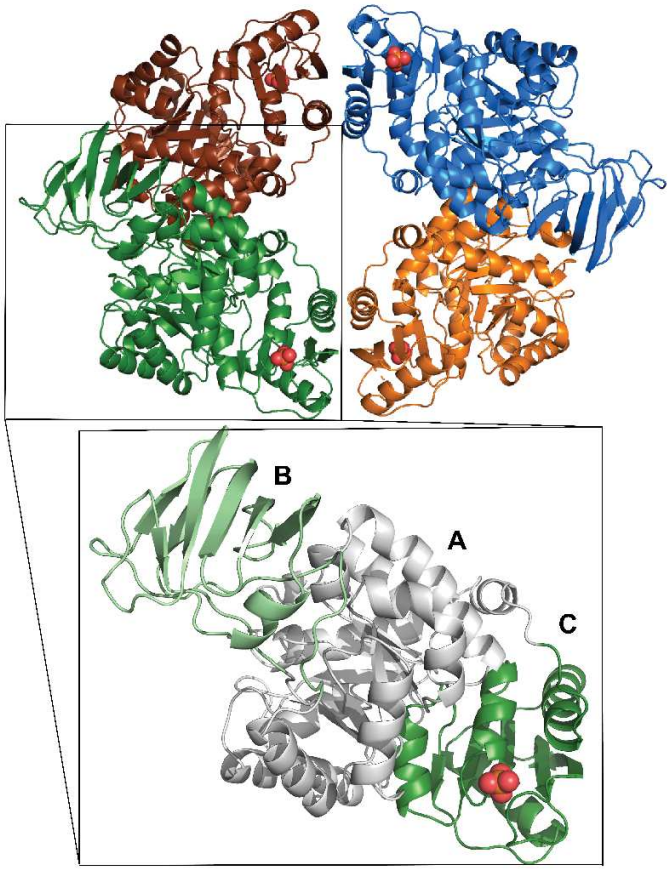


Figure 3. The tetramer and monomer structure of ZmPYK. The dimers are made up of two monomers colored brown and green or 2 monomers colored blue and orange. These two sets of dimers form the tetramers, with tetramerization interfaces between the brown and blue monomers and the green and orange monomers. The phosphate ions bound in the site comparable to a phosphate moiety of FBP bound to hlPYK are shown as spheres. *Inset:* A monomer of ZmPYK with the A, B and C domains highlighted.

Monomeric structure of ZmPYK. Like other pyruvate kinases, ZmPYK monomers are composed of three domains (**Figure 3**). The A domain, residues 1-77 and 176-352, contains the active site. The B domain interrupts the A domain, consisting of residues 78-175. In many structures of PYK homologs, the B domain is mobile and cannot always be modeled. Indeed, in this structure of ZmPYK, the B domain is only modeled in 6 of the 8 monomers, with insufficient density in monomers E and F. The C domain, which contains the site analogous to the FBP-binding pocket in hlPYK, is made up of residues 353-475.

Tetrameric assembly. The crystallographic software PISA[33] (proteins, interfaces, structures and assemblies) predicts two tetrameric assemblies formed by chains A - D and E – H, respectively, with one assembly shown in **Figure 3**. While this tetrameric assembly is homologous to other pyruvate kinase tetramers (**Figure 4**), it has less interface surface area. The ZmPYK monomer is 68 amino acids smaller than hlPYK; a majority of the difference is due to the tetramerization interface. The first 25 residues of hlPYK are disordered in the majority of structures determined, and these have no corresponding residues in ZmPYK. At residue 10, the ZmPYK A domain begins to align well with hlPYK's A domain at that homologue's residue 55. These residues in hlPYK are a significant part of the tetramerization interface. In addition, there is a helix in hlPYK (residues 402 to 415 of the C domain) that is in proximity to residues 26-55, which are also part of the tetramerization interface. By comparison, these residues in ZmPYK (343-354) form a short helix followed by a loop, reducing the size of the tetramerization interface. PISA calculates the dimeric interfaces of hlPYK to be ~2750 $Å^2$. In ZmPYK, the dimerization interfaces are significantly smaller: ~1660 $Å^2$ (see **Figure 3** caption). The tetramer interface in ZmPYK comprises ~800 $Å^2$, whereas this same interface in hlPYK is ~1450 $Å^2$. Although

ZmPYK crystallized by packing in the same tetrameric form as hlPYK, the smaller interacting surface area suggests the reason that the dimeric state is predominant in solution. Typically, a single interface with less than 1000 Å^2 surface area would be considered a crystal packing interface and not a clear biological assembly.

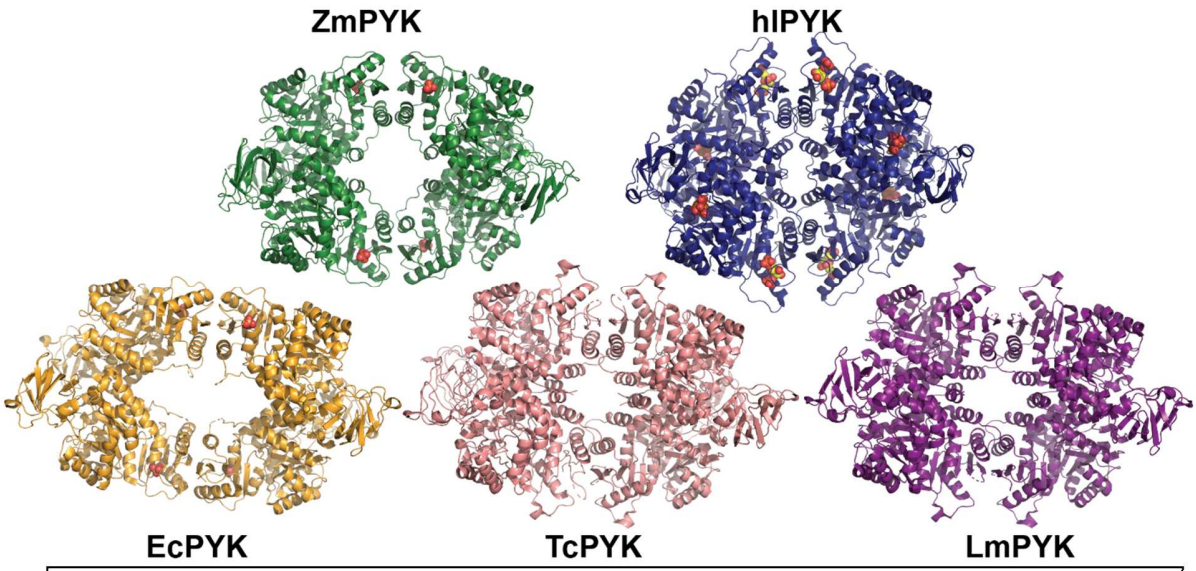


Figure 4. Comparison of structural homologues with ZmPYK.

Comparison to structural homologues. The nearest structural homolog is *E. coli* type I pyruvate kinase (PDB:4YNG), with root mean squared deviation (r.m.s.d.) of 1.42-1.58 Å over 444-451 Cα, depending on which monomers are being compared (**Figure 4**). Two other close structural homologues are from *Trypanosoma cruzi* (PDB:4KRZ), with an r.m.s.d. of 1.70 Å over 452 Cα and *Leishmania mexicana* (PDB:3QV8), with an r.m.s.d. of 1.65 Å over 444 Cα. The r.m.s.d. comparison for ZmPYK and hlPYK gave a value of 2.6 Å over 416 Cα. With the exception of ZmPYK, all of these homologues are allosterically regulated and tetrameric. Interestingly, the two prokaryotic proteins (ZmPYK and EcPYK) do not have the N-terminal residues that assist with tetramerization. Nevertheless, EcPYK does bury surface area suitable for tetramerization (each interface is ~2000 Å^2) and is tetrameric in solution [34]. All of the homologues depicted here

encode a complete glycolytic pathway, except ZmPYK, which uses the Entner-Doudoroff pathway [35, 36] (**Figure 1**).

Phosphate ion in "allosteric" site. The allosteric binding site is largely conserved in ZmPYK, with a phosphate ion bound in the location occupied by the 6-phosphate of FBP in the homologue structures (**Figure 5**). There are two differences at the 1-phosphate position that may account for FBP not serving as an allosteric modulator for ZmPYK. Arginine 501 of hlPYK makes an ionic interaction with the 1-phosphate. This residue in ZmPYK is leucine 436, and cannot perform the same role. Indeed, the helix in which arginine 501 resides and the loop that precedes it (residues 423-443 of ZmPYK) have shifted significantly making the allosteric binding pocket smaller. In addition, ZmPYK glutamate 383 has replaced hlPYK threonine 445. In hlPYK Thr-445 faces away from the binding cavity. The longer Glu-383 of ZmPYK could occlude the binding site of the 1-phosphate of FBP, or cause charge repulsion preventing binding.

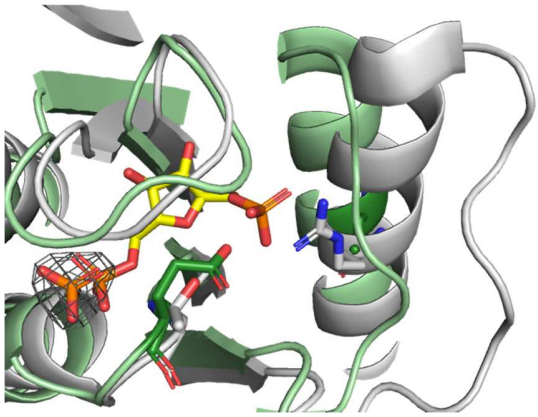


Figure 5. The allosteric site. The allosteric site of hlPYK (gray) with FBP bound (yellow) is overlaid with the same site for ZmPYK (green) and the bound phosphate ion. The phosphate ion of the ZmPYK structure is surrounded by the $2F_o$-F_c map contoured at 1.5σ. Note that Arg501 that chelates the 1-phosphate in hlPYK is L346 in ZmPYK reducing binding energy for the sugar phosphate. Also of note, Thr445 of hlPYK is E383 in ZmPYK, which may cause charge repulsion with the 1-phosphate and/or occlude the 1-phosphate binding site.

Removal of phosphate from ZmPYK. Purine nucleoside phosphorylase (PNP) was used to scrub phosphate from ZmPYK samples. PNP catalyzes the conversion of a purine nucleoside and phosphate into adenine and ribose 1-phosphate. PNP and excess adenosine were incubated with

the ZmPYK sample, and the ZmPYK was subsequently further purified by nickel affinity. To verify the removal of phosphate from the sample, ZmPYK was denatured and tested for phosphate using the MesG assay. None was detected (**Supplemental Figure 1**). This indicates that our scrubbing method was successful in removing phosphate (the phosphate was bound sufficiently tightly such that it could only be removed with the use of PNP or protein unfolding).

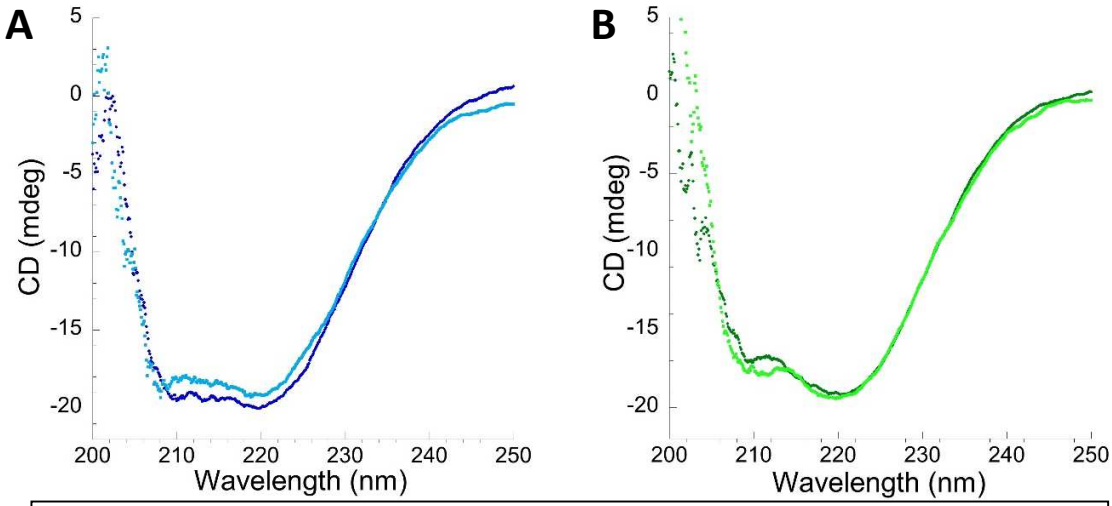


Figure 6. Circular dichroism spectra. A. hlPYK in the absence (dark blue) and presence (light blue) of FBP. **B.** ZmPYK in the absence (dark green) and presence (light green) of phosphate.

Circular Dichroism spectra. To examine the effect of FBP or phosphate on the overall fold of hlPYK and ZmPYK, CD spectroscopy was performed. When the CD spectra of hlPYK were recorded in the absence and presence of FBP, the traces were comparable and indicated a mix of α-helices and β-sheets (**Figure 6A**). When the analogous experiment was performed with ZmPYK, in the absence and presence of phosphate, the shape of the curve was also comparable (**Figure 6B**). The presence of substrates or products (phosphoenolpyruvate and pyruvate) did not appreciably change the spectra; however, ATP and ADP altered the traces at wavelengths below 215 nm, most likely due to signal interference by these molecules (**Supplemental Figure 2**). For this reason, 222 nm was the wavelength chosen for melting curves.

Circular Dichroism melting curves. To examine enzyme stability in the presence and absence of FBP or phosphate, the melting temperatures were measured by CD spectroscopy. The melting curves for hlPYK were determined in the absence and presence of FBP. These traces show a single melting phase and were fit to a sigmoidal equation with a single exponential term (**Figure 7**). The melting temperatures range from 61.4 to 66.9 °C, with the effector FBP not significantly altering the melting temperatures (**Table 3**). However, the binding of substrates and products changed the thermal stability of hlPYK, as noted by the change in slope of the curves (**Supplemental Figure 3**). For example, pyruvate promotes thermal stability when compared to the apo protein or other bound ligands (light and dark purple traces in **Figure 7**).

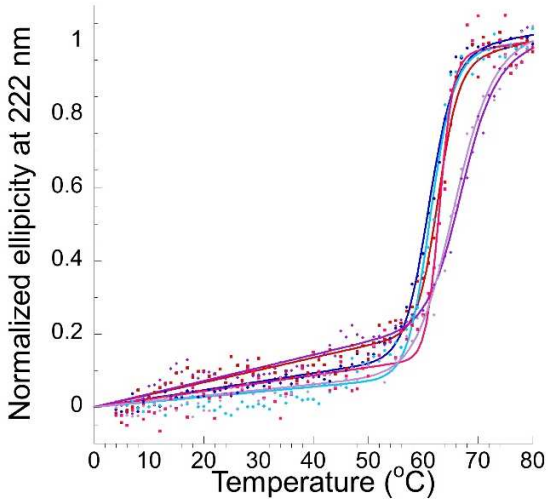


Figure 7. hlPYK melting curves. Melting curves are shown for apo-hlPYK (red), hlPYK-FBP (pink), hlPYK-PEP (blue), hlPYK-PEP-FBP (cyan), hlPYK-PYR (purple), hlPYK-PYR-FBP (lavender).

The melting curves for ZmPYK were determined in the absence and presence of phosphate (**Figure 8**). Interestingly, the curves were sigmoidal, but were fit using two exponential terms to account for the two distinct melting temperatures: 38 ± 4 °C (without phosphate) and 36 ± 2 °C (with phosphate) for the early transition, and 49 ± 1 °C (without phosphate) and 50.1 ± 0.4 °C (with phosphate) for the late transition (**Table 4**). The early transition appears to represent ~20% of the normalized signal, with the

Table 3: Melting Curve data for hlPYK

		Tm1 (°C)	Slope Tm1
hlPYK w/o FBP			
	apo	61.4 ± 0.1	28 ± 1
	ATP	63.4 ± 0.3	37 ± 4
	ADP	63.2 ± 0.2	33 ± 6
	PEP	65.21 ± 0.04	51 ± 4
	pyruvate	66.9 ± 0.3	20 ± 4
hlPYK w/ FBP			
	apo	61.6 ± 0.2	34 ± 3
	ATP	63.9 ± 0.4	34 ± 5
	ADP	63.5 ± 0.6	52 ± 4
	PEP	66.7 ± 0.4	60 ± 11
	pyruvate	65.3 ± 0.4	24 ± 4

remaining 80% for the second transition. The traces overlay well (dark and light green in **Figure 8**), with the slopes also being comparable. Due to the observed oligomerization differences in the size exclusion chromatography and the crystal structure, one possibility is that the first transition arises from changes in secondary structure that are coincident with the conversion of the protein to a dimeric state, a feature that would be found for ZmPYK but not hlPYK. However, this is one possible interpretation among many, including the impact of the mobile B domains or general differences in overall stability. When substrates (phosphoenolpyruvate or ADP) or products (pyruvate or ATP) are included, the primary difference is in the fraction of change in ellipticity accounted for by the initial phase. The melting temperature for the first phase ranges from 33 to 40 °C and for the second phase from 49 to 50.1 °C. There is one exception: ADP causes an increase in melting temperature of the second, significant unfolding phase (52 – 52.4 °C) (orange and yellow traces

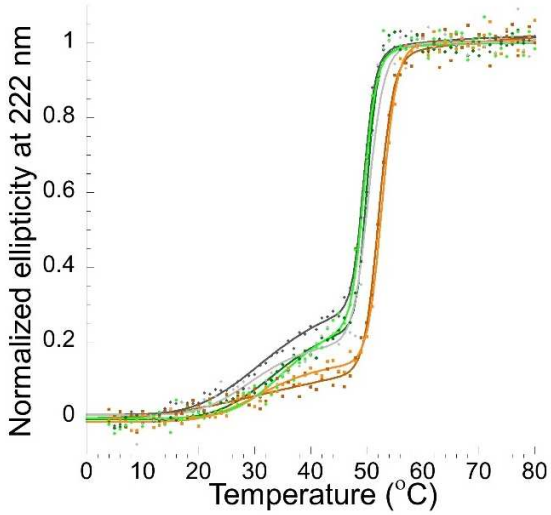


Figure 8. ZmPYK melting curves. Melting curves are shown for apo-ZmPYK (dark green), ZmPYK-PO_4 (light gren), ZmPYK-ADP (orange), ZmPYK-ADP-PO_4 (yellow), ZmPYK-PYR (dark gray), ZmPYK-PYR-PO_4 (light gray).

in **Figure 8**), and eliminates the initial unfolding phase in the sample without phosphate. The black and grey traces in **Figure 8** represent the traces with pyruvate, which show the most predominant initial unfolding phase. All traces can be found in **Supplemental Figure 4**.

DISCUSSION

In this work, we report the X-ray crystallographic structure of ZmPYK to 2.4 Å. While gel filtration chromatography confirms the dimeric state to be the most prevalent in solution, the protein crystallized as a tetramer. ZmPYK crystallizes using the standard tetramerization interfaces observed for PYK homologues; however, the amount of buried surface area is considerably smaller than that documented in other species, consistent with solution dimerization. In other words, the protein is dimeric in solution, but crystallized as a tetramer because this provides an easily accessible low energy path to crystal formation. Circular dichroism measurements suggest that there is an initial melting transition, that could be interpreted as being derived from the dimer-tetramer oligomerization equilibrium, from B domain mobility or other general differences in protein stability between ZmPYK and hlPYK.

Table 4: Melting Curve data for ZmPYK

		Tm1 (°C)	Tm2 (°C)	Slope Tm1	Slope Tm2
ZmPYK w/o PO4					
	apo	38 ± 4	49 ± 1	10 ± 10	48 ± 8
	ATP	37 ± 1	49.2 ± 0.2	8 ± 3	52 ± 7
	ADP		52.0 ± 0.5		38 ± 4
	PEP	40 ± 2	49.6 ± 0.1	9 ± 1	64 ± 16
	pyruvate	33 ± 3	49.5 ± 0.2	6 ± 1	46 ± 13
ZmPYK w/ PO4					
	apo	36 ± 2	50.1 ± 0.4	12 ± 2	51 ± 9
	ATP	40 ± 1	49.9 ± 0.1	7 ± 1	54 ± 9
	ADP	30 ± 3	52.4 ± 0.5	5 ± 1	39 ± 2
	PEP	38 ± 2	49.8 ± 0.1	9 ± 1	53 ± 8
	pyruvate	36 ± 1	50.1 ± 0.2	6 ± 1	75 ± 7

Neither the final ZmPYK gel filtration purification step for crystallization nor the crystallization conditions contain phosphate ions, so we were surprised to observe a phosphate bound in the site analogous to the 6-phosphate site of FBP in hlPYK. This led us to speculate that the phosphate ion could be serving as an allosteric regulator in the place of FBP or could be serving a structural role. Circular dichroism spectra and melting curves show no significant changes upon binding of the effector FPB to hlPYK nor for binding of phosphate to ZmPYK. Our data are therefore inconclusive as to the role of phosphate in ZmPYK, but the simplest explanation is that the phosphate is derived from the cellular cytoplasm of the *E. coli* during heterologous expression, and remains bound throughout purification. A current hypothesis is that the Entner-Doudoroff pathway evolutionarily pre-dates glycolysis [37]. This phosphate binding site may represent the initial phases for developing an allosteric site for regulation of pyruvate kinase in higher order species. Alternatively, this site could play a functional under some as yet unknown condition.

Many studies of enzyme function rely on mutational analyses in which catalytic residues are substituted for amino acids of different biochemical properties and the kinetic parameters are compared between wildtype and variant forms. For example, converting the catalytic nucleophile, say the serine of a serine protease, to a residue that is incapable of serving as a nucleophile, such as an alanine, results in the abolishment of catalytic activity. These residues serve structurally or functionally important roles, are very highly conserved, and are considered "toggle" positions (activity is on or off). Some non-conserved positions in allosterically regulated proteins have been shown to have "rheostat" activity: substituting these positions with all possible variations shows that activity can be dialed up or down like a dimmer switch, sometimes to values higher than wildtype [11, 18, 21, 38-40]. The model systems for these studies have

been hlPYK, aldolase A, and the LacI/GalR protein family. Interestingly, changes in sequence at rheostat residues have been shown to coincide with phylogenetic branch points (non-conserved positions that correlate with speciation events within a family phylogenetic tree) [41]. We recently reported that using the same methodology for identifying rheostat positions in allosteric enzymes (phylogenetic branch points) was not successful in ZmPYK [42].

When this work began, we had expected to map rheostat residues onto the solved structure of ZmPYK, but in other work we have recently shown that of all of the predicted rheostat residues (mutations associated with phylogenetic branch points), only one turned out to show moderate rheostat behavior: Proline 226, which is found in Domain A at the interface with Domain B [42]. All other predicted sites had toggle or neutral (no change in kinetic parameters) behavior. This may be because of the low melting temperature of ZmPYK relative to hlPYK. The CD data indicate that ZmPYK is less thermally stable than hlPYK, promoting a theory that rheostat residues are not observed because the protein does not have a similar dynamic range for calibrating activity before it becomes unstable. We propose that tetramerization of PYK affords sufficient structural stability to accommodate rheostat tuning of allosteric behavior. An alternative hypothesis is that only enzymes evolutionarily adapted to tune their activity using allosteric modulators will possess rheostat positions. We hypothesize that ZmPYK will make an excellent structural model to serve as a negative control for allosteric structural changes observed in hlPYK as this research moves forward.

ACKNOWLEDGEMENTS

We thank Pierce T. O'Neil (KUMC) for comments on the manuscript. This work was funded by grants from the W. M. Keck Foundation (to LSK, AWF, and ALL), from the National Institutes of General Medical Sciences (GM11589 to LSK and AWF, GM147635 to LSK, GM127665 to ALL). JSM was supported by a National Institutes of Health Graduate Training Program in the Dynamic Aspects of Chemical Biology Grant T32 GM008545 and by an American Heart Association Predoctoral Fellowship PRE33960374. CLW was supported as a K-INBRE Scholar by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under grant number P20 GM103418.

Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Contract No. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research, and by the NIH and NIGMS (including P41 GM103393). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIH or NIGMS. Thank you to the staff at the SSRL for their generous assistance.

REFERENCES

[1] Holyoak, T., Zhang, B., Deng, J., Tang, Q., Prasannan, C. B., and Fenton, A. W. (2013) Energetic coupling between an oxidizable cysteine and the phosphorylatable N-terminus of human liver pyruvate kinase, *Biochemistry 52*, 466-476.

[2] Alontaga, A. Y., and Fenton, A. W. (2011) Effector analogues detect varied allosteric roles for conserved protein-effector interactions in pyruvate kinase isozymes, *Biochemistry 50*, 1934-1939.

[3] Carlson, G. M., and Fenton, A. W. (2016) What Mutagenesis Can and Cannot Reveal About Allostery, *Biophys J 110*, 2809.

[4] Chappell, B. M., and Fenton, A. W. (2020) The phosphate moiety of phosphoenolpyruvate does NOT contribute to allosteric regulation of liver pyruvate kinase by fructose-1,6-bisphosphate(), *Arch Biochem Biophys 695*, 108633.
[5] Fenton, A. W. (2012) Identification of allosteric-activating drug leads for human liver pyruvate kinase, *Methods Mol Biol 796*, 369-382.
[6] Fenton, A. W. (2013) Are all regions of folded proteins that undergo ligand-dependent order-disorder transitions targets for allosteric peptide mimetics?, *Biopolymers 100*, 553-557.
[7] Fenton, A. W., and Alontaga, A. Y. (2009) The impact of ions on allosteric functions in human liver pyruvate kinase, *Methods Enzymol 466*, 83-107.
[8] Fenton, A. W., and Hutchinson, M. (2009) The pH dependence of the allosteric response of human liver pyruvate kinase to fructose-1,6-bisphosphate, ATP, and alanine, *Arch Biochem Biophys 484*, 16-23.
[9] Fenton, A. W., and Tang, Q. (2009) An activating interaction between the unphosphorylated n-terminus of human liver pyruvate kinase and the main body of the protein is interrupted by phosphorylation, *Biochemistry 48*, 3816-3818.
[10] Harris, R. A., and Fenton, A. W. (2019) A critical review of the role of M2PYK in the Warburg effect, *Biochim Biophys Acta Rev Cancer 1871*, 225-239.
[11] Hodges, A. M., Fenton, A. W., Dougherty, L. L., Overholt, A. C., and Swint-Kruse, L. (2018) RheoScale: A tool to aggregate and quantify experimentally determined substitution outcomes for multiple variants at individual protein positions, *Hum Mutat 39*, 1814-1826.
[12] Ishwar, A., Tang, Q., and Fenton, A. W. (2015) Distinguishing the interactions in the fructose 1,6-bisphosphate binding site of human liver pyruvate kinase that contribute to allostery, *Biochemistry 54*, 1516-1524.
[13] Johnson, L. E., Ginovska, B., Fenton, A. W., and Raugei, S. (2019) Chokepoints in Mechanical Coupling Associated with Allosteric Proteins: The Pyruvate Kinase Example, *Biophys J 116*, 1598-1608.
[14] Martin, T. A., Wu, T., Tang, Q., Dougherty, L. L., Parente, D. J., Swint-Kruse, L., and Fenton, A. W. (2020) Identification of biochemically neutral positions in liver pyruvate kinase, *Proteins 88*, 1340-1350.
[15] McFarlane, J. S., Ronnebaum, T. A., Meneely, K. M., Chilton, A., Fenton, A. W., and Lamb, A. L. (2019) Changes in the allosteric site of human liver pyruvate kinase upon activator binding include the breakage of an intersubunit cation-pi bond, *Acta Crystallogr F Struct Biol Commun 75*, 461-469.
[16] Pendergrass, D. C., Williams, R., Blair, J. B., and Fenton, A. W. (2006) Mining for allosteric information: natural mutations and positional sequence conservation in pyruvate kinase, *IUBMB Life 58*, 31-38.
[17] Prasannan, C. B., Tang, Q., and Fenton, A. W. (2012) Allosteric regulation of human liver pyruvate kinase by peptides that mimic the phosphorylated/dephosphorylated N-terminus, *Methods Mol Biol 796*, 335-349.
[18] Swint-Kruse, L., Martin, T. A., Page, B. M., Wu, T., Gerhart, P. M., Dougherty, L. L., Tang, Q., Parente, D. J., Mosier, B. R., Bantis, L. E., and Fenton, A. W. (2021) Rheostat functional outcomes occur when substitutions are introduced at nonconserved positions that diverge with speciation, *Protein Sci 30*, 1833-1853.

[19] Tang, Q., and Fenton, A. W. (2017) Whole-protein alanine-scanning mutagenesis of allostery: A large percentage of a protein can contribute to mechanism, *Hum Mutat 38*, 1132-1143.
[20] Tang, Q., Villar, M. T., Artigues, A., Thyfault, J. P., Apte, U., Zhu, H., Peterson, K. R., and Fenton, A. W. (2019) Mutational mimics of allosteric effectors: a genome editing design to validate allosteric drug targets, *Sci Rep 9*, 9031.
[21] Wu, T., Swint-Kruse, L., and Fenton, A. W. (2019) Functional tunability from a distance: Rheostat positions influence allosteric coupling between two distant binding sites, *Sci Rep 9*, 16957.
[22] Xu, Q., Tang, Q., Katsonis, P., Lichtarge, O., Jones, D., Bovo, S., Babbi, G., Martelli, P. L., Casadio, R., Lee, G. R., Seok, C., Fenton, A. W., and Dunbrack, R. L., Jr. (2017) Benchmarking predictions of allostery in liver pyruvate kinase in CAGI4, *Hum Mutat 38*, 1123-1131.
[23] Swings, J., and De Ley, J. (1977) The biology of Zymomonas, *Bacteriol Rev 41*, 1-46.
[24] Seo, J. S., Chong, H., Park, H. S., Yoon, K. O., Jung, C., Kim, J. J., Hong, J. H., Kim, H., Kim, J. H., Kil, J. I., Park, C. J., Oh, H. M., Lee, J. S., Jin, S. J., Um, H. W., Lee, H. J., Oh, S. J., Kim, J. Y., Kang, H. L., Lee, S. Y., Lee, K. J., and Kang, H. S. (2005) The genome sequence of the ethanologenic bacterium Zymomonas mobilis ZM4, *Nat Biotechnol 23*, 63-68.
[25] Pawluk, A., Scopes, R. K., and Griffiths-Smith, K. (1986) Isolation and properties of the glycolytic enzymes from Zymomonas mobilis. The five enzymes from glyceraldehyde-3-phosphate dehydrogenase through to pyruvate kinase, *Biochem J 238*, 275-281.
[26] Steiner, P., Fussenegger, M., Bailey, J. E., and Sauer, U. (1998) Cloning and expression of the Zymomonas mobilis pyruvate kinase gene in Escherichia coli, *Gene 220*, 31-38.
[27] McPhillips, T. M., McPhillips, S. E., Chiu, H. J., Cohen, A. E., Deacon, A. M., Ellis, P. J., Garman, E., Gonzalez, A., Sauter, N. K., Phizackerley, R. P., Soltis, S. M., and Kuhn, P. (2002) Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines, *J Synchrotron Radiat 9*, 401-406.
[28] Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., and Bricogne, G. (2011) Data processing and analysis with the autoPROC toolbox, *Acta Crystallogr D Biol Crystallogr 67*, 293-302.
[29] McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software, *J Appl Crystallogr 40*, 658-674.
[30] Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr D Biol Crystallogr 66*, 213-221.
[31] Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallogr D Biol Crystallogr 66*, 486-501.
[32] Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography, *Acta Crystallogr D Biol Crystallogr 66*, 12-21.

[33] Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions, *Acta Crystallogr D Biol Crystallogr 60*, 2256-2268.
[34] Waygood, E. B., and Sanwal, B. D. (1974) The control of pyruvate kinases of Escherichia coli. I. Physicochemical and regulatory properties of the enzyme activated by fructose 1,6-diphosphate, *J Biol Chem 249*, 265-274.
[35] Maugeri, D. A., Cannata, J. J., and Cazzulo, J. J. (2011) Glucose metabolism in Trypanosoma cruzi, *Essays Biochem 51*, 15-30.
[36] Saunders, E. C., DP, D. E. S., Naderer, T., Sernee, M. F., Ralton, J. E., Doyle, M. A., Macrae, J. I., Chambers, J. L., Heng, J., Nahid, A., Likic, V. A., and McConville, M. J. (2010) Central carbon metabolism of Leishmania parasites, *Parasitology 137*, 1303-1313.
[37] Romano, A. H., and Conway, T. (1996) Evolution of carbohydrate metabolic pathways, *Res Microbiol 147*, 448-455.
[38] Fenton, A. W., Page, B. M., Spellman-Kruse, A., Hagenbuch, B., and Swint-Kruse, L. (2020) Rheostat positions: A new classification of protein positions relevant to pharmacogenomics, *Med Chem Res 29*, 1133-1146.
[39] Fenton, K. D., Meneely, K. M., Wu, T., Martin, T. A., Swint-Kruse, L., Fenton, A. W., and Lamb, A. L. (2022) Substitutions at a rheostat position in human aldolase A cause a shift in the conformational population, *Protein Sci 31*, 357-370.
[40] Ruggiero, M. J., Malhotra, S., Fenton, A. W., Swint-Kruse, L., Karanicolas, J., and Hagenbuch, B. (2021) A clinically relevant polymorphism in the Na(+)/taurocholate cotransporting polypeptide (NTCP) occurs at a rheostat position, *J Biol Chem 296*, 100047.
[41] Miller, M., Vitale, D., Kahn, P. C., Rost, B., and Bromberg, Y. (2019) funtrp: identifying protein positions for variation driven functional tuning, *Nucleic Acids Res 47*, e142.
[42] Page, B. M., Martin, T. A., Wright, C. L., Fenton, L. A., Villar, M. T., Artigues, A., Tang, Q., Lamb, A. L., Fenton, A. W., and Swint-Kruse, L. (in revision) Going rogue: predictions from multiple sequence aligments do not extrapolate to a homolog with an isolated evoluationary trajectory, *Protein Sci*.