

# **An Overview of Structure, Function and Regulation of Pyruvate Kinases**

Norbert Schormann<sup>1</sup>†, Katherine L. Hayden<sup>2</sup>†, Paul Lee<sup>3</sup>, Surajit Banerjee<sup>4</sup> & Debasish Chattopadhyay<sup>3\*</sup>

<sup>1</sup> Department of Biochemistry, University of Alabama at Birmingham, Birmingham, AL 35294

<sup>2</sup> Department of Chemistry and Physics, Birmingham-Southern College, Birmingham, AL 35254

<sup>3</sup> Department of Medicine, University of Alabama at Birmingham, Birmingham, AL 35294

<sup>4</sup> Northeastern Collaborative Access Team and Department of Chemistry and Chemical Biology, Cornell University, Argonne, IL 60439

\*Corresponding author: [dchatop@uabmc.edu](mailto:dchatop@uabmc.edu)

†Contributed equally

## Supplementary Material: Materials & Methods, Tables, Figures and Movies

### Materials & Methods

1. Crystallization and Data Collection
2. Crystal Structure Determination and Refinement

### Supplementary Tables

Table S1: Crystal structures of PyKs (with references).

Table S2: Pairwise protein sequence comparisons of PyKs.

Table S3: Hinge region sequences.

Table S4: Site interactions in representative PyKs.

Table S5: Data collection and refinement statistics for holo-*CpPyK* (6P0Y).

Table S6: Canonical dimers in *L. mexicana*, *T. cruzi*, *M. tuberculosis*, *S. cerevisiae*, *E. coli* PyKs versus covalent dimer in *C. parvum* PyK.

### Supplementary Figures

Figure S1. Close-up view of the A-domain.

Figure S2. Close-up view of the C-domain.

Figure S3. Inter-subunit interactions involving effector loop residues at the C-C interface.

Figure S4. Interactions between monomers at the A-A interface.

Figure S5. Comparison of ADP and metal ion binding sites in holo-*CpPyK* and *LmPyK* substrate activator complex.

Figure S6. Effector loop conformation in *TcPyK* in the effector-free and effector-bound states.

Figure S7. Rock and lock model for allosteric transition in Trypanosomatid PyKs.

- A. Dimers across the C-C interface in *TcPyK* FDP complex (4KS0) subunits.
- B. Assembly of *LmPyK* (substrate plus activator complex) tetramer R-state (3HQP).
- C. Close-up view of the interactions between subunits across the C-C interface.

Figure S8. Allosteric transitions in Human PyK-M2.

- A. C-C interface in human PyK-M2 in the absence of the activator (3SRH) in the T-state.
- B. Cartoon diagram showing the C-C interface in human PyK-M2 OXL, ATP and FBP bound state (4FXF).
- C. Priming the neighboring active site in the R-state.

Figure S9. Effector loop in holo-CpPyK (*6P0Y*).

Figure S10. Ordering of the effector loop in holo-CpPyK structure stabilizes the C-C dimer interface.

- A. Apo-CpPyK tetramer.
- B. Tetramer in holo-CpPyK structure.
- C. Close up view of the interactions across the C-C interface.

#### Supplementary Material

M1: Quicktime movie showing the morph between apo- and holo-complex of *LmPyK* (*3HQN* and *3HQO*, respectively)

M2: Quicktime movie showing the morph between apo- and holo-complex of *CpPyK* (*4DRS* and *6P0Y*, respectively)

## Materials & Methods

### 1. Crystallization and Data Collection

*Cryptosporidium parvum* pyruvate kinase (*CpPyK*) was subjected to crystallization trials using Art Robbins Gryphon robot and commercial screens. The concentrated protein (at 10mg/ml) was incubated for at least an hour with 1mM ADP, 5mM MgCl<sub>2</sub> and 5mM KCl. Initial hits (from 20% of PEG 3350, PEG 4000, PEG 6000, PEG 8000 and PEG 10000 in acetate, Mes and Hepes buffer at pH 4.6, 6.5 and 7.5, respectively) were obtained from the Qiagen PEGSuite crystallization screen, with the best hit conditions being further refined to routinely obtain crystals. Briefly, 1 µl of reservoir solution was mixed with 1 µl of incubated protein solution in a hanging drop vapor diffusion setup using Linbro boxes. *CpPyK* was flash frozen in buffer of the crystallization condition (100mM Hepes, pH 7.5, 10-15% PEG 8000) supplemented by 25% glycerol. Diffraction data were collected using an ADSC Quantum 315 CCD detector at the NE-CAT 24-ID-E beamline of the Advanced Photon Source (APS) in Chicago. Data for *CpPyK* were integrated, merged and scaled with XDS (Kabsch, 2010a; Kabsch, 2010b) followed by Aimless (Evans, 2006) in CCP4 (Winn *et al.*, 2011). Data collection parameters are listed in Supplementary Table S5.

### 2. Crystal Structure Determination and Refinement

The structure was solved by molecular replacement using Phaser (McCoy *et al.*, 2007) with a monomer of the apo form (4DRS) as initial search model. Ligands (ADP and metal ions) and water molecules were placed into previously unmodeled blobs of mFo-DFc difference electron density after initial refinement of the two protein subunits. Refinement was carried out by a combination of Refmac5 (Murshudov *et al.*, 1997) in CCP4 and Phenix (Afonine *et al.*, 2012), using TLS (Winn *et al.*, 2001) with isotropic B-factor refinement. For ligand (ADP) refinement we used Mogul (Mogul; Cambridge Structural Database) restraints obtained from the Grade web server (<http://grade.globalphasing.org>). For model building we used Coot (Emsley & Cowtan, 2004). Figures are created with PyMol (Version 2.2.0; Schrödinger LLC). The model quality was validated using a combination of QualityCheck (<https://smb.slac.stanford.edu/jcsg/QC/>), Phenix and wwPDB ([https://validate-rcsb-2.wwpdb.org/](https://validate.rcsb-2.wwpdb.org/)). The validations consist of Molprobity (Chen *et al.*, 2010) results (geometry, clashes etc.) for the protein residues and listings about the fit to the electron density of protein residues, water molecules, metal ions (K<sup>+</sup>, Mg<sup>2+</sup>) and ligands (ADP). Identified conserved Mg<sup>2+</sup> and K<sup>+</sup> sites in *CpPyK* were also verified using the CheckMyMetal web server ([https://csgid.org/csgid/metal\\_sites/](https://csgid.org/csgid/metal_sites/)). Refinement parameters are listed in Supplementary Table S5.

### 3. References

- Kabsch, W. (2010a). *Integration, scaling, space-group assignment and post-refinement*. Acta Cryst. D66:133-144.
- Kabsch, W. (2010b). *XDS*. Acta Cryst. D66:125-132.
- Afonine, P.V. *et al.* (2012). *Towards automated crystallographic structure refinement with phenix.refine*. Acta Cryst. D68:352-367.
- Emsley, P., Cowtan, K. (2004). *Coot: model-building tools for molecular graphics*. Acta Cryst. D60:2126-2132.
- Murshudov, G.N., Vagin, A.A., Dodson, E.J. (1997). *Refinement of macromolecular structures by the maximum-likelihood method*. Acta Cryst. D53:240-255.
- Winn, M.D., Isupov, M.N., Murshudov G.N. (2001). *Use of TLS parameters to model anisotropic displacements in macromolecular refinement*. Acta Cryst. D57: 122 - 133.
- The PyMOL Molecular Graphics System, Version 2.2.0 Schrödinger, LLC.
- Mogul - A knowledge-based library of molecular geometry derived from the Cambridge Structural Database (CSD).
- Chen, V.B. *et al.* (2010). *MolProbity: all-atom structure validation for macromolecular crystallography*. Acta Cryst. D66: 12-21.
- McCoy, A.J. *et al.* (2007). *Phaser crystallographic software*. J. Appl. Cryst. 40: 658-674.
- Evans, P.R. (2006). *Scaling and assessment of data quality*. Acta Cryst. D62: 72-82.
- Winn, M.D. *et al.* (2011). *Overview of the CCP4 suite and current developments*. Acta Cryst. D67: 235-242.

**Table S1. Crystal structures of PyKs.**

<b>Species</b>	<b>PDB ID</b>	<b>Resolution [Å]</b>	<b>Active Site Ligands</b>	<b>Effector/ other Ligands</b>	<b>Reference</b>
<i>E. coli</i>	4YNG	2.28	none	SO4	1)
<i>E. coli</i> R292D	1E0T	1.80	none	SO4	2)
<i>E. coli</i> R2171L	1E0U	2.00	none	SO4	2)
<i>E. coli</i>	1PKY	2.50	none	none	3)
<i>G. stearo thermophilus</i>	2E28	2.40	none	SO4	4)
<i>S. aureus</i>	3T0T	3.10	I30	none	5)
<i>S. aureus</i>	3T05	3.05	none	none	6)
<i>S. aureus</i>	3T07	3.30	09C	none	6)
<i>M. tuberculosis</i>	5WRP	2.85	none	PO4	7)
<i>M. tuberculosis</i>	5WS8	2.62	MG,OXL	none	7)
<i>M. tuberculosis</i>	5WS9	1.90	MG,K,OXL,ATP	AMP	7)
<i>M. tuberculosis</i>	5WSA	2.85	MG,OXL	G6P	7)
<i>M. tuberculosis</i>	5WSB	2.25	MG,K,OXL	AMP,G6P	7)
<i>M. tuberculosis</i>	5WSC	2.40	MG,K,OXL	AMP,G6P	7)
<i>T. brucei</i>	4KCT	1.95	MG,K,PYR	FDP	8)
<i>T. brucei</i>	4KCU	2.35	MG,K,MLT	FDP	8)
<i>T. brucei</i>	4KCV	2.18	MG,K,AKG	FDP	8)
<i>T. brucei</i>	4KCW	2.50	MG,K,OXL	FDP	8)
<i>T. brucei</i>	4HYV	2.30	MG,K,PEP	FDP	9)
<i>T. cruzi</i>	4KRZ	2.50	K	none	10)
<i>T. cruzi</i>	4KS0	2.80	MG,K,OXL	FDP	10)
<i>T. cruzi</i>	3QV9	2.10	K,QV7,PO4	none	11)
<i>T. gondii</i>	3GG8	2.21	none	SO4	12)

<i>T. gondii</i>	3EOE	2.30	none	none	12)
<i>L. mexicana</i>	3SRK	2.65	K,PTK,LSA	none	13)
<i>L. mexicana</i>	3PP7	2.35	K,SVR,PTK	none	11)
<i>L. mexicana</i>	3QV6	2.85	K,QV6	none	11)
<i>L. mexicana</i>	3QV7	2.70	K,SO4	QV7,QV8	11)
<i>L. mexicana</i>	3QV8	2.45	S62	none	11)
<i>L. mexicana</i>	3HQN	2.00	K	SO4	14)
<i>L. mexicana</i>	3HQO	3.40	K,MG,OXL,ATP	none	14)
<i>L. mexicana</i>	3HQP	2.30	K,MG,OXL,ATP	FDP	14)
<i>L. mexicana</i>	3HQQ	5.00	none	FDP	14)
<i>L. mexicana</i>	3IS4	2.10	PTK	none	15)
<i>L. mexicana</i>	3KTX	2.10	PTK	none	15)
<i>L. mexicana</i> E451W	2PM2	3.30	SO4	SO4	16)
<i>L. mexicana</i> E451W	2PM3	3.10	none	none	16)
<i>L. mexicana</i>	1PKL	2.35	none	SO4	17)
<i>C. parvum</i>	4DRS	2.50	ACT	SO4	18)
<i>C. parvum</i>	3MA8	2.64	CIT	SO4	19)
<i>C. parvum</i>	6P0Y	2.60	K,MG,ADP	none	this review
<i>P. falciparum</i>	3KHD	2.70	none	none	19)
<i>P. areophilum</i>	3QTG	2.20	none	SO4	20)
<i>A. aegypti</i>	6DU6	3.51	none	FBP	21)
<i>S. cerevisiae</i>	1A3W	3.00	K,MN,PG	FBP	22)
<i>S. cerevisiae</i>	1A3X	3.00	K,MN,PG	none	22)
Cat (M1)	1PKM	2.60	none	none	23)
Cat (M1)	1PYK	2.60	none	none	24)
Rabbit (M1)	3N25	2.41	MN,K,PYR	none	25)
Rabbit (M1)	2G50	1.65	NA,MN,PYR	Ala	26)
Rabbit (M1)	1F3W	3.00	K,MN,PYR	none	27)

<i>Rabbit</i> (M1) S402P	<i>1F3X</i>	2.80	K,MN,PYR	none	27)
<i>Rabbit</i> (M1)	<i>1AQF</i>	2.70	K,MG,PEQ	none	28)
<i>Rabbit</i> (M1)	<i>1A5U</i>	2.35	NA,MG,OXL	none	29)
<i>Rabbit</i> (M1)	<i>1A49</i>	2.10	K,MG,OXL,ATP	none	29)
<i>Rabbit</i> (M1)	<i>1PKN</i>	2.90	K,MN,PYR	none	29)
<i>Human</i> liver	<i>4IP7</i>	1.80	MN,ADN,FLC	FBP	30)
<i>Human</i> liver	<i>4IMA</i>	1.95	MN,ADN,FLC	FBP	30)
<i>Human</i> (M2)	<i>6GG4</i>	2.46	K	Phe	31)
<i>Human</i> (M2)	<i>6GG5</i>	3.20	K	Trp	31)
<i>Human</i> (M2)	<i>6GG6</i>	2.96	K,MG	Ser	31)
<i>Human</i> (M2) R399E	<i>5X0I</i>	2.64	K,MG,PYR	FBP,Ser	32)
<i>Human</i> (M2) S437Y	<i>6B6U</i>	1.35	K,MG,OXL	B3P	33)
<i>Human</i> (M2)	<i>5X1V</i>	2.10	7XX	FBP	34)
<i>Human</i> (M2)	<i>5X1W</i>	3.00	7Y0	FBP	34)
<i>Human</i> (M2)	<i>4YJ5</i>	2.41	MG,PYR	FBP,Ser	35)
<i>Human</i> (M2) Y105E	<i>4QG6</i>	3.21	none	Pro	36)
<i>Human</i> (M2) R399E	<i>4QG9</i>	2.38	MG	ACT	36)
<i>Human</i> (M2) K305Q	<i>4QG8</i>	2.30	K,MG,MLI	none	36)
<i>Human</i> (M2) K422R	<i>4QGC</i>	2.30	K	none	36)
<i>Human</i> (M2) K422R	<i>4RPP</i>	2.59	none	FBP	36)
<i>Human</i> (M2) C424A	<i>4WJ8</i>	2.87	K,MG,OXD	FBP	37)
<i>Human</i> (M2)	<i>4JPG</i>	2.33	1OX	FBP	38)



<i>Human</i> (M2)	<i>4FXF</i>	2.55	K,MG,OXL,ATP	FBP	39)
<i>Human</i> (M2)	<i>4FXJ</i>	2.90	none	Phe	39)
<i>Human</i> (M2)	<i>4B2D</i>	2.30	MG	FBP,Ser	40)
<i>Human</i> (M2)	<i>4GIN</i>	2.30	MG,OXL	NZT	41)
<i>Human</i> (M2)	<i>3U2Z</i>	2.10	O7T	FBP	42)
<i>Human</i> (M2)	<i>3SRD</i>	2.90	K,MG,OXL	FBP	43)
<i>Human</i> (M1)	<i>3SRF</i>	2.85	K,MG,PYR	none	43)
<i>Human</i> (M2)	<i>3SRH</i>	2.60	none	none	43)
<i>Human</i> (M2)	<i>3ME3</i>	1.95	3SZ	FBP	42)
<i>Human</i> (M2)	<i>3H6O</i>	2.00	D8G	FBP	44)
<i>Human</i> (M2)	<i>3GQY</i>	1.85	DZG	TLA,FBP	44)
<i>Human</i> (M2)	<i>3GR4</i>	1.60	ADP,DYY	FBP	44)
<i>Human</i> (M2) S437Y	<i>3G2G</i>	2.00	none	SO4	44)
<i>Human</i> (M2)	<i>3BJF</i>	2.03	K,MG,OXL	FBP	45)
<i>Human</i> (M2)	<i>3BJT</i>	2.50	MG,OXL	none	45)
<i>Human</i> (M2)	<i>1T5A</i>	2.80	K,MG,OXL,PO4	FBP	46)
<i>Human</i> (M2)	<i>1ZJH</i>	2.20	none	none	47)
<i>Human</i> R486W erythrocyte	<i>2VGI</i>	2.87	K,MN,PGA	FBP	48)
<i>Human</i> T384M erythrocyte	<i>2GVF</i>	2.75	K,MN,PGA	FBP	48)
<i>Human</i> R479H erythrocyte	<i>2VGG</i>	2.74	K.MN.PGA	FBP	48)
<i>Human</i> erythrocyte	<i>2VGB</i>	2.70	K,MN,PGA	FBP	48)
<i>Human</i> D499N liver	<i>6NN4</i>	2.15	PEP	FBP	49)
<i>Human</i> W527H liver	<i>6NN5</i>	2.26			49)

<i>Human</i> GGG liver	<i>6NN7</i>	2.32	FLC	49)
<i>Human</i> S531E liver	<i>6NN8</i>	2.42		49)

97 deposited structures (PDB ID in italics); 4 unreleased structures (as of June 2019)

### Legend (Ligand IDs)

MN = Mn<sup>2+</sup>; K = K<sup>+</sup>; MG = Mg<sup>2+</sup>; PYR = Pyruvic acid; OXL = Oxalate; CIT = Citric acid; ADN = Adenosine; PGA = 2-Phosphoglycolic acid; FBP = Fructose-1,6-diphosphate; FDP = Fructose-2,6-diphosphate; PO4 = Phosphate; SO4 = Sulfate; ADP = Adenosine-5'-diphosphate; ATP = Adenosine-5'-triphosphate; OXD = Oxalic acid; MLT = Malate; PEP = Phosphoenolpyruvate; AKG = 2-Oxoglutaric acid; PTK = Pyrene-1,3,6,8-tetrasulfonic acid; AMP = Adenosine monophosphate; G6P = Glucose-6-phosphate; ACT = Acetate; NA = Na<sup>+</sup>; MLI = Malonate; PEQ = Phospholactate; FLC = Citrate; SVR = 'Inhibitor'; LSA = 1,2-Benzisothiazol-3(2H)-one 1,1-dioxide; S62 = 1,3-benzothiazole-2,5-disulfonic acid; QV6 = 'Inhibitor'; QV7 = 'Inhibitor'; QV8 = 'Inhibitor'; I30 = 'Inhibitor'; 09C = 'Inhibitor'; 7XX = 'Inhibitor'; 7Y0 = 'Inhibitor'; 07T = 'Inhibitor'; 1OX = 'Inhibitor'; 3SZ = 'Inhibitor'; TLA = Tartaric acid; NZT = 'Inhibitor'; D8G = 'Inhibitor'; DZG = 'Inhibitor'; DYY = 'Inhibitor'

## References (Table S1)

- 1) Donovan, K.A., Atkinson, S.C., Kessans, S.A., Peng, F., Cooper, T.F., Griffin, M.D., Jameson, G.B., Dobson, R.C. (2016) *Acta Crystallogr D Struct Biol* **72**: 512-519.
- 2) Valentini, G., Chiarelli, L., Fortin, R., Speranza, M. L., Galizzi, A., Mattevi, A. (2000) *J Biol Chem* **275**: 18145-18152.
- 3) Mattevi, A., Valentini, G., Rizzi, M., Speranza, M.L., Bolognesi, M., Coda, A. (1995) *Structure* **3**: 729-741.
- 4) Suzuki, K., Ito, S., Shimuzu-Ibuka, A., Sakai, H. (2008) *J Biochem* **144**: 305-312.
- 5) Axerio-Cilies, P., et al. (2012) *ACS Chem Biol* **7**: 350-359.
- 6) Zoraghi, R., et al. (2011) *J Biol Chem* **286**: 44716-44725.
- 7) Zhong, W., Cui, L., Goh, B.C., Cai, Q., Ho, P., Chionh, Y.H., Yuan, M., Sahili, A.E., Fothergill-Gilmore, L.A., Walkinshaw, M.D., Lescar, J., Dedon, P.C. (2017) *Nat Commun* **8**: 1986-1986.
- 8) Zhong, W., Morgan, H.P., Nowicki, M.W., McNae, I.W., Yuan, M., Bella, J., Michels, P.A., Fothergill-Gilmore, L.A., Walkinshaw, M.D. (2014) *Biochem J* **458**: 301-311.
- 9) Zhong, W., Morgan, H.P., McNae, I.W., Michels, P.A., Fothergill-Gilmore, L.A., Walkinshaw, M.D. (2013) *Acta Crystallogr D Struct Biol* **69**: 1768-1779.
- 10) Morgan, H.P., Zhong, W., McNae, I.W., Michels, P.A., Fothergill-Gilmore, L.A., Walkinshaw, M.D. (2014) *R Soc Open Sci* **1**: 140120-140120.
- 11) Morgan, H.P., McNae, I.W., Nowicki, M.W., Zhong, W., Michels, P.A., Auld, D.S., Fothergill-Gilmore, L.A., Walkinshaw, M.D. (2011) *J Biol Chem* **286**: 31232-31240.
- 12) Bakszt, R., Wernimont, A., Allali-Hassani, A., Mok, M.W., Hills, T., Hui, R., Pizarro, J.C. (2010) *Plos One* **5**: e12736-e12736.
- 13) Morgan, H.P., et al. (2012) unpublished (3SRK)
- 14) Morgan, H.P., McNae, I.W., Nowicki, M.W., Hannaert, V., Michels, P.A.M., Fothergill-Gilmore, L.A., Walkinshaw, M.D. (2010) *J Biol Chem* **285**: 12892-12898.
- 15) Morgan, H.P., McNae, I.W., Hsin, K.Y., Michels, P.A., Fothergill-Gilmore, L.A., Walkinshaw, M.D. (2010) *Acta Crystallogr Sect F* **66**: 215-218.

- 16) Tulloch, L.B., Morgan, H.P., Hannaert, V., Michels, P.A.M., Fothergill-Gilmore, L.A., and Walkinshaw, M.D. (2008) *J Mol Biol* **383**: 615-626.
- 17) Rigden, D.J., Phillips, S.E.V., Michels, P.A.M., Fothergill-Gilmore, L.A. (1999) *J Mol Biol* **291**: 615-635.
- 18) Cook, W.J., Senkovich, O., Aleem, K., Chattopadhyay, D. (2012) *Plos One* **7**: e46875-e46875.
- 19) Wernimont, A.K., et al.: SGC (2010) unpublished (3MA8, 3KHD)
- 20) Solomons, J.T.G., Johnson, U., Schoenheit, P., Davies, C. (2012) unpublished (3QTG)
- 21) Petchampai, N., Murillo-Solano, C., Isoe, J., Pizarro, J.C., Scaraffia, P.Y. (2018) *Insect Biochem Mol Biol* **104**: 82-90.
- 22) Jurica, M.S., Mesecar, A., Heath, P.J., Shi, W., Nowak, T., Stoddard, B.L. (1998) *Structure* **6**: 195-210.
- 23) Allen, S., Muirhead, H. (1996) *Acta Cryst D* **52**: 499-504.
- 24) Stuart, D.I., Levine, M., Muirhead, H., Stammers, D.K. (1979) *J Mol Biol* **134**: 109-142.
- 25) Fenton, A.W., Johnson, T.A., Holyoak, T. (2010) *Protein Sci* **19**: 1796-1800.
- 26) Williams, R., Holyoak, T., McDonald, G., Gui, C., Fenton, A.W. (2006) *Biochemistry* **45**: 5421-5429.
- 27) Wooll, J.O., Friesen, R.H.E., White, M.A., Watowich, S.J., Fox, R.O., Lee, J.C., Czerwinski, E.W. (2001) *J Mol Biol* **312**: 525-540.
- 28) Larsen, T.M., Benning, M.M., Wesenberg, G.E., Rayment, I., Reed, G.H. (1997) *Arch Biochem Biophys* **345**: 199-206.
- 29) Larsen, T.M., Laughlin, L.T., Holden, H.M., Rayment, I., Reed, G.H. (1994) *Biochem* **33**: 6301-6309.
- 30) Holyoak, T., Zhang, B., Deng, J., Tang, Q., Prasannan, C.B., Fenton, A.W. (2013) *Biochemistry* **52**: 466-476.
- 31) Yuan, M., McNae, I.W., Chen, Y., Blackburn, E.A., Wear, M.A., Michels, P.A.M., Fothergill-Gilmore, L.A., Hupp, T., Walkinshaw, M.D. (2018) *Biochem J* **475**: 1821-1837.
- 32) Wang, W.C., Chen, T.J. (2017) unpublished (5X0I)
- 33) Srivastava, D., Razzaghi, M., Henzl, M.T., Dey, M. (2017) *Biochemistry* **56**: 6517-6520.
- 34) Matsui, Y., Yasumatsu, I., Asahi, T., Kitamura, T., Kanai, K., Ubukata, O., Hayasaka, H., Takaishi, S., Hanzawa, H., Katakura, S. (2017) *Bioorg Med Chem* **25**: 3540-3546.

- 35) Liu, J.S., Wu, C.W., Wang, W.C. (2015) unpublished (4YJ5)
- 36) Wang, P., Sun, C., Zhu, T., Xu, Y. (2015) *Protein Cell* **6**: 275-287.
- 37) Mitchell, T., Yuan, M., McNae, I., Morgan, H., Walkinshaw, M.D. (2014) unpublished (4WJ8)
- 38) Guo, C., Linton, A., Jalaie, M., Kephart, S., Ornelas, M., Pairish, M., Greasley, S., Richardson, P., Maegley, K., Hickey, M., Li, J., Wu, X., Ji, X., Xie, Z. (2013) *Bioorg Med Chem Lett* **23**: 3358-3363.
- 39) Morgan, H.P., O'Reilly, F.J., Wear, M.A., O'Neill, J.R., Fothergill-Gilmore, L.A., Hupp, T., Walkinshaw, M.D. (2013) *Proc Natl Acad Sci USA* **110**: 5881-5886.
- 40) Chaneton, B., Hillmann, P., Zheng, L., Martin, A.C.L., Maddocks, O.D.K., Chokkathukalam, A., Coyle, J.E., Jankevics, A., Holding, F.P., Vousden, K.H., Frezza, C., O'Reilly, M., Gottlieb, E. (2012) *Nature* **491**: 458-462.
- 41) Kung, C., et al. (2012) *Chem Biol* **19**: 1187-1198.
- 42) Anastasiou, D., et al. (2012) *Nat Chem Biol* **8**: 839-847.
- 43) Morgan, H.P., et al. (2012) unpublished (3SRD, 3SRF, 3SRH)
- 44) Hong, B., et al.: SGC (2009) unpublished (3GR4, 3GQY, 3G2G, 3H6O)
- 45) Christofk, H.R., Vander Heiden, M.G., Wu, N., Asara, J.M., Cantley, L.C. (2008) *Nature* **452**: 181-186.
- 46) Dombrauckas, J.D., Santarsiero, B.D., Mesezar, A.D. (2005) *Biochem* **44**: 9417-9429.
- 47) Atanassova, A., Choe, J., Arrowsmith, C., Edwards, A., Sundstrom, M., Bochkarev, A., Park, H. (2005) unpublished (1ZJH)
- 48) Valentini, G., Chiarelli, L.R., Fortin, R., Dolzan, M., Galizzi, A., Abraham, D.J., Wang, C., Bianchi, P., Zanella, A., Mattevi, A. (2002) *J Biol Chem* **277**: 23807-23814.
- 49) McFarlane, J.S., Ronnebaum, T.A., Meneely, K.M., Chilton, A., Fenton, A.W., Lamb, A.L. (2019) *Acta Crystallogr., Sect. F* **75**: 461-469.

**Table S2. Pairwise protein sequence comparisons of PyKs.**

	<i>CpPyK</i>	<i>LmPyK</i>	<i>TcPyK</i>	<i>EcPyK</i>	<i>MtPyK</i>	<i>ScPyK</i>	<i>HsM2PyK</i>	<i>TgPyK</i>
<i>CpPyK</i>		44//52/37/26	43//52/36/28	43//48/45/24	35//41/28/22	40//45/32/32	43//50/37/25	57//59/59/46
<i>LmPyK</i>			76//86/60/66	42//54/29/28	37//46/29/25	49//60/42/34	51//64/36/30	42//55/30/26
<i>TcPyK</i>				43//54/29/27	39//46/31/21	49//60/37/35	50//63/34/28	42//52/34/25
<i>EcPyK</i>					39//47/31/28	44//54/31/29	48//58/33/35	46//57/33/30
<i>MtPyK</i>						35//47/18/22	37//46/27/25	35//44/24/26
<i>ScPyK</i>							52//58/38/42	42//52/28/34
<i>HsM2PyK</i>								43//56/33/26
<i>TgPyK</i>								

Sequence Identities of Domains [in %]: All//A(A1,A2)/B/C {using “LALIGN” for All; “EMBOSS/NEEDLE” for A/B/C (domains)}

Assignment of Domains based on Structure-based Sequence Alignment

*CpPyK*: *6POY* (subunit A); domain A (A1:33-112; A2:212-389), domain B (113-211), domain C (390-526); All//A,B,C (33-526)

*LmPyK*: *3HQP* (subunit A); domain A (A1:18-88; A2:187-356), domain B (89-186), domain C (357-498); All//A,B,C (18-498)

*TcPyK*: *4KS0* (subunit A); domain A (A1:15-88; A2:187-358), domain B (89-186), domain C (359-499); All//A,B,C (15-499)

*EcPyK*: *IPKY* (subunit A); domain A (A1:1-70; A2:171-340), domain B (71-170), domain C (341-470); All//A,B,C (1-470)

*MtPyK*: *5WSB* (subunit A); domain A (A1:1-70; A2:168-336), domain B (71-167), domain C (337-472); All//A,B,C 1-472)

*ScPyK*: *IA3W* (subunit A); domain A (A1:19-88; A2:189-360), domain B (89-188), domain C (361-500); All//A,B,C (19-500)

*HsM2PyK*: *4FXF* (subunit A); domain A (A1:43-105; A2:221-387), domain B (106-220), domain C (388-530); All//A,B,C 43-530)

*TgPyK*: *3EOE* (subunit A); domain A (A1:59-124; A2:224-393), domain B (125-223), domain C (394-531); All//A,B,C (59-531)

Lalign (<https://www.ebi.ac.uk/Tools/psa/lalign/>); Emboss/Needle ([https://www.ebi.ac.uk/Tools/psa/emboss\\_needle/](https://www.ebi.ac.uk/Tools/psa/emboss_needle/))

### Supplementary Table S3

Conserved sequence in the hinge region (*LmPyK* residues 83-92)

KPYK2_ECOLI	41	HGSPEDHKMRADKVVREIAAKLG-----RHVAILGDLQGPKIRVSTFKEGKV---FLNIG	91
5WSB_Mtb	38	HGDYDDHKVAYERVRVASDATG-----RAVGVLADLQGPKIRLGRFASGAT---HWAEG	88
KPYK1_YEAST	54	HGSYEHKSVIDNARKSEELYPG-----RPLAIALDTKGPEIIRTGTTTNDV--DYPIPPN	106
KPYK2_YEAST	56	HGSYEFHQSVIENAVKSEQQFP-----RPLAIALDTKGPEIIRTGRTLNDQ--DLYIPVD	108
KPYK_LEIME	55	HGSHEYHQTINNVRQAAAELG-----VNIAIALDTKGPEIIRTGQFVGGDA---VMERG	105
Q4D9Z4_TRYCC	55	HGSHEYHQTINNLRAAATELG-----AHIGLALDTKGPEIIRTGLFKDGGI---ALAPG	105
KPYR_HUMAN	121	HGSHEYHAESIANVREAVESFAGSPLSYRPAIALDTKGPEIIRTGILQGGPESEVELVKG	180
KPYM_HUMAN_M1	78	HGTHEYHAETIKNVRTATESFASDPILYRPVAVALDTKGPEIIRTGLIKGSGTAEVELKKG	137
KPYM_HUMAN_PKM2	78	HGTHEYHAETIKNVRTATESFASDPILYRPVAVALDTKGPEIIRTGLIKGSGTAEVELKKG	137
CpPyK	79	HGDHESHFKTLQNIREEAAKARPH-----STVGIMLDTKGPEIIRTGMLEGGK--PIELKAG	131
Q969A2_TOXGO	91	HGDHETHARTVQNIQEAMKQRPE-----ARLAILLDTKGPEIIRTGFLKDHK--PITLQOG	143
C6KTA4_PLAF7	72	HGSHEDEHKEMFNNVLKAQELRPN-----CLLGMLLDTKGPEIIRTGFLKNK---EVHLKEG	123
KPYK1_ECOLI	37	HGDYAEHQRIQNLNRNVMK--KTG-----KTAAILLDTKGPEIIRTMKLEGGN--DVSLKAG	88
KPYK_STAA8	37	HGSHEEHKGRIDTIRKVAK-RLD-----KIVAILLDTKGPEIIRTHNMKDGII--IE-LERG	87
		** * .: * :***:** .	

KPYK2\_ECOLI: *E. coli* PyK2; 5WSB\_Mtb: *M. tuberculosis* PyK1; KPYPK1\_YEAST: Yeast PyK1; KPYPK2\_YEAST: Yeast PyK2;

KPYK\_LEIME: *L. mexicana* PyK; Q4D9Z4\_TRYCC: *T. cruzi* PyK1; KPYPK\_HUMAN: Human PyK-R;

KPYM\_HUMAN\_M1: Human PyK-M1; KPYPK\_HUMAN\_PKM2: Human PyK-M2; CpPyK: *C. parvum* PyK;

Q969A2\_TOXGO: *T. gondii* PyK1; C6KTA4\_PLAF7: *P. falciparum* PyK1; KPYPK1\_ECOLI: *E. coli* PyK2; KPYPK\_STAA8: *S. aureus* PyK1

**Table S4. Site interactions in representative PyKs.**

	<i>3HQP</i>	<i>4KSO</i>	<i>4FXF</i>	<i>1A3W</i>	<i>5WSB</i>	<i>6POY</i>
<b>Active Site</b>						
<b>ATP</b>	P29,R49,N51, H54,Y59,H60, R90,R175, D264,S330, G331,A334, K335; <b>OXL</b> , K <sup>+</sup> ,Mg <sup>2+</sup> (1,2)		T50,P53,R73, N75,H78,Y83, H84,R120, K207,D296, S362; <b>OXL</b> , K <sup>+</sup> ,Mg <sup>2+</sup> (1,2)			
<b>ADP</b>						T51,P54,R74, N76,H79, R116,S362, G363,A366, N367; Mg <sup>2+</sup>
<b>OXL</b>	K238,E240, A261,G263, D264,T296; Mg <sup>2+</sup> (1),ATP	K239,E241, A262,R263, G264,D265, T297; Mg <sup>2+</sup>	K270,E272, A293,R294, G295,D296, T328; Mg <sup>2+</sup> (1)		K218,E220, A241,R242, G243,D244, T276; Mg <sup>2+</sup>	
<b>PGA</b>				K240,E242, A263,R264, G265,D266, T298; K <sup>+</sup> , Mn <sup>2+</sup>		
<b>K<sup>+</sup> (1)</b>	N51,S53, D83,T84,S211	N52,S54, D84,T85,S212	N75,S77,D113, T114	N51,S53,D84, T85; <b>PGA</b>	N35,S37, D67	N76,S78, D109,T110
<b>K<sup>+</sup> (2)</b>	Q354,S355, L357,E359					
<b>Mg<sup>2+</sup> (1)</b>	E240,D264; <b>OXL</b> ,ATP	E241,D265; <b>OXL</b>	E242,D266; <b>OXL</b>		E220,D244; <b>OXL</b>	
<b>Mg<sup>2+</sup> (2)</b>	<b>ATP</b>		<b>ATP</b>			D170; <b>ADP</b>
<b>Mn<sup>2+</sup></b>				E242,D266; <b>PGA</b>		
<b>Effector Site</b>						
<b>AMP</b>					R351,F373, T374,Q375, S376,T379, A397,W398, V416,K418, M419,M425, A449,G450, P453,G454, T455,V456, G457,S458, T459	



G6P					L233,E267, N268,K270, H345,R348, T349,G352, R382,R385	
FDP	L399,S400, N401,T402, R404,S405, K453,R456, H480,A481, V485,K486, G487,A489	L400,S401, N402,S403, R405,S406, K454,R457, H481,A482, V486,K487, G488,Y489, P490				
FBP			L431,T432, K433,S434, S437,W482, R489,G514, R516,P517, G518,S519, G520,F521, T522	L401,S402, T403,S404, G405,T406, T407,W452, R459,G484, G490,H491, S492		

Ligand and metal binding sites grouped by 'Active site' and 'Effector site' with interacting residues (one-letter code) listings (highlighted in yellow: H-bonding residues) for PyK enzymes (*Lm*PyK: 3HQP; *Tc*PyK: 4KSO; *Hs*M2PyK: 4FXF; *Sc*PyK: 1A3W; *Mtb*PyK: 5WSB; *Cp*PyK: 6POY). Ligands and metal ions are highlighted in bold.

Protein residues that interact with ligands (substrates, substrate analogs, effectors) are from PDBsum (<http://www.ebi.ac.uk/pdbsum/>) with H-bonds and non-bonded contacts defined according to established criteria (H-bonds: Hydrogen-Acceptor atom distance < 2.7Å, Donor-Acceptor atom distance < 3.3Å, Donor-Hydrogen-Acceptor angle > 90°; Non-bonded contacts: ≤ 3.9Å).

**Table S5. Data collection and refinement statistics for holo CpPyK (6P0Y).***Data collection statistics*

Space Group	P2 <sub>1</sub> 2 <sub>1</sub> 2
Unit Cell Parameters [Å]	a = 83.45, b = 105.30, c = 133.50
Resolution [Å]	48.98 - 2.60 (2.72 - 2.60)
Unique reflections	36773 (4439)
Redundancy	4.9 (4.9)
Completeness [%]	99.5 (99.9)
Mean I/σ(I)	8.9 (1.3)
R <sub>merge</sub> [%]	18.5 (141.4)
R <sub>pim</sub> [%]	9.1 (69.9)
CC <sub>1/2</sub>	0.995 (0.510)

*Refinement statistics*

Resolution [Å]	48.98 - 2.60 (2.67 - 2.60)
No. of reflections	36730 (2692)
R <sub>work</sub> [%]	21.7 (33.5)
R <sub>free</sub> [%]	26.8 (37.9)
No. of protein atoms	7487
No. of ligand atoms	60 (2 Mg <sup>2+</sup> , 2 K <sup>+</sup> , 2 Cl <sup>-</sup> ; 2 ADP)
No. of waters	225
Wilson B-factor [Å <sup>2</sup> ]	38.4
Average B-factors [Å <sup>2</sup> ]	
Overall	50.7
Protein, Ligands, Waters	51.2, 46.7, 34.1
Correlation coefficient Fo-Fc	0.94
Correlation coefficient Fo-Fc free	0.90
<i>rmsd</i> bonds [Å]	0.014
<i>rmsd</i> angles [°]	1.72
Molprobit Score	1.32
Clash Score	3.02
Ramachandran	
favored	96.6%
outliers	0.3% (3)

**Table S6. Canonical dimers in *L. mexicana*, *T. cruzi*, *M. tuberculosis*, *S. cerevisiae*, *E. coli* PyKs versus covalent dimer in *C. parvum* PyK.**

Species	Ligands	ASA [Å <sup>2</sup> ]	BSA [Å <sup>2</sup> ]	$\Delta G_{\text{diss}}$ kcal/mol	$\Delta G_0$ kcal/mol	Interface residues	H-bonds	Salt bridges	Disulfide bonds
<i>Cp</i> PyK apo '4DRS'	GOL/SO <sub>4</sub> / ACT	39490	4171	34.3	17.1	109	15	7	2
<i>Cp</i> PyK holo '6P0Y'	Mg/K/ADP/ Cl	40663	5614	36.6	18.3	114	20	8	2 (4)
<i>Lm</i> PyK apo '3HQN'	K/GOL/SO <sub>4</sub>	38611	4444	21.5	10.7	117	13	8	-
<i>Lm</i> PyK holo '3HQO'	Mg/K/OXL/ ATP	37834	7789	25.9	12.9	154	25	4	-
<i>Tc</i> PyK apo '4KRZ'	K/GOL	39122	4335	25.7	12.9	120	22	12	-
<i>Mtb</i> PyK apo '5WRP'	PO <sub>4</sub>	34934	2519	0.5	-	63	11	11	-
<i>Mtb</i> PyK holo '5WS8'	Mg/OXL	36046	3599	6.7	3.4	86	14	6	-
<i>Sc</i> PyK holo '1A3X'	Mn/K/PGA	37234	5596	21.6	10.8	136	23	-	-
<i>Ec</i> PyK apo '1PKY'		35982	2114	7.4	3.7	60	13	13	-

Calculations and listings are based on PDBePISA (<https://www.ebi.ac.uk/pdbe/pisa/>).

$\Delta G_0$  is the standard free energy of dissociation into monomeric units (ground state), in kcal/mol.

$\Delta G_{\text{diss}}$  indicates the free energy of assembly dissociation, in kcal/M. The free energy of dissociation corresponds to the free energy difference between dissociated and associated states. Positive values of  $\Delta G_{\text{diss}}$  indicate that an external driving force should be applied in order to dissociate the assembly, therefore assemblies with  $\Delta G_{\text{diss}} > 0$  are thermodynamically stable.

Reference: E. Krissinel and K. Henrick (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372, 774-797.

## Legends for Supplementary Figures

### **Figure S1. Close-up view of the A-domain.**

The A-domain of *LmPyK* consists of residues 18-88 and 187-356.

Secondary structural elements of the A-domain of *LmPyK* are labeled in the cartoon drawing.

The A-domain has a  $(\beta/\alpha)_8$  barrel structure. ATP is shown in stick model.  $K^+$  and  $Mg^{2+}$  are shown as violet and green spheres. Asp83 which coordinates  $K^+$ , and Glu88 which is implicated in  $K^+$ -dependency of PyKs, are also shown in stick model and labeled.

### **Figure S2. Close-up view of the C-domain.**

The C-domain of *LmPyK* consists of residues 357-496.

Secondary structural elements of the C-domain of *LmPyK* are labeled in the cartoon drawing.

FDP bound in the effector binding pocket is shown in stick model.

### **Figure S3. Inter-subunit interactions involving effector loop residues at the C-C interface.**

A close up view of the C-C interface in the *LmPyK* structure (*3HQP*). C-domains of the monomers forming the small interface are depicted in light cyan and cyan in the cartoon diagram. Sidechains of Asp482 and Lys484 of each monomer form hydrogen bonds with sidechains of Arg493 and Glu498 of the other monomer at the C-C interface. These residues and the FDP molecules are shown as stick models.

### **Figure S4. Interactions between monomers at the A-A interface.**

Interactions between monomers (shown in light yellow and light orange) across the *r*-axis form the large interface. Mainly residues in  $A\alpha 6$ ,  $A\alpha 7$  and  $A\alpha 8$  helices participate in the interactions. N-helix of each monomer (shown in teal and grey) also shows minor interactions with the A-domain of a neighboring monomer. Portion of the ATP molecule bound to one subunit is visible.

**Figure S5. Comparison of ADP and metal ion binding sites in holo-*CpPyK* and *LmPyK* substrate activator complex.**

Close up view of the active sites of *CpPyK* and *LmPyK*. Cartoon drawing shows superposition of holo-*CpPyK* (*6P0Y*) in blue and *LmPyK* (*3HQP*) in light orange.

*CpPyK*: ADP stick model (C atoms in blue); K<sup>+</sup> (blue) and Mg<sup>2+</sup> (cyan) spheres.

*LmPyK*: ATP stick model (C atoms rose); K<sup>+</sup> (light yellow) and Mg<sup>2+</sup> (light orange) spheres.

The *LmPyK* structure contains an additional Mg<sup>2+</sup> ion.

**Figure S6. Effector loop conformation in *TcPyK* in the effector-free and effector-bound states.**

Cartoon diagram showing superposition of the C-domain residues in *TcPyK* without FDP (*4KRZ*; light orange), FDP-bound state (*4KS0*; cyan) and in a dye-bound state (*3QV9*; light blue). The effector FDP shown in stick model (C: green) is bound in the effector-binding pocket. Lys454 and Arg457 sidechains move closer to the effector-binding site and make contacts with the 2'-phosphate group. The FDP 6'-phosphate interacts with the residues in the phosphate-binding loop (magenta). The effector loop (dark purple) is disordered in the effector-free structure (*4KRZ*) and becomes ordered when FDP is bound (*4KS0*). This loop is also ordered in the structure (*3QV9*) bound with Ponceau-S dye, although the effector-site is empty.

**Figure S7. Rock and lock model for allosteric transition in Trypanosomatid PyKs.**

A. Dimers across the C-C interface in *TcPyK* FDP complex (*4KS0*) subunits; A (light yellow) and B (pale cyan) are shown in cartoon drawing. The effector loop in the A and B subunits are colored in orange and slate, respectively. Asp483 on each effector loop interacts with Arg494 on the molecule across the C-C interface, thus locking the tetramer in a stable assembly. The effector loop is disordered in the absence of FDP.

- B. Assembly of *LmPyK* (substrate plus activator complex) tetramer R-state (*3HQP*). Cartoon drawing shows subunits colored by chain. C-C interface and binding site for FDP are labeled.
- C. Close-up view of the interactions between subunits across the C-C interface. In addition to the interactions between Asp482 and Arg493 of the monomers across the interface, interactions of Lys484 and Glu498 between the subunits further stabilize the tetramer.

**Figure S8. Allosteric transitions in Human PyK-M2.**

- A. C-C interface in human PyK-M2 in the absence of the activator (*3SRH*) in the T-state. Trp481 side chain orients inwards upon FBP binding. Lys421 of each chain inserts into the hole formed by residues 390-420 (shown in orange and cyan) of the other monomer across the C-C interface. Cartoon drawings for chains across the C-C interface are in white and pale cyan.
- B. Cartoon diagram showing the C-C interface in human PyK-M2 OXL, ATP and FBP bound state (*4FXF*). Trp482 sidechain is oriented inward.
- C. Priming the neighboring active site in the R-state. Interactions at the A-A interface between A (white) and D (light yellow) subunits in human PyK-M2 bound with OXL, ATP and FBP (*4FXF*). Network of interactions among the active site ligand (OXL), Mg<sup>2+</sup> ion (green sphere), residues 294-296 of the A-chain and Arg342 of the D-chain in the R-state. Interactions of Arg342 stabilizes the glycine rich Aα6' helix. Lack of this inter-subunit interaction may be responsible for the inactivity of the monomer.

**Figure S9. Effector loop in holo-CpPyK (*6P0Y*).**

- A. Effector loop becomes ordered in holo-CpPyK structure. In both monomers in the asymmetric unit the entire loop (residues 508-517) could be built unambiguously. In

both monomer Cys513 and Cys517 were linked by a disulfide bond. Fo-Fc omit map (Polder) was generated using the Phenix program suite (Afonine. *et al.*, 2012) and contoured at  $3\sigma$  level. Residues 508-517 were omitted from both molecules in the asymmetric unit for creating the map. Residues in the region are shown in stick model (C: salmon) and the only ordered residues (507 and 518) in the apo-CpPyK structure are also shown (C: cyan). The disulfide bond between Cys513-Cys517 is indicated.

B. Omit map contoured at  $3\sigma$  level for the effector loop residues in monomer B in holo-CpPyK.

**Figure S10. Ordering of the effector loop in holo-CpPyK structure stabilizes the C-C dimer interface**

- A. Cartoon diagram showing apo-CpPyK tetramer (molecules colored by chain). Residues in the effector loop (508-517) were missing.
- B. Tetramer in holo-CpPyK structure in a similar orientation as in A. Interactions between the monomers at the C-C interface is stabilized by hydrophobic interactions between residues in the loop (including Pro514) which inserts into a hydrophobic area in the monomer across the interface. Pro514 is shown as dark purple blue spheres.
- C. Close up view of the interactions between the residues in the effector loop with those in the neighboring monomer across the C-C interface. Also shown the interaction between Glu433 and Lys510 of the same subunit. These interactions are absent from the apo-CpPyK structure because of the disorder in the loop.

**Figure S1**

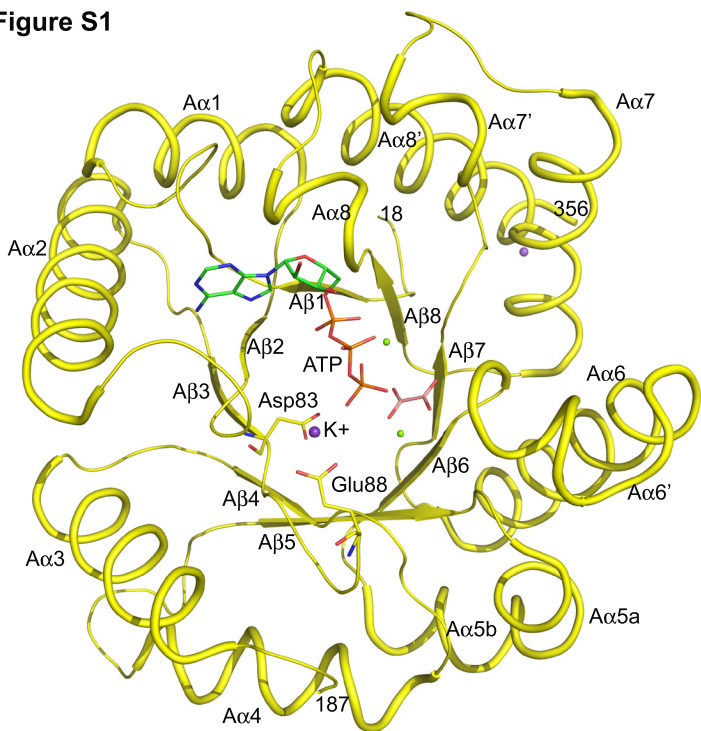
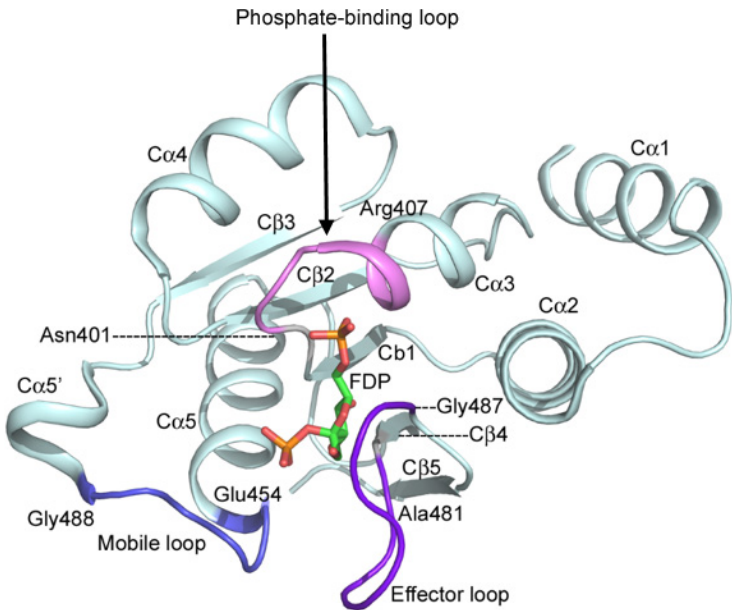




Figure S2



**Figure S3**

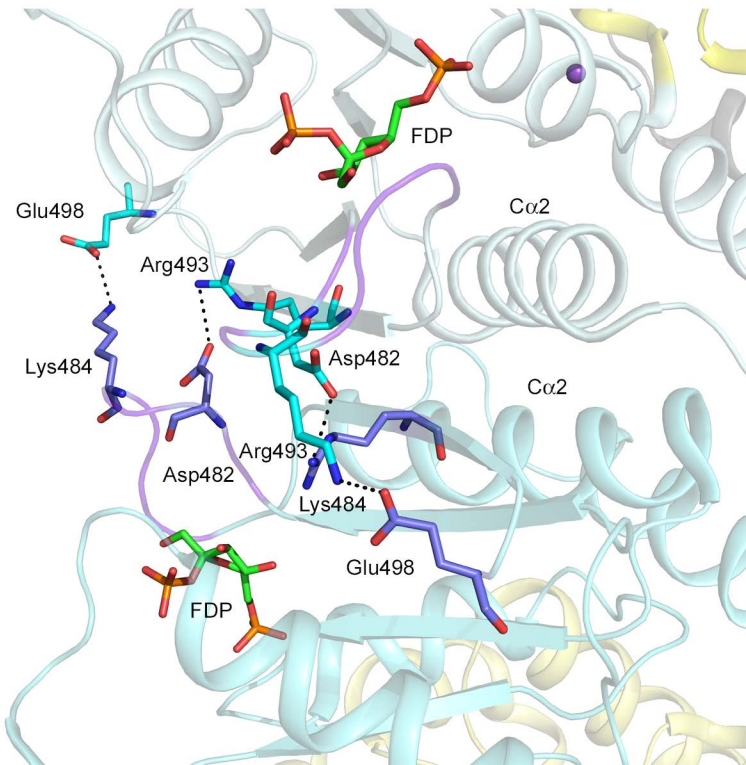


Figure S4

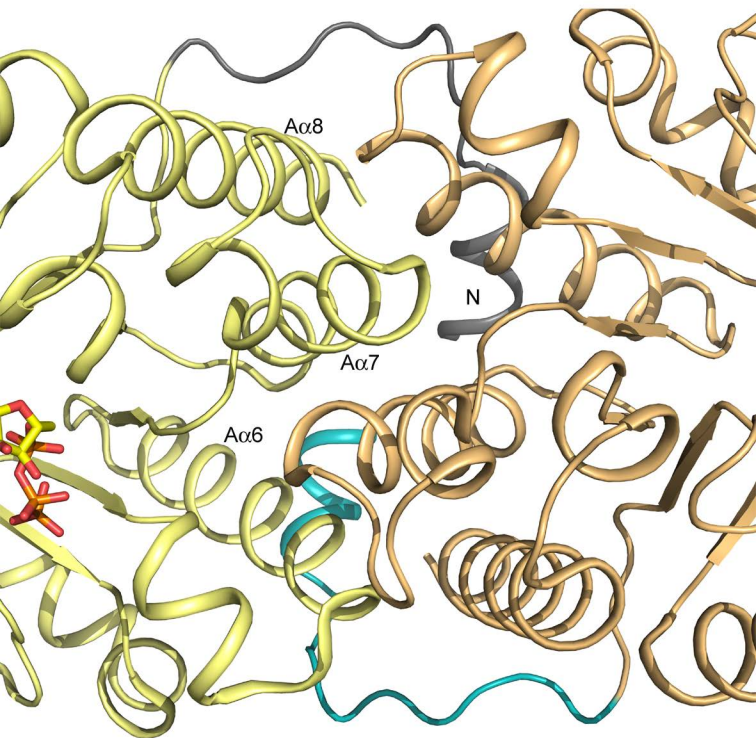
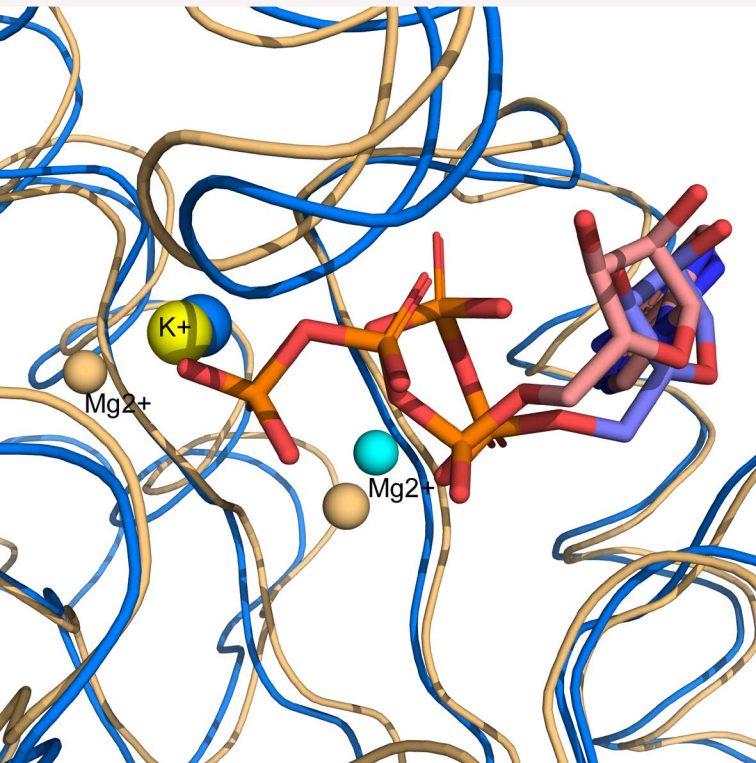
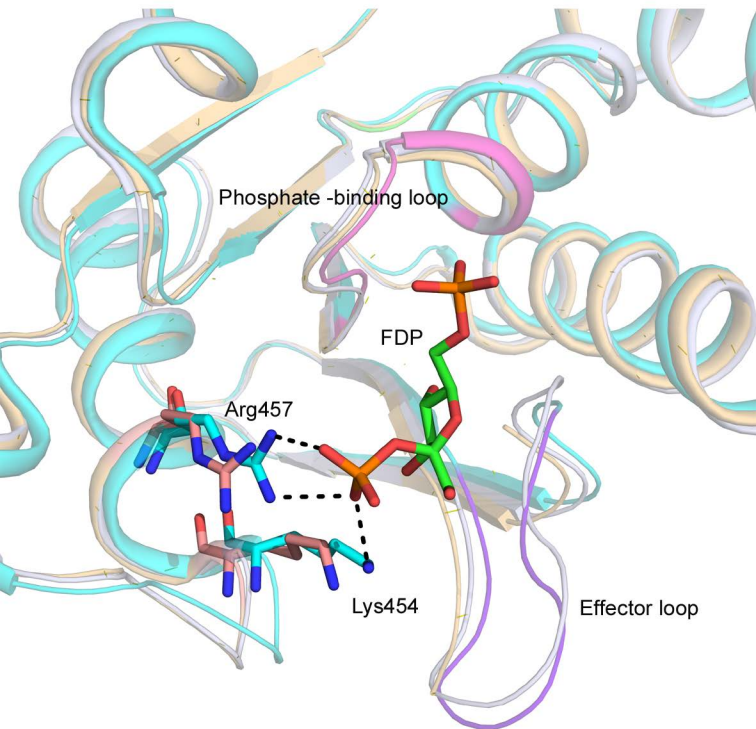


Figure S5

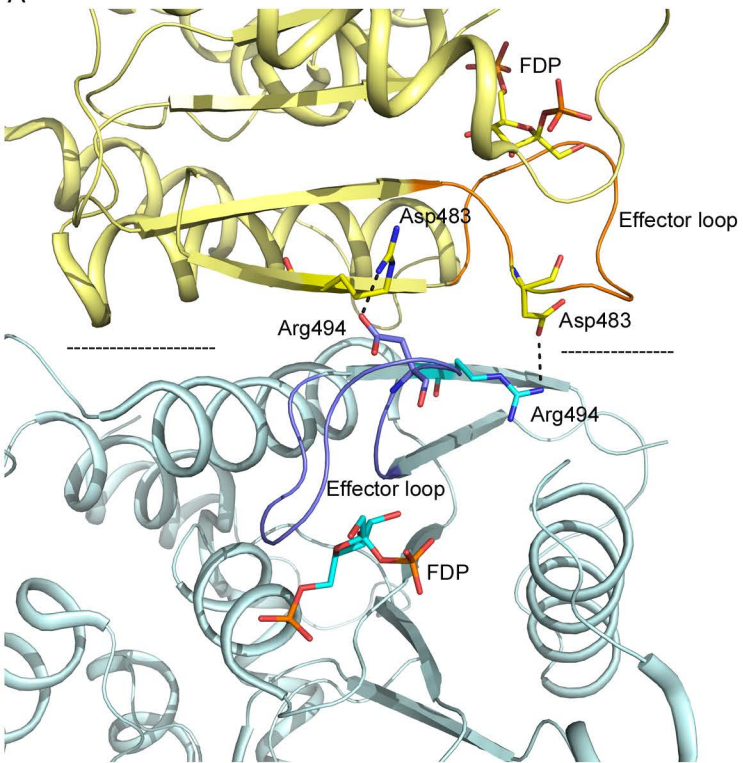


**Figure S6**

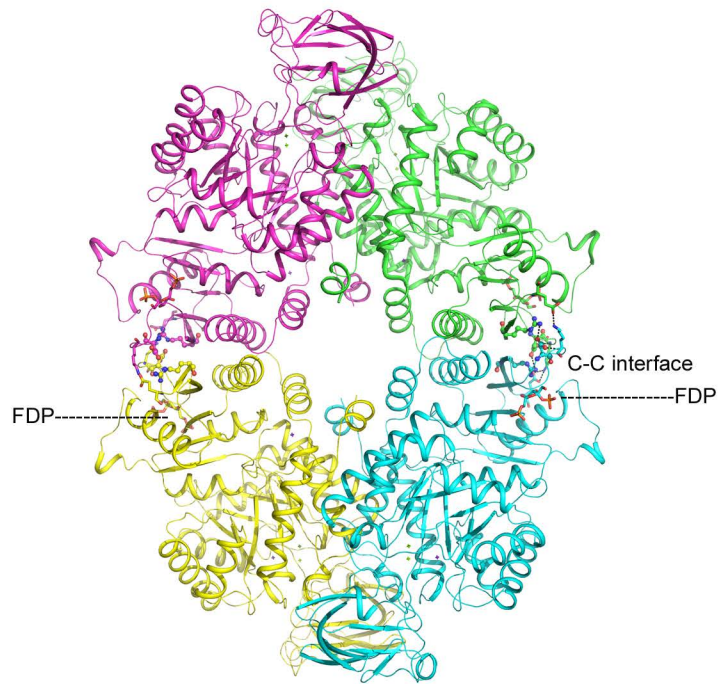


**Figure S7**

**A**



**B**



**C**

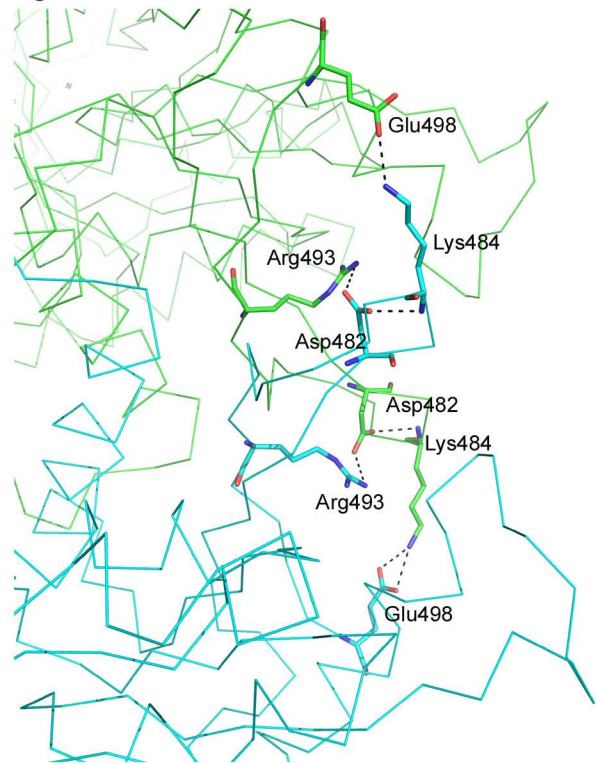
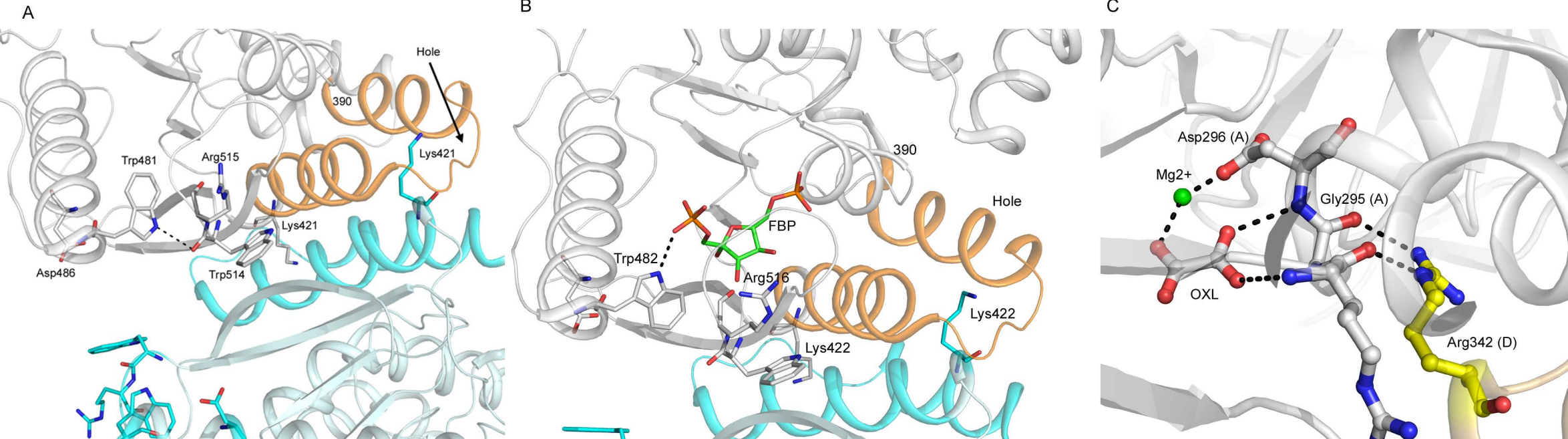
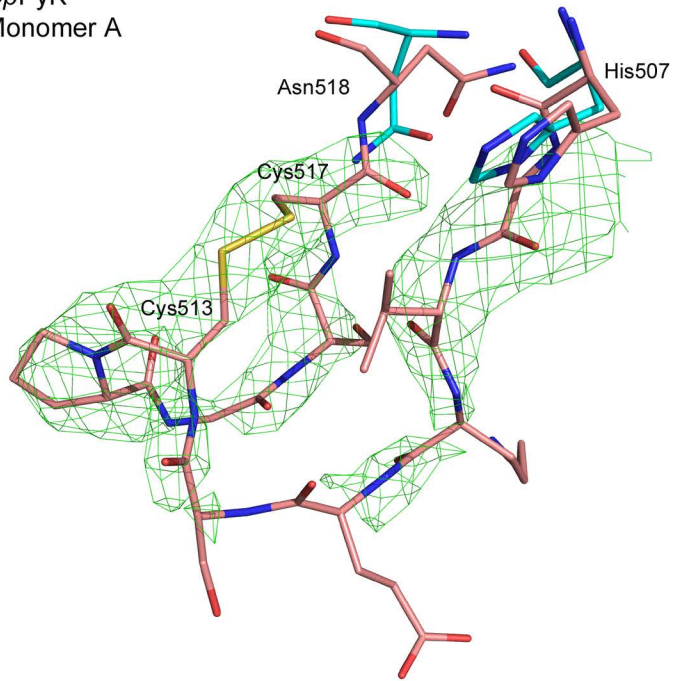


Figure S8

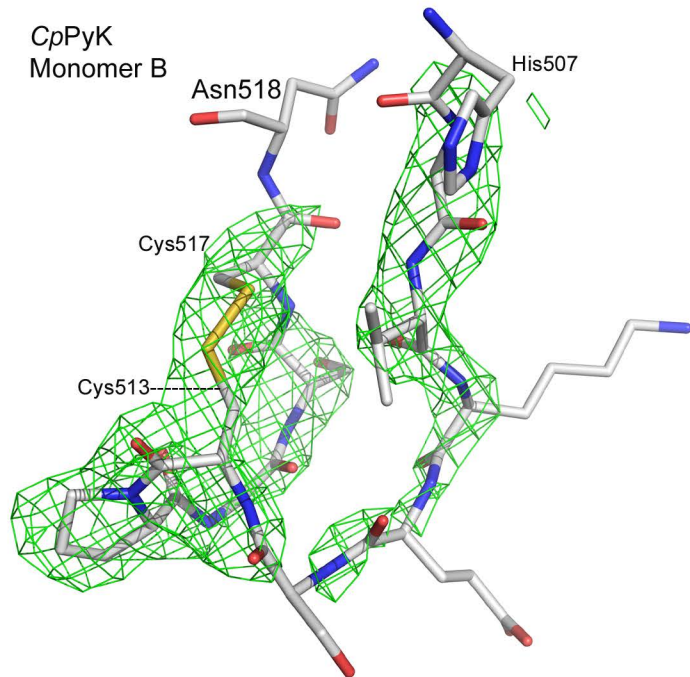


**Figure S9**

*CpPyK*  
Monomer A



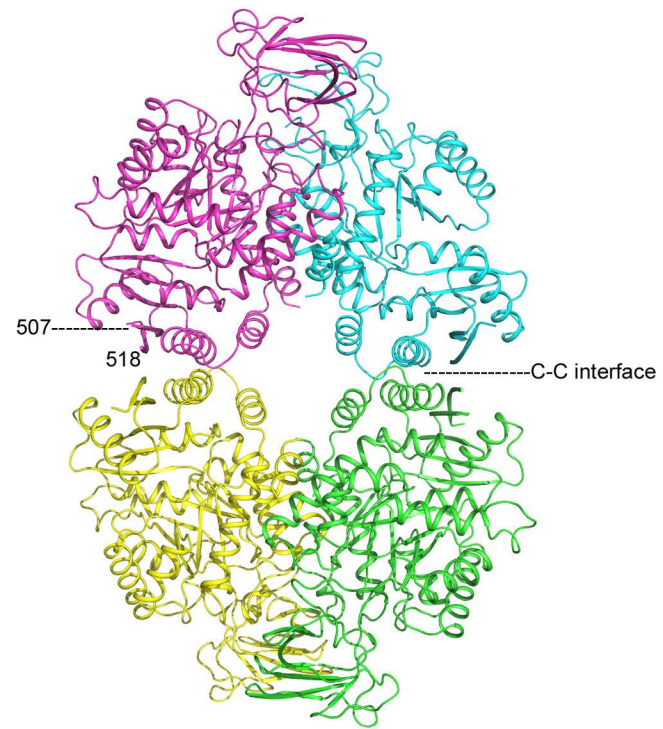
*CpPyK*  
Monomer B



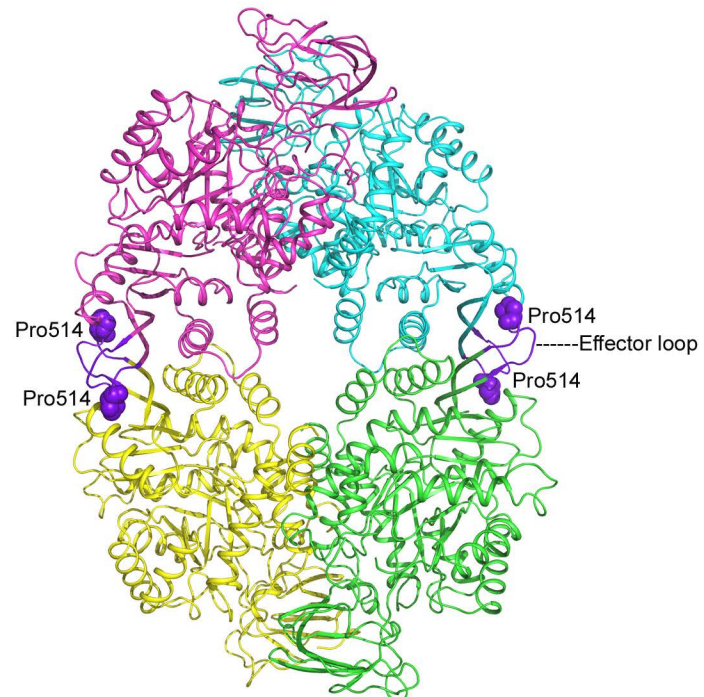


**Figure S10**

**A**



**B**



**C**

