Structural and Biochemical Insights into Previously Uncharacterized Activity of Nucleotide-Dependent Phosphoenolpyruvate Carboxykinases

by

Sarah Alexis Erica Barwell

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Biology

Waterloo, Ontario, Canada, 2024

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Andrew Gulick Professor University of Buffalo
Supervisor(s)	Todd Holyoak Associate Professor Biology
Internal Member	David Rose Adjunct Professor Biology
Internal Member	Barb Moffatt Adjunct Professor Biology
Internal-external Member	Thorsten Dieckmann Associate Professor Chemistry

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Phosphoenolpyruvate carboxykinase (PEPCK) enzymes are central to glucose metabolism, with their main role in catalyzing one of the first steps of gluconeogenesis. In addition to its primary role, PEPCK is also involved, whether directly or indirectly, in glyceroneogenesis, amino acid metabolism, and lipogenesis. The PEPCK enzyme is responsible for the conversion of oxaloacetic acid to phosphoenolpyruvate using a triphosphate nucleotide as a phosphoryl donor. This dissertation focuses on human cytosolic PEPCK and *Escherichia coli* PEPCK as representative enzymes of the PEPCK family.

The hcPEPCK enzyme is well characterized and extensively studied. Recent research suggests a new role for this enzyme, and potentially other GTP-dependent PEPCKs, as a protein kinase. hcPEPCK is thought to phosphorylate the ER protein INSIG, resulting in a cascade effect leading to increased lipogenesis. Expanding upon the information in the literature, phosphomimetic mutants were used to study the effects of phosphorylation on hcPEPCK, and its subsequent interactions with the INSIG protein.

The ATP-dependent PEPCKs are less studied than their GTP counterparts, and some of the research in the literature is not up to date. Historically, the bulk of the research has involved GTP-dependent PEPCK. Using *E. coli* PEPCK as a proxy for the ATP-PEPCK class, a full kinetic and structural characterization was performed to further our understanding on these enzymes and compare differences and similarities between the classes. In addition, a new ^{31}P-NMR assay was designed to address the hypothesis that ATP-dependent PEPCK can catalyze two additional reactions: the conversion of phosphoenolpyruvate to pyruvate (pyruvate kinase-like activity), and the nucleotide-dependent catalysis of phosphoglycolate.

Acknowledgements

Once again, I would like to acknowledge my supervisor, Dr. Todd Holyoak, for allowing me the opportunity to work in his laboratory six years ago when I first started my M.Sc. degree. He was the one who offered me the chance to continue working in this field and do my Ph.D. research in the same lab. I am grateful for the opportunity to attend graduate school, which would not be an option without the funding and guidance provided by Todd. Todd is a great supervisor and colleague, who has encouraged me to attend conferences and seminars where I found my courage to enjoy giving presentations and even earned a few awards because of it.

I would also like to thank my committee members, Dr. David Rose and Dr. Barb Moffatt, for offering their unique perspectives and advice towards my thesis project, as well as finding time in their schedules to attend meetings, answer emails and read my reports. My thesis would not complete without their feedback and critiques. Both of my committee members are easy to talk to and get along with.

Special thanks go to April Wettig, the administrative coordinator of graduate studies for the Department of Biology. She has a wealth of knowledge on almost everything a graduate student might possibly need to know, and is always quick to answer any inquires in a timely manner. She was essential in ensuring I met degree requirements on time. A quick acknowledgement and heartfelt thanks to Dr. Carol Cameron in the Chemistry department, who allowed me to work in her office when our lab was under construction and renovation. I would also like to acknowledge both Janet Venne and Blossom Yan, NMR technicians from the Department of Chemistry, who trained me on the 300 MHz NMR spectrometer.

Acknowledgements would not be complete without a nod towards my fellow graduate students in the Holyoak Lab, both new and old. The ‘old guard’ of the lab were quick to lend a hand explaining new techniques and how to use laboratory equipment when I first began my degree, and we still bounce ideas off each other to this day, even if some of them have left the lab. My peers were essential in creating a friendly and supportive environment, and it was a wonderful experience heading out for a lab coffee break once a week.

Dedication

I dedicate my thesis to my mother, for her support and enthusiasm towards the sciences. Though she has passed, I know she would be proud of how far I've come. Her long-time passion for the sciences encouraged me to get this far in my studies, and I will hold her memory with me through future endeavours, hoping to inspire others like she did.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Å	angstroms
A280	absorbance at 280 nm
A340	absorbance at 340 nm
ADP	adenosine diphosphate
AI	auto-induction
Akt1	Protein kinase B/Akt kinase
AMP	adenosine monophosphate
ASU	asymmetrical unit
ATP	adenosine triphosphate
ATPyS	adenosine 5'-O-(3-thio)triphosphate
BL21(DE3)	competent E. coli expression cell line
BSP	β-sulfopyruvate
C-PEPCK	cytosolic phosphoenolpyruvate carboxykinase
CCP4	Collaborative Computational Project No.4
Coot	Crystallographic object-oriented toolkit
CuKα	rotating copper anode – source of incident x-ray beam
DIALS	Diffraction Integration for Advanced Light Sources
dNTP	deoxynucleotide triphosphate
DSF	Differential Scanning Fluorimetry
DTT	dithiothreitol
ε	molar extinction coefficient
EC	Enzyme Commission number

FFT	Fast Fourier Transform
FRET	fluorescence resonance energy transfer
GDP	guanosine diphosphate
GTP	guanosine triphosphate
HCC	hepatocellular carcinoma
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His6	six histidine tag
IMAC	immobilized metal affinity column
INSIG	insulin-induced gene 1 protein
IPTG	isopropyl β- d-1-thiogalactopyranoside
Kan	kanamycin
kcat	turnover number; amount of substrate converted into product per unit time
kDa	kilodaltons
keV	kiloelectron volts
K_I	kinetic inhibition constant
K_M	Michaelis constant, equivalent to the substrate concentration at half V_{max}
LB	lysogeny broth
LB-Kan	bacterial growth media containing lysogeny broth and kanamycin antibiotic
LDH	lactate dehydrogenase
MAD	multi-wavelength anomalous dispersion
MDH	malate dehydrogenase
M-PEPCK	mitochondrial phosphoenolpyruvate carboxykinase
MOLREP	molecular replacement program

Mpro	SARS-CoV-2 main protease
MW	molecular weight
NADH	nicotinamide adenine dinucleotide
Ni-NTA	nickel nitrilotriacetic acid resin
NPS	nitrogen, phosphate and sulphate salts
OAA	oxaloacetic acid
OD600	optical density or absorbance at 600 nm
P6DG	polyacrylamide 1000 - 6000 MW fractionation range desalting gel
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEG	polyethylene glycol
PEP	phosphoenolpyruvate
PGA	phosphoglycolate
PK	pyruvate kinase
Pi	inorganic phosphate
PPi	pyrophosphate
rcf	relative centrifugal force
rcPEPCK	rat cytosolic phosphoenolpyruvate carboxykinase
rpm	rotations per minute
RT	ambient/room temperature
SCAP	SREBP cleavage activating protein
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEC	size exclusion chromatography

SREBPs	sterol regulatory element binding proteins
SUMO	small ubiquitin-like modifier
TCEP	tris(2-carboxyethyl)phosphine
TOP10	competent *E. coli* cloning cell line
TS	transition state of a chemical reaction
Vmax	maximum reaction rate in enzyme kinetics
WT	wildtype
XL10	competent *E. coli* cloning cell line

Chapter 1: Introduction to phosphoenolpyruvate carboxykinases and a brief review of the literature

1.1 Historical characterization of PEPCK in the literature

PEPCK enzymes, as we know them today, were first reported on in the early 1950's and were initially named oxalacetic carboxylases. The earliest known report by Utter and Kurahashi describes the enzyme isolated from chicken liver, and references the fixation of CO_2 generating OAA, requiring a nucleotide diphosphate as a cofactor.[1,2] Subsequent studies showed that the enzyme preferentially catalyzed the reaction in the direction of net OAA decarboxylation.[3] Up to this point in time, it was assumed all PEPCK enzymes belonged to the same class and used ATP as their phosphoryl donor nucleotide, until nucleotide specificity was determined in 1957, and that chicken liver PEPCK preferentially used GTP nucleotide instead.[4] By the 1960's these enzymes were renamed phosphoenolpyruvate carboxykinases due to their mechanism of action and their role in gluconeogenesis, and further characterization of PEPCK enzymes continued since then.

PEPCK enzymes have primarily been characterized to both understand their metabolic role *in vivo*, as well as being used as model enzymes to study more general structure-function relationships and the role of dynamic processes in catalytic function. As the field of enzymology grows, our understanding of PEPCK function and the enzyme family as whole increases steadily. Our collective knowledge of PEPCK has expanded to include three different families, ATP- (EC 4.1.1.49), GTP- (EC 4.1.1.32) and the recently re-classified PP_i-dependent PEPCKs (EC 4.1.1.38). Structural insight into the family was first achieved when the first PEPCK crystal structure, derived from *E. coli*, was reported by Matte et al., in 1996.[5,6] The first GTP-dependent structure of the human cytosolic form was later solved in 2002.[7] Lastly, only two crystal

structures of the PP_i-dependent PEPCK enzyme have been submitted as of date, the first of which was PP_i-PEPCK from *Actinomyces israelii* in 2019, with the latter from *Propionibacterium freudenreichii* being determined only recently in 2023.[8]

1.2 PEPCK's role in gluconeogenesis and central metabolism

Phosphoenolpyruvate carboxykinase, or PEPCK for short, is a key metabolic enzyme known primarily for its role near the beginning of the gluconeogenic pathway. Through this pathway, PEPCK is involved in regulating glucose levels by catalyzing the rate-limiting conversion of oxaloacetic acid (OAA) to phosphoenolpyruvate (PEP), ultimately regenerating the supply of glucose in the cell.[9] PEP has the greatest high-energy phosphate bond available to a biological system, and the conversion to pyruvate (via pyruvate kinase – Figure 1A) at the end of glycolysis is generally unidirectional *in vivo*.[9–11] The PEPCK-catalyzed reaction is therefore one of two reactions needed to regenerate the supply of PEP in the cell, the other being the pyruvate carboxylase catalyzed conversion of pyruvate to OAA (Figure 1B).[10] These three enzymes, pyruvate kinase, pyruvate carboxylase and PEPCK lie at the center of the pyruvate – PEP – OAA node that is central to primary metabolism and carbon flux.[12,13]

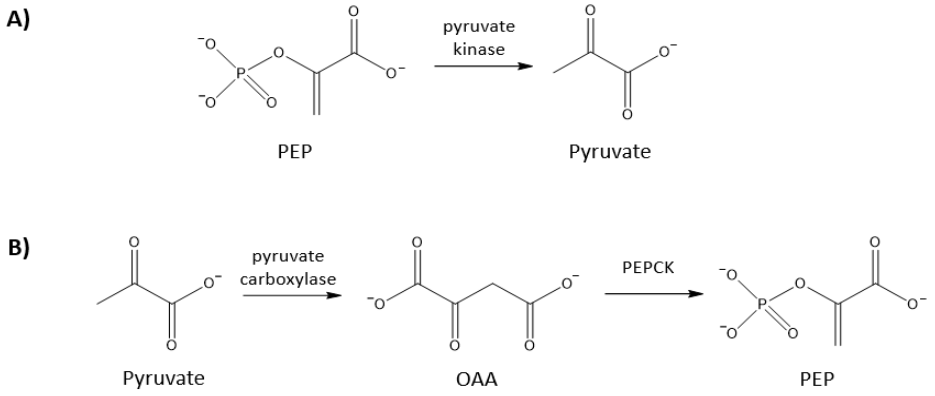


Figure 1: Carbon flux: the PEP-pyruvate-OAA node
A) The irreversible formation of pyruvate from PEP catalyzed by pyruvate kinase, as the last step of glycolysis. **B**) Regeneration of PEP in the gluconeogenic pathway, catalyzed by pyruvate carboxylase and PEPCK.

In eukaryotes, the majority of the gluconeogenic pathway occurs in the cytoplasm, with the exception of the initial conversion of pyruvate to OAA by pyruvate carboxylase, which occurs in the mitochondria.[9,14] Since the rest of the gluconeogenic enzymes that convert PEP to glucose-6-phosphate reside in the cytosol, the missing link is PEPCK. GTP-dependent PEPCK enzymes exist in two forms, cytosolic (C-PEPCK) and mitochondrial (M-PEPCK), although cytosolic PEPCK is thought to be the primary contributor to gluconeogenesis.[15,16] Some species have only one form, while in others PEPCK is split between the two isozymes.[17,18] There is high sequence conservation between the two isozymes (68% identity, 86% similarity between the human isoforms), and their structures are nearly identical.[19,20] Henceforth, all discussion of GTP-dependent PEPCK will refer to that of its cytosolic form.

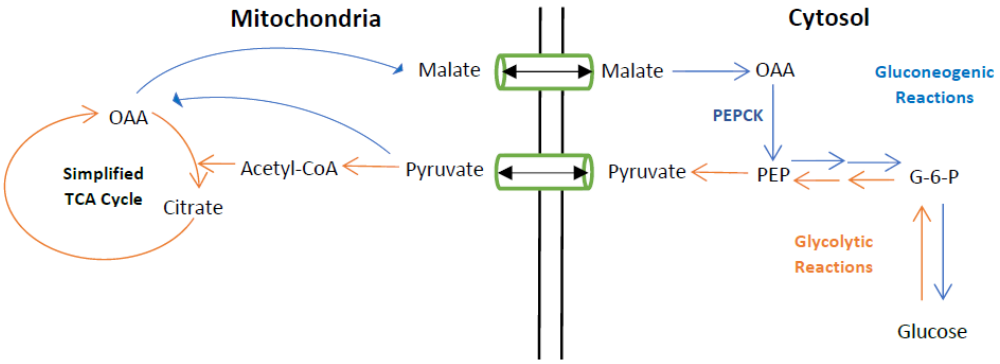


Figure 2: Simplified reaction map of glucose metabolism in the cell.
Glycolytic and gluconeogenic reactions, colour-coded in orange and blue, respectively, occur partially in the mitochondrial matrix (left) and partially in the cytosol (right). Pyruvate, malate and PEP can cross the inner mitochondrial membrane via specialized transport proteins, while OAA cannot traverse the membrane.

The PEPCK-catalyzed conversion of OAA to PEP involves the transfer of a 'high energy' phosphoryl group (Figure 3). PEPCK uses a donor nucleoside triphosphate (NTP) to provide the phosphate group, and transfers it onto the enolate form of pyruvate (formed as an intermediate), generating the PEP product.[21] Although this reaction is largely unidirectional as seen in the gluconeogenic pathway *in vivo*, the reaction is thermodynamically reversible *in vitro*.[14]

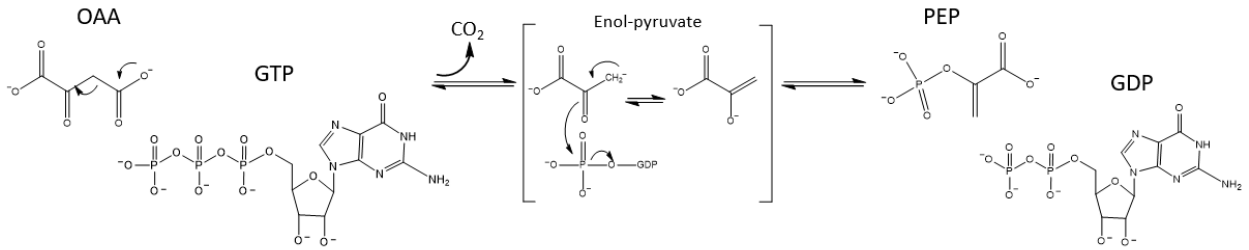


Figure 3: PEPCK-catalyzed reaction in higher eukaryotes.
The PEPCK-catalyzed reaction written left to right in the physiological direction of gluconeogenesis. Decarboxylation occurs first in the stepwise reaction, followed by the phosphorylation of the intermediate by GTP.[21]

1.3 Structural classification of different PEPCK enzymes

Phosphoenolpyruvate carboxykinases belong to the lyase family of enzymes (EC 4), specifically the carboxy-lyases (EC 4.1.1). As mentioned earlier, the PEPCK-catalyzed reaction requires a phosphoryl donor to phosphorylate the enolate intermediate and produce PEP. In higher eukaryotes, including humans, guanosine triphosphate (GTP) is used as the phosphate donor. However, prokaryotic organisms and some lower eukaryotes use adenosine triphosphate (ATP) instead. This has resulted in further classification of PEPCK enzymes into two distinct families: ATP-dependent PEPCK (EC 4.1.1.49) and GTP-dependent PEPCK (EC 4.1.1.32). As a general

rule, an organism will have either of the two types of PEPCK, but not both, though there are exceptions.[22] Both ATP-dependent and GTP-dependent PEPCKs are found across all three branches of life, including archaea.[21] While it has been observed generally that higher eukaryotes (human, rat, chicken) have GTP-dependent PEPCK and lower eukaryotes/prokaryotes (plants, yeast, *E. coli*, etc.) have ATP-dependent PEPCK, it is not always the case, as several bacterial species have been discovered to possess the GTP-dependent isozyme. Current research is not yet decided upon whether these two classes arose from convergent or divergent evolution, though current evidence suggests it is likely the latter.[8,23]

ATP-dependent PEPCK and GTP-dependent PEPCK are not very similar from a sequence conservation perspective, as the amino acid percent identity is only ~20%. However, these two isozymes share global structural homology (Figure 4), with a globular, bi-lobal structure consisting of N- and C- terminal domains of roughly equivalent size, with the active site situated in the junction of where these two domains meet.

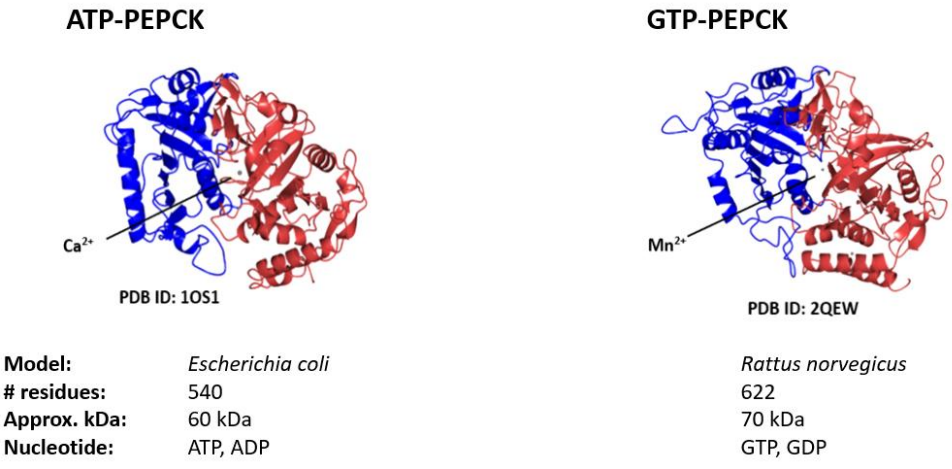


Figure 4: Structural comparison of ATP and GTP-dependent PEPCKs
Two representative structures were chosen from the PDB to represent the ATP- and GTP-dependent classes of PEPCK. The N- and C-terminal domains are coloured blue and red, respectively.

Both isozymes undergo dynamic changes from open to closed states upon substrate binding, analogous to a clam closing its shell to protect a pearl within. Much of the analysis that follows has been determined via the rat cytosolic enzyme. The PEPCK active site contains three catalytically active mobile loops: the P loop that is involved in nucleotide binding, the R loop that coordinates the OAA/PEP substrate, and the Ω loop that protects the active site from unwanted chemistry after substrate binding (Figure 5).[24,25] The Ω-loop, aptly named for its resemblance to the Greek letter omega, exists in a dynamic equilibrium between open and closed states.[24] Consistent with the dynamic requirements of the lid, the Ω-loop is a flexible element of the structure, and thus in some crystal structures the loop is so disordered that it can not be modelled unless the crystal lattice has "captured" it in a stable conformation, usually via the binding of substrates or other substrate analogs.

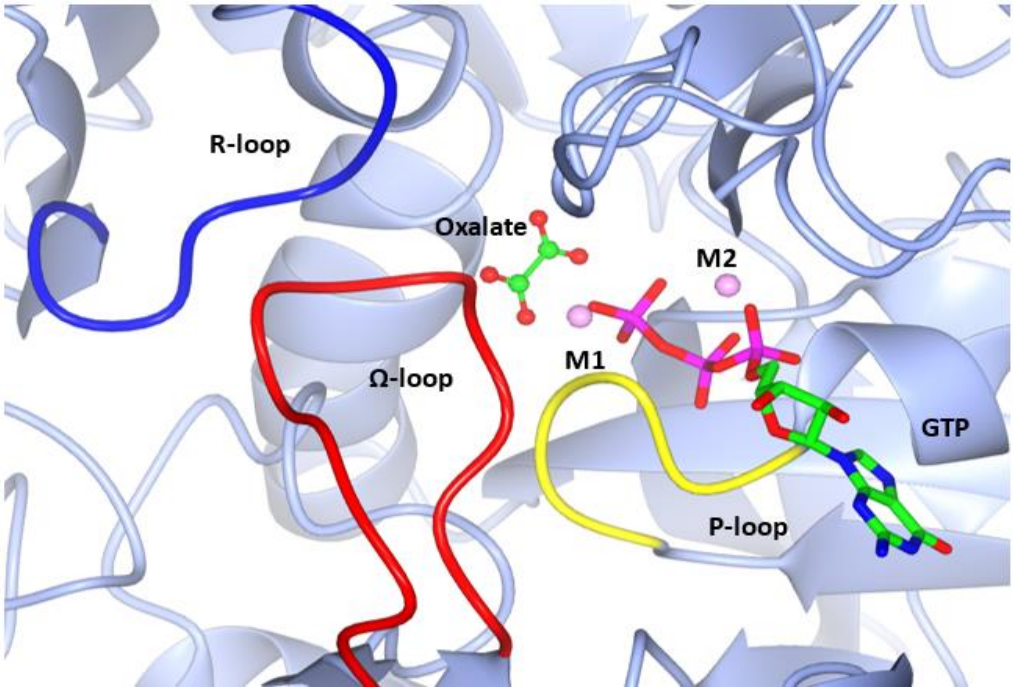


Figure 5: Catalytically active mobile loops in the PEPCK active site
Ribbon structure of rat cytosolic PEPCK is coloured ice blue. The R-loop is coloured blue, the P-loop is coloured yellow, and the omega loop is coloured red. Both the oxalate ligand and the GTP nucleotide are coloured by atom-type and rendered as ball-and-stick models. Manganese ions are coloured light pink and represent the M1 and M2 metals. This diagram was generated using CCP4MG.[26]

The active site also contains two different divalent metal cations, referred to as the M1 and M2 metals, respectively.[24] Binding of the M1 metal, typically manganese, generates the functional holoenzyme. The M2 metal, usually magnesium, binds as a metal-nucleotide complex, serving to stabilize the negative charge on the phosphates of the di/trinucleotide.

Superposition of ATP-PEPCK and GTP-PEPCK protein structures (Figure 6) reveals that the three catalytically active mobile loops are nearly identical in structure. The R and P loops, involved in the binding of the OAA/PEP and nucleotide substrates, respectively, and the omega loop lid are superposed nearly perfectly. Of the low percentage of amino acids that are identical, it stands to reason that they include the residues that are directly involved in substrate and product binding. A pairwise sequence alignment compares the mobile loop sequences between ATP and GTP-dependent PEPCK (Figure 7).

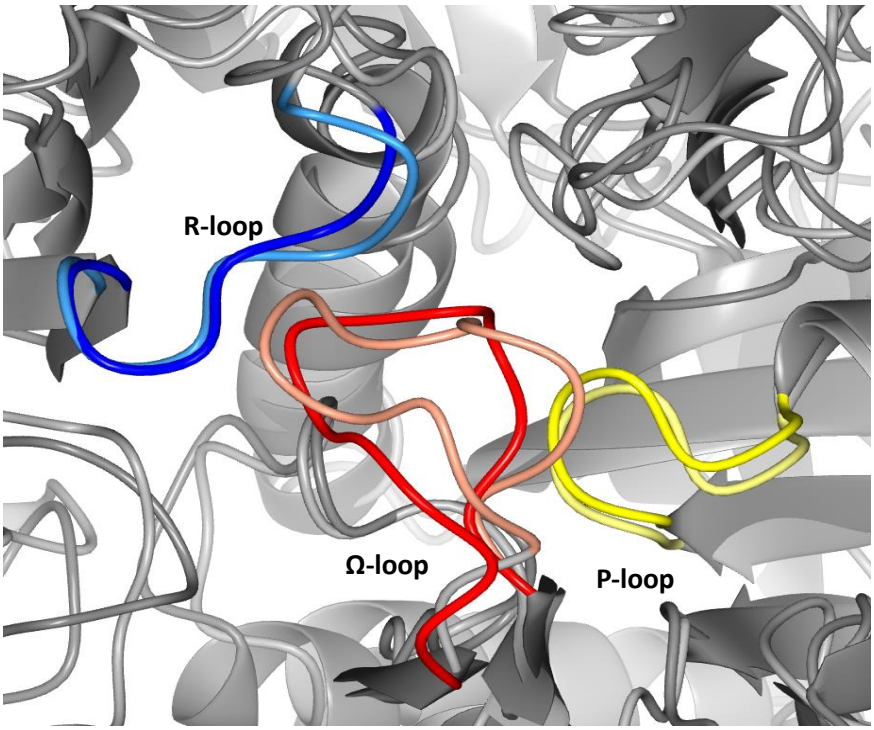


Figure 6: Superimposition of the active sites of *E. coli* and rat cytosolic PEPCK
R loop is coloured blue, P loop is coloured yellow, and omega loop is coloured red. GTP-PEPCK structure is represented by the darker coloured loops, while the ATP-PEPCK structure is represented by the brighter version (light blue, pale yellow and pink). This diagram was generated using CCP4MG.[26]

Figure 7: Pairwise sequence alignment of EcPEPCK and hcPEPCK
The three catalytically active mobile loops characteristic of PEPCK enzymes are highlighted in a pairwise sequence alignment. The R-loop sequence is highlighted in blue, the P-loop in yellow, and the Ω-loop in red. The alignment for EcPEPCK and hcPEPCK was generated using the EMBOSS needle global sequence alignment web tool from EMBL-EBI. [27,28]

Although both the ATP and GTP enzymes share the same global PEPCK core structure (RMSD of 3.31Å for human cytosolic and *E.coli* PEPCK, TM-score of 0.85)[29], their quaternary state can differ wildly by organism. Generally, GTP-dependent PEPCKs are monomeric, while the ATP-dependent enzymes have a greater variability in their oligomeric state.[21] *E. coli* PEPCK, one of the more characterized ATP-dependent PEPCKs, is monomeric, while PEPCKs isolated from plants are typically tetrameric. The highest level of oligomerization seen in PEPCK enzymes is in *Urochloa panicoides* (fodder grass), in which PEPCK exists as a hexamer.[30]

Another lesser known class of enzymes, first characterized as PEP-carboxytrans-phosphorylases in the 1970's, has recently been reclassified as inorganic pyrophosphate (PPi)-dependent PEPCK (EC 4.1.1.38), resulting in a third family of PEPCK enzymes. PP_i-PEPCK enzymes are more narrowly distributed than their nucleotide-dependent counterparts, and are absent in archaea.[8,23] As their name suggests, these related enzymes use pyrophosphate as their phosphoryl donor rather than a nucleoside triphosphate. PP_i-PEPCK enzymes are twice as large as their nucleotide-dependent cousins, with roughly 1400 amino acid residues. At first glance they do not appear structurally homologous to other PEPCK enzymes, but upon closer observation they share the same core N- and C-terminal domains that make up the typical PEPCK protein structure (Figure 8). In contrast to the majority of nucleotide-dependent PEPCKs, kinetic analysis of the PP_i-PEPCK enzyme from *P. freudenreichii* reveals a preference for the direction of carbon fixation (OAA synthesis).[31] From an evolutionary standpoint, these enzymes are very intruiging, since they share the same core PEPCK structure, yet are very distant from both ATP- and GTP-dependent PEPCKs. Previous studies have looked into the evolutionary relationship between the three classes of PEPCK enzymes, suggesting that the two nucleotide-dependent enzymes share a common ancestor, while the pyrophosphate-dependent enzyme's origins are questionable, due to the sequence identity between PP_i-PEPCK and NTP-PEPCKs being lower than 10%.[23,31] The most recent research suggests that the PP_i-PEPCK enzymes also share a common ancestor, due to the structural similarity of the active site, including key amino acid residues involved in substrate and metal binding.[8,31]

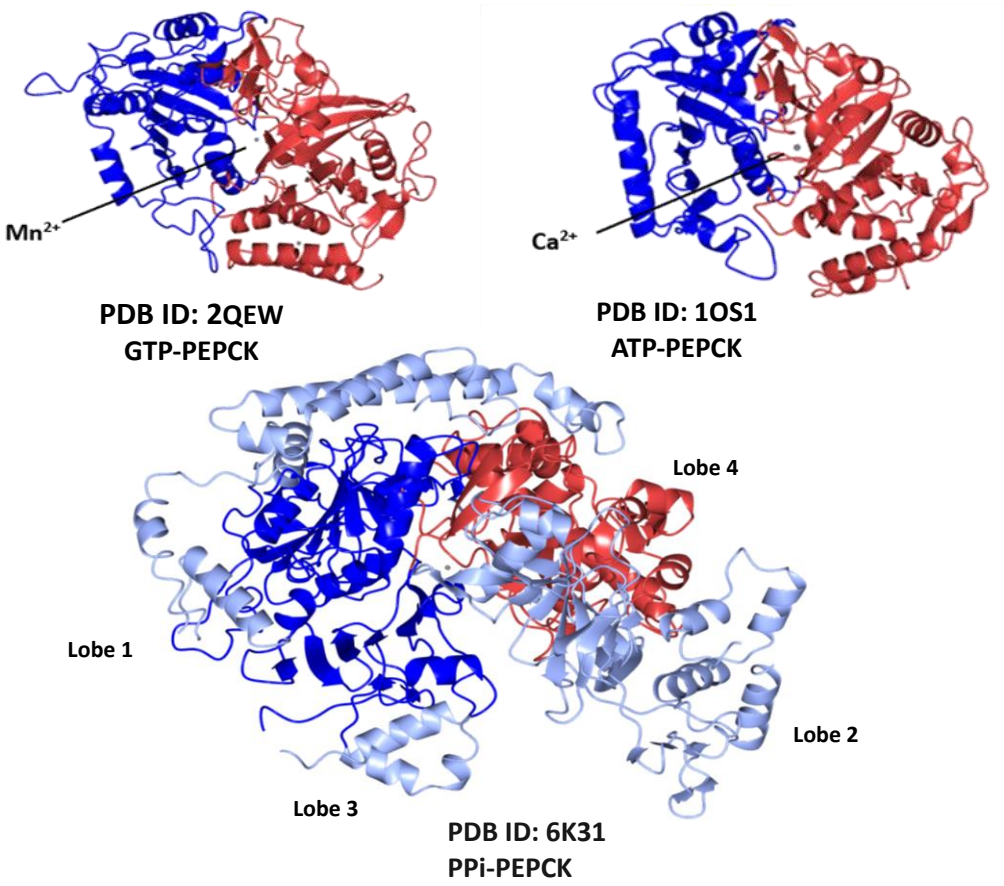


Figure 8: Structural comparison of the three classes of PEPCK enzymes
Three ribbon structures are depicted to demonstrate the structural similarities between ATP, GTP and PP_i-dependent PEPCK enzymes. The proteins are colour-coded blue and red to represent the N-terminal and C-terminal domains, respectively. This diagram was generated using CCP4MG.[26]

Aside from the core PEPCK domain structure, the PP_i-dependent enzymes contain four external 'lobe' domains, of which their function is still not fully understood. Two of the lobes, lobe 2 and lobe 3 (the largest and smallest lobes, respectively) are thought to make up the dimer interface, upon binding small metabolites including OAA, malate and fumarate.[31] Inhibition of PP_i-PEPCK activity upon dimerization is thought to be a form of enzyme regulation, as unlike either ATP or GTP-PEPCKs, PP_i-PEPCK appears to migrate between the two quaternary states, going from the active monomer to an inactive dimer. Although the nucleotide-dependent enzymes can have

higher quaternary states, the use of dimerization specifically as a potential mechanism of regulation is thought to be unique to the PP_i-dependent class of PEPCK enzymes.

1.4 Auxiliary roles for PEPCK in metabolism

PEPCK, as a metabolic enzyme is largely known for its role in glucose metabolism, converting OAA into PEP and ultimately upregulating the level of glucose in the cell. Since PEPCK controls this rate-limiting step of gluconeogenesis, gene expression of this enzyme is tightly regulated by various hormones to ensure homeostasis of these interconnected metabolic pathways. PEPCK is regulated at the gene level by stimulation via cAMP and glucagon, while being inhibited by glucose and insulin (Figure 9).[32–34]

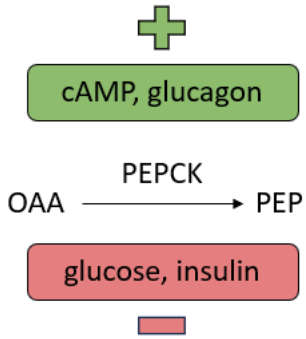


Figure 9: Regulation of PEPCK gene expression by hormones.
Through regulation of PEPCK gene expression, gluconeogenic activity is stimulated by cAMP and glucagon (as shown in green), and inhibited by glucose and insulin (as shown in red).

The PEPCK-catalyzed reaction of OAA → PEP sits at the junction of the citric acid cycle and gluconeogenesis, and is considered a rate-limiting step of the latter pathway.[9] The citric acid cycle is the central process of energy metabolism; however, intermediates of the cycle also function as substrates in other interconnected biosynthetic pathways, including amino acid synthesis, glyceroneogenesis and fatty acid synthesis.[35] These cataplerotic processes are carefully balanced with anaplerotic processes that replenish these intermediates. An example of such

anaplerosis is the recycling of amino acid skeletons into the citric acid cycle.[36] Catabolism of amino acids produces various gluconeogenic precursors that can be disposed of via the citric acid cycle. To prevent accumulation of these anions in the mitochondria, cataplerotic enzymes are required to remove these excess intermediates from the cycle. PEPCK, for example, can remove these anions in the form of OAA. Lastly, cataplerotic enzymes can ensure the citric acid cycle does not run unchecked, needlessly wasting energy in the overproduction of ATP and the reducing agent NADH. As PEPCK catalyzes the production of PEP, it also acts as a feeder reaction for downstream metabolic processes. In addition to gluconeogenesis, PEPCK plays a role in glyceroneogenesis, amino acid synthesis, and the aforementioned recycling of recycling of carbon skeletons.[35]

In contrast to PEPCK catalyzing the conversion of OAA to PEP as the first step of gluconeogenesis, pyruvate kinase catalyzes the last step of glycolysis, converting PEP to pyruvate.[9] During gluconeogenesis, pyruvate kinase is not stimulated by the various sugar metabolites found in the glycolytic pathway, and is instead inhibited by MgATP to prevent PEP from being converted back into pyruvate.[37] Interestingly, a study by Zamboni et al. showed that when pyruvate kinase is knocked out in *B. subtilis* bacteria, ATP-dependent PEPCK is able to work in the reverse of its normal gluconeogenic function, converting PEP to OAA.[38] This is not completely surprising, as some microorganisms are known to use their PEPCK enzymes to convert PEP to OAA for carbon fixation, which is especially useful in an anaerobic environment.[39–41] Beyond their canonical activity, there is even evidence that suggests some ATP-dependent PEPCK enzymes, such as those from *Saccharomyces cerevisiae*, have the ability to mimic pyruvate kinase activity and catalyze the conversion of PEP to pyruvate while generating ATP from ADP.[42]

As mentioned earlier, PEPCK expression is inhibited by the hormone insulin. Normally secreted in response to high blood glucose levels, insulin represses further gluconeogenic activity by PEPCK. In patients who are insulin-deficient, PEPCK is overexpressed. In fact, overexpression of the PEPCK enzyme in mutant mice has been proven to result in symptoms of type II diabetes, including the characteristic hyperglycemia.[43]

Overexpression of PEPCK is also seen in cancer patients. Cancer cells are able to hijack metabolic processes in order to obtain the nutrients needed to support tumour cell proliferation. It can also be noted that cancer cells favour anaerobic glucose metabolism over oxidative metabolism.[44] Therefore, it is not surprising that gluconeogenic enzymes like PEPCK are overexpressed or upregulated to increase citric acid cycle flux. By increasing the rate of both gluconeogenesis and the citric acid cycle, cancer cells are able to utilize alternative metabolic routes to fuel cell growth, even when glucose is unavailable.[45] As is characteristic of cancerous tumours, metabolic pathways are upregulated to provide nutrients for cell proliferation and metastasis. It has been established that PEPCK has a role in gluconeogenesis and glyceroneogenesis via its canonical carboxy-lyase activity (belonging to EC 4.1.1), however recent studies have suggested that human cytosolic PEPCK moonlights as a protein kinase in the SREBP pathway leading to increased lipogenesis in human hepatocellular carcinoma (HCC) cells.[46,47] It has been hypothesized that PEPCK phosphorylates the insulin-induced 1 protein (INSIG) as part of a kinase cascade.[46] This moonlighting activity has not been fully characterized in PEPCK as of yet, and it remains to be seen if this putative role for the human enzyme is unique or is conserved in other GTP-dependent PEPCKs.

1.5 Inhibition of PEPCK enzymes

Insight has been given into PEPCK regulation on several levels, from regulation of PEPCK gene expression by various hormones, to phosphorylation and other post-translational modifications and even regulation at the protein level, via small molecules, anions and other inhibitors. The active site of PEPCK enzymes includes two binding pockets, one for OAA/PEP binding, and the other for the nucleotide.[24,48] Therefore, it is not surprising that analogs of either the substrate or the nucleotide have been characterized to bind with relatively high affinity. Substrate analogs, including dicarboxylates modified from those found in the TCA cycle, phosphoryl monocarboxylates and sulfonyl monocarboxylates were analyzed for inhibitory action.[49] Some of the analogs that yielded the greatest inhibition include oxalate, phosphonoformate, phosphoglycolate, sulfopyruvate and sulfoacetate.[49] From a structural standpoint, all of these inhibitors are analogs of either the OAA substrate, the PEP product, or in the case of oxalate, an analog of the enolate intermediate. The OAA/PEP binding pocket contains two distinct subsites adjacent to each other, the OAA subsite and the PEP subsite, with the OAA subsite located in closer proximity to the M1 metal. Structural analogs of OAA and PEP bind to one of these subsites depending on their chemical structure.

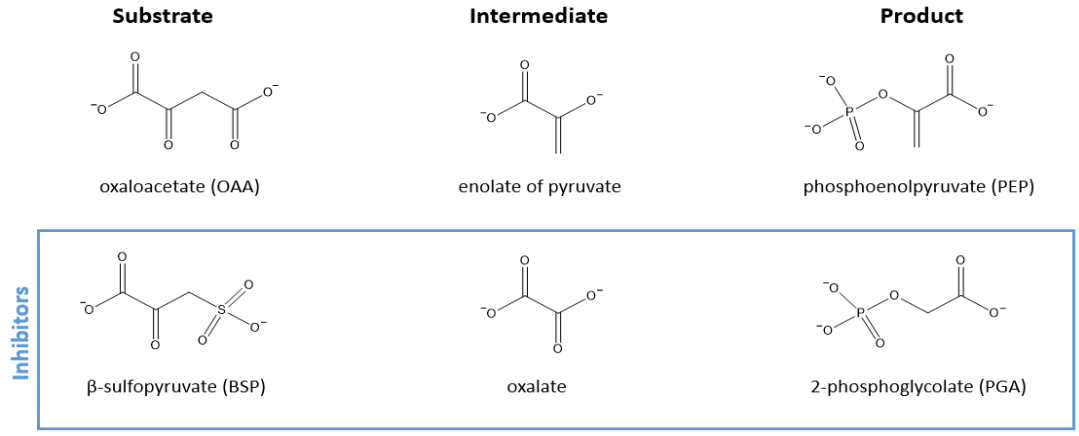


Figure 10: Substrates of the PEPCK-catalyzed reaction and some of their analogous inhibitors.[50] Inhibitors are highlighted in blue and listed under their respective analogs. Compounds were generated using ChemDraw from the ChemOffice+ suite.[51,52]

A recent study characterized a novel molecule to bridge these two subsites together, resulting in an inhibitor with increased potency against PEPCK. This inhibitor, 3-[(carboxymethyl)thio]picolinic acid, or CMP for short, uses 3-mercaptopicolinic acid (MPA – historically known as an inhibitor of PEPCK[53–55]) as a chemical scaffold, together with an extended carboxy tail to demonstrate occupation of both subsites simultaneously.[48] By occupying both of these subsites, CMP shows the possibility of creating novel inhibitors with increased selectivity for PEPCK.

1.6 Summary and Objectives

In summary, aside from the structural information in the PDB databank, the majority of information we have on PEPCK has been done on the GTP-dependent enzymes, with emphasis on the rat and human cytosolic enzymes. The two nucleotide-dependent enzymes have been studied independently of each other, and no comparative studies have been done between the two classes using the same experimental techniques. The following chapters of the thesis address the similarities and differences between ATP and GTP-dependent PEPCK enzymes, using *E. coli* and human cytosolic enzymes respectively. A combined kinetic and structural approach was used to elucidate key information on these enzymes, and prospective new roles for both the human and *E. coli* enzymes are addressed.

Chapter 2: hcPEPCK moonlights as a protein kinase potentially leading to increased lipogenesis via the SREBP pathway

2.1 Introduction – INSIG, lipogenesis and PEPCK

PEPCK enzymes are well known for their role in gluconeogenesis and glyceroneogenesis as metabolic kinases, or carboxy-lyases/decarboxylases. These enzymes have been thoroughly studied for their role in the conversion of OAA to PEP. However, GTP-PEPCKs may exhibit protein moonlighting, as described in a recent study.[46] It has been suggested that human cytosolic PEPCK (hcPEPCK) has an alternative role as a protein kinase in a signal cascade that results in increased lipogenesis.[46] Upon phosphorylation by an AKT1 kinase, PEPCK activity is thought to switch from its typical gluconeogenic pathway and instead phosphorylate the INSIG protein.[46] The INSIG protein, or insulin-induced gene 1 protein, while expressed in all tissues, is overexpressed in the liver and is involved in the regulation of cholesterol homeostasis.[56] Aptly named, the INSIG protein is regulated by insulin. INSIG is a transmembrane protein that is localized to the endoplasmic reticulum, that regulates sterol biosynthesis via a negative feedback loop.[56]

Sterol biosynthesis is regulated by the SREBP pathway, wherein the genes that encode for biosynthetic enzymes are upregulated via transcription factors known as sterol regulatory element-binding proteins (SREBPs).[57] This pathway has a unique feature – termed regulatory intramembrane proteolysis (RIP) – that allows for the SREBPs to become activated in times where cholesterol concentrations are low.[58] Inactivated SREBP proteins are found in the ER, where they exist as a complex with another transmembrane protein, SCAP (SREBP cleavage-activating protein). SCAP proteins are able to escort SREBPs to the Golgi apparatus via COPII vesicles, where proteases S1P and S2P cleave the functional N-terminal domain off the SREBP

transmembrane protein.[59,60] Then the activated SREBP can be translocated into the nucleus where it can act as a transcription factor, upregulating biosynthetic genes and leading to lipogenesis.[61,62]

Cholesterol homeostasis is a delicate balance, since cholesterol (and other sterols) are needed for everything from structural components of a cell to several different hormones. However, when cholesterol is too abundant, it can aggregate and clog arteries, causing atherosclerosis and increased risk of myocardial infarction. Therefore, when cholesterol biosynthesis is regulated via a negative feedback loop – the presence of cholesterol itself represses biosynthetic gene expression.[58] Both SCAP and INSIG proteins are capable of binding sterols like cholesterol, and upon binding cholesterol, they undergo a conformational change resulting in the formation of the SCAP-INSIG complex (Figure 11).[60,63,64] Complexed with INSIG, SCAP proteins can no longer escort SREBPs to the Golgi for activation, therefore repressing further biosynthetic gene expression and in turn generating unneeded cholesterol when sterol levels are already plentiful.

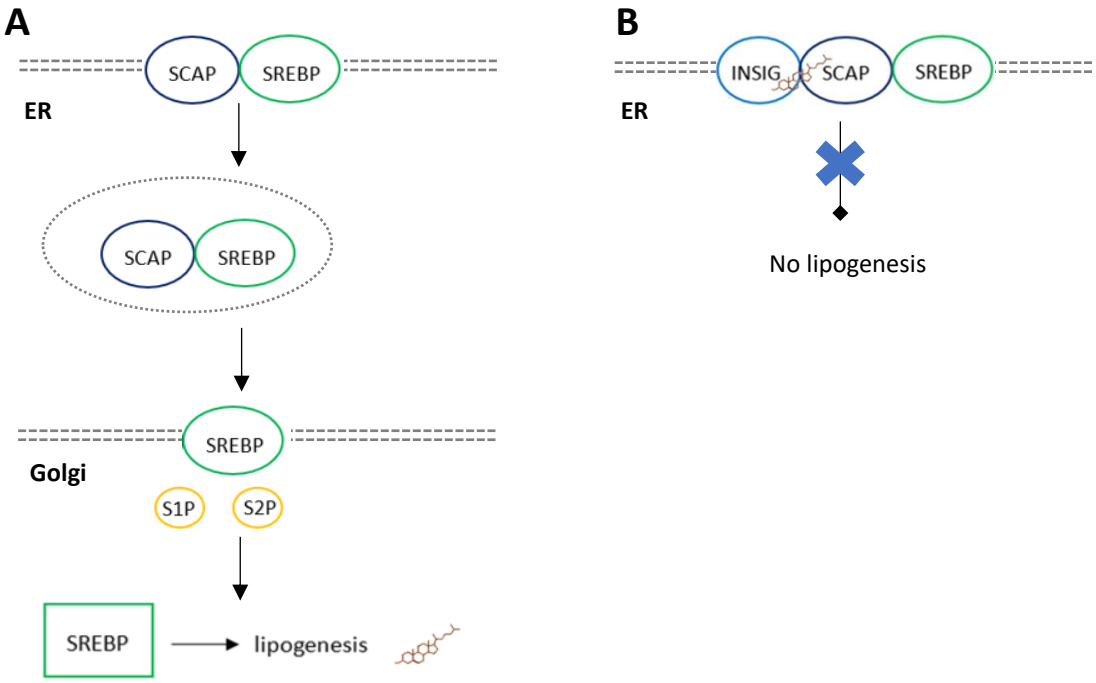


Figure 11: Negative feedback loop of SREBP pathway
A) In the absence of cholesterol, SCAP proteins escort SREBPs to the Golgi to be activated by S1P and S2P proteases. The active SREBPs translocate to the nucleus, where they upregulate cholesterol biosynthesis. B) In the presence of cholesterol, cholesterol binds to INSIG and SCAP proteins, forming a complex and halting the SREBP pathway via a negative feedback loop.

However, in certain conditions where PEPCK is overexpressed, it has been suggested that PEPCK acts to phosphorylate the INSIG protein, causing INSIG to be unable to bind to SCAP efficiently, allowing for the negative feedback loop of the SREBP pathway to be abolished and leading to uncontrolled lipogenesis (Figure 12).[46] Conditions like these lend itself to cancerous cell proliferation, and in fact PEPCK levels are often elevated in tumour cells, as expected.[45,62]

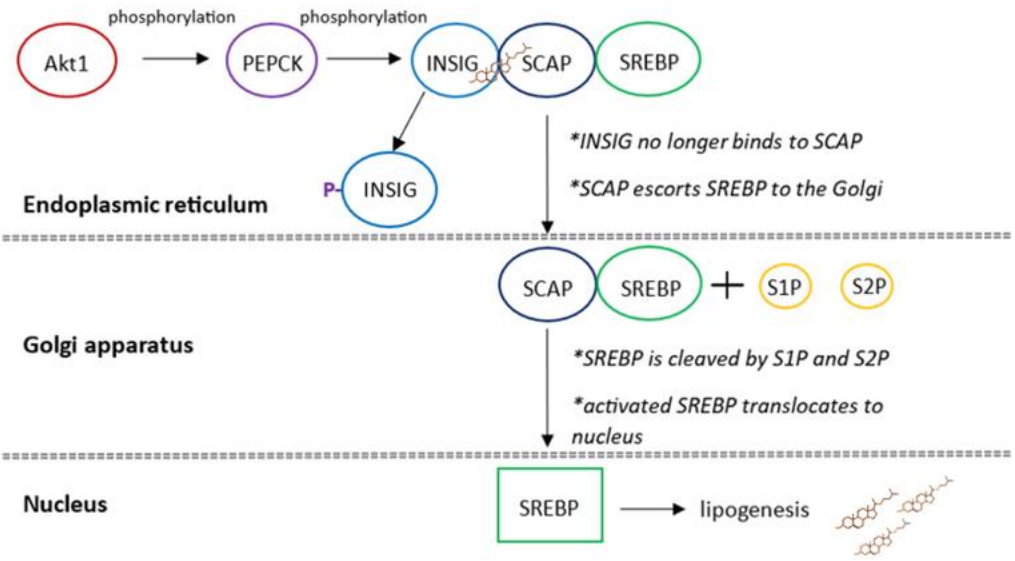


Figure 12: A simplified diagram of the proposed kinase cascade leading to increased lipogenesis
In HCC cells even where lipid levels are high, PEPCK is thought to phosphorylate INSIG1/2, resulting in the INSIG protein being unable to bind to the SCAP protein. The SREBP pathway is free to proceed as though in the absence of cholesterol, bypassing the negative feedback loop and leading to abnormally high levels of lipids in the cell.

Using co-immunoprecipitation analyses and mass spectrometry, it has been suggested that PEPCK is phosphorylated at Ser90 by AKT1, allowing it to be translocated to the ER.[46] Once there, it phosphorylates the INSIG protein (Ser207 and Ser151 in INSIG1 and INSIG2, respectively). There are two forms of the INSIG protein, INSIG1 and INSIG2, and these isozymes are nearly structurally identical.[59,65] The mobile loop peptide sequence is conserved in both isozymes: FDR***S***RSGF.[46] This sequence containing the phosphorylated serine (S151) is located on one of the few cytosolic loops of the INSIG protein (Figure 13).

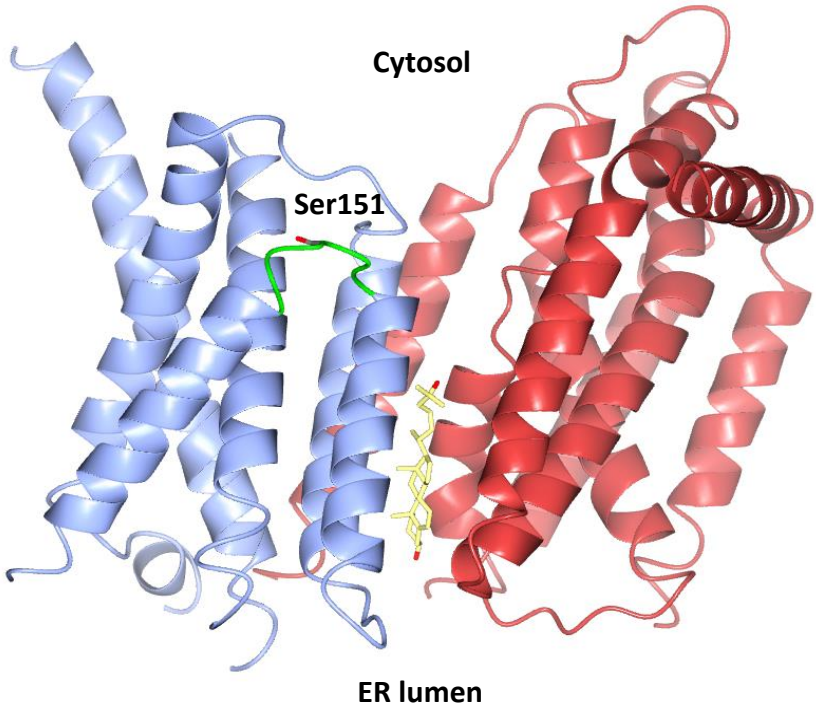


Figure 13: Ribbon diagram of the human INSIG2-SCAP complex with 25-hydroxyl cholesterol
The INSIG protein is coloured blue, with the mobile loop depicted in green. The SCAP protein is coloured red, and the sterol ligand, 25-hydroxyl cholesterol is coloured yellow. The ribbon structure was generated using CCP4MG, using the PDB file (PDB ID: 6M49) obtained from the PDB.[26,64]

Xu et al. states that phosphorylation of hcPEPCK at Ser90 results in decreased affinity for its oxaloacetate substrate, reducing its ability to catalyze the canonical PEPCK-catalyzed reaction in gluconeogenesis.[46] Ser90 is located on the catalytically active R-loop in the PEPCK active site, the mobile loop that coordinates OAA/PEP binding (Figure 14). It is also noted that phosphomimetic mutants of PEPCK – S90D or S90E mutants that resemble phosphorylated serine (Figure 15) – also show decreased affinity for the OAA substrate, "turning off" gluconeogenic activity and therefore the production of PEP.[46] Therefore, in this chapter, gluconeogenic activity of the wildtype and phosphomimetic S90E mutant of human cytosolic PEPCK will be analyzed to further our understanding of PEPCK's potential role as a protein kinase in SREBP-dependent lipogenesis.

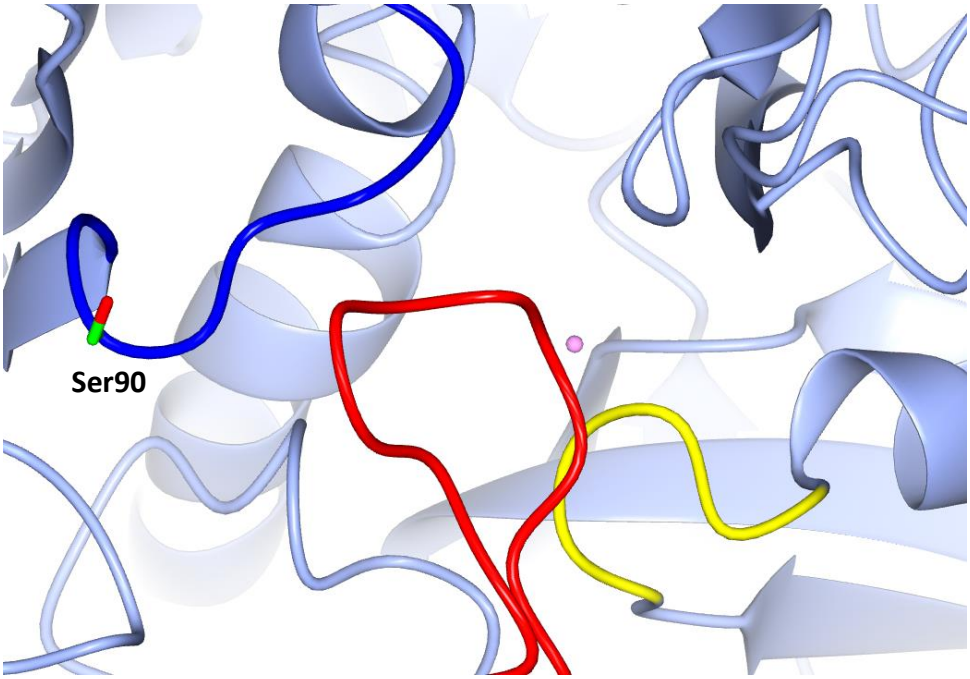


Figure 14: Location of Ser90 in the active site of hcPEPCK
Ser90 is located on the R-loop in the hcPEPCK active site. The R-loop is coloured blue, the P-loop is coloured yellow and the omega loop is coloured red. The catalytically active M1 metal, a manganese cation, is depicted as a pink sphere. This ribbon structure was generated using CCP4MG.[26]

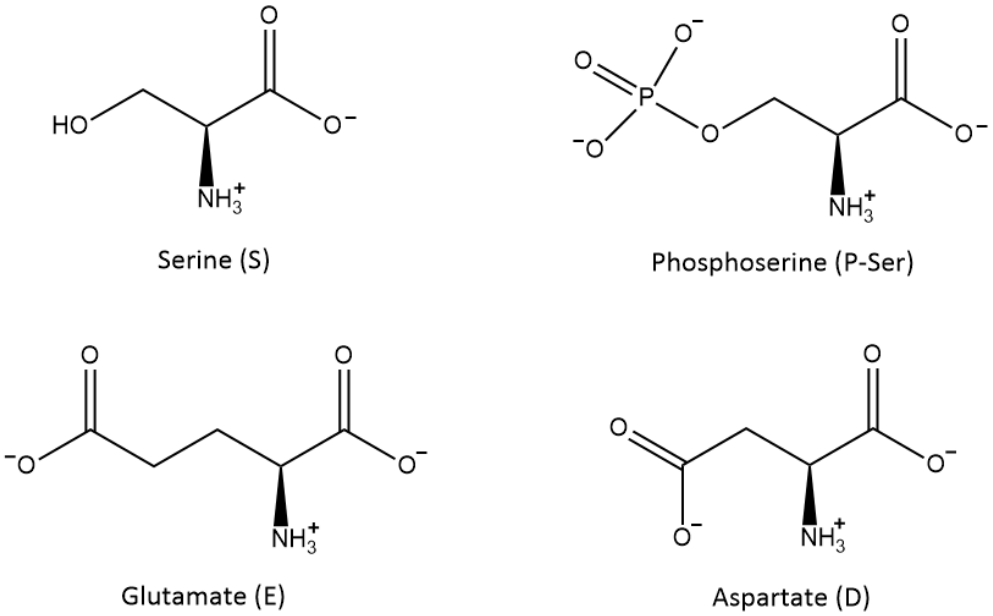


Figure 15: A structural comparison of serine, phosphoserine and their phosphomimetic substitutions at neutral pH
Glutamate and aspartate are structural analogs of phosphorylated serine due to their long negatively charged sidechains that mimic the negatively charged phosphate group.

2.2 Materials & Methods

2.2.1 Materials

Nickel-NTA resin was from UBP-Bio and P-6DG resin was from Bio-Rad. PEP potassium salt, NADH disodium salt (>96% purity), TCEP and DTT were from ChemImpex, GTP disodium salt (>90% purity) was from BioShop, GDP disodium salt (95% purity) was from CombiBlocks, oxalic acid (98% purity) was from Sigma Aldrich, and OAA (>97% purity) was from Millipore Sigma. INSIG-2 peptide [amino acid sequence FDR*S*RSGF] was chemically synthesized (95% purity) and obtained from BioBasics or Genscript. SYPRO Orange 5000X fluorescent dye was from Invitrogen (Thermo Fisher).

2.2.2 Enzymes

The coupling enzymes pyruvate kinase (PK), lactate dehydrogenase (LDH) and malate dehydrogenase (MDH) were purchased from Calzyme Laboratories. The plasmid expressing His6-SUMO protease was a gift from C. Lima (Sloan-Kettering Institute, New York, NY) and SUMO protease was expressed and purified as previously described.[24]

2.2.3 Cloning & transformation of WT and S90E hcPEPCK

The gene sequence of both wildtype and phosphomimetic (S90E) human cytosolic PEPCK was synthesized and cloned into a pE-SUMO-star (Kan) vector (LifeSensors PE-1101-0020) by GenScript. The resulting plasmid allows for the expression of human PEPCK with a N-terminal HIS_6-SUMO fusion with the tag cleavable by a SUMO protease, yielding authentic PEPCK

protein with no additional amino acids. The plasmid was transformed into BL21(DE3) *Escherichia coli* cells via the heat shock method. In brief, ~50 ng of plasmid DNA was added to a tube of 100 µl of chemically competent BL21(DE3) cells and placed on ice for 30 minutes. The mixture of DNA and cells was immersed in a 42°C water bath for 45 seconds before immediately being placed back on ice for 5 minutes. 1 mL of LB media was added to the transformed cells, and they were allowed to grow at 37°C for 1 hour before being plated onto LB-Kan media for growth overnight at 37°C. Individual colonies were chosen to create glycerol cell stocks of human cytosolic PEPCK (hcPEPCK). Pilot expression of these cell stocks was performed to validate protein over-expression and solubility.

2.2.4 hcPEPCK protein expression and purification

Expression of the protein was initiated by inoculating 200mL of LB media supplemented with 50 µg/mL kanamycin with BL21(DE3) cell stocks containing the HIS_6-SUMO-PEPCK construct. The flasks were incubated with shaking (170rpm) at 37°C overnight. The overnight culture was divided into four aliquots and subsequently used to inoculate four flasks of 950 mL ZYP-5052 autoinduction media (AI)[66] again supplemented with 50 µg/mL of kanamycin. These AI cultures were then incubated while shaking at 170rpm for 20-24 hours, at 20°C to induce expression of the recombinant HIS_6-SUMO-PEPCK proteins. After protein expression, the culture media was centrifuged at 5000 rcf for ten minutes to harvest the cells and the resulting cell pellets were stored at -80°C.

The purification protocol for both WT and S90E hcPEPCK constructs was identical to that of the rcPEPCK protein, and has been published previously.[67] Briefly, the cell pellets were resuspended

in 25 mM HEPES pH 7.5, 10% glycerol, 300 mM NaCl, 10 mM imidazole and 2 mM TCEP (Buffer 1) and lysed via two passes through a French Press cylinder at 1000 psi. The cell lysate was centrifuged at 15000 rcf for 45 minutes and the cell pellet was discarded. The soluble lysate was added to Ni-NTA resin that had been equilibrated with Buffer 1, and left to incubate at 4°C while stirring for one hour. After incubation, the resin lysate slurry was poured into a clean column. The recombinantly His-tagged hcPEPCK protein bound to the Ni-NTA resin, and the rest of the cellular proteins in the lysate were washed off the column using Buffer 1 until the absorbance (A_{280}) of the eluent reached zero. The PEPCK protein was then eluted in 10 mL fractions from the Ni-NTA column using 25 mM HEPES pH 7.5, 300 mM imidazole and 2 mM TCEP (Buffer 2). Each of the fractions had their absorbance measured, and were pooled together if the measured A_{280} was > 0.2. The pooled fractions were concentrated to under 10 mL via a stirred cell concentrator pressurized with N_2 gas, using a filter with a pore size cut-off of 30 kDa. A P-6DG size exclusion column was setup with 1 cm Chelex 100 resin at the top to remove metals present in the sample. The concentrated protein was added to the equilibrated P-6DG column for buffer exchange into 25 mM HEPES pH 7.5, 2 mM TCEP (Buffer 3). Fractions were collected and monitored for protein presence via A_{280} > 0.2. The HIS_6-SUMO tag was then cleaved using SUMO protease. Incubation with the SUMO protease took place overnight at 4°C with gentle magnetic stirring. SUMO protease and cleaved HIS_6-SUMO tag were removed from the sample via another pass through the Ni-NTA column.

The purified hcPEPCK was concentrated to ~20 mg/mL via filter centrifugation, and the final protein concentration was determined using a NanoDrop spectrophotometer (A_{280} of 0.1%: 1.5 mg/mL^{-1}). The concentrated protein was then flash-frozen into 30 µl aliquots and stored at -80°C.

2.2.5 hcPEPCK crystallization

Wildtype and S90E hcPEPCK GTP + oxalate and GTP-only crystals were grown using the hanging drop vapour diffusion method at ambient temperature, using a ratio of 4 µl protein solution to 2 µl mother liquor. The reservoir solution consisted of 18-28% PEG 3350, 100 mM HEPES (pH 7.5), and 10-40 mM $MnCl_2$ as an additive. Water was added to each of the wells for a final volume of 700 µl. 2 µl of the reservoir solution was pipetted onto 4 µl of protein solution (2-8 mg/mL hcPEPCK, 25 mM HEPES pH 7.5, 10 mM DTT, 1 mM GTP and 1 mM oxalate) on a siliconized cover slide. Crystals with an imperfect cubic morphology grew after 24 hrs, and continued to grow in size for up to one week. The best crystals grew in conditions consisting of 26% PEG 3350 and 20 mM $MnCl_2$. Crystals were harvested and soaked for 10 minutes in cryoprotectant consisting of 30% PEG 3350, 10% PEG 400, 100 mM HEPES pH 7.5, 20 mM $MnCl_2$, 10 mM DTT, 1 mM GTP and 1 mM oxalate, before being cryo-cooled by immersion in liquid nitrogen.

The INSIG 8-mer peptide used in both crystallization and kinetics was prepared as a 100 mM solution dissolved in 1M HEPES buffer at pH 7.5, using its molecular weight of 970.5 g/mol. The phosphomimetic S90E hcPEPCK + INSIG crystal tray was initially prepared using the same reservoir solutions as the GTP + oxalate co-crystal structures, with the protein solution containing 0.5-2 mM INSIG peptide instead of GTP and oxalate. Crystals were harvested and soaked in cryoprotectant as mentioned above (containing 1 mM INSIG) before being cryo-cooled in liquid nitrogen.

Additional crystallographic conditions were robotically screened using two protein solutions (either 5 or 10 mg/mL S90E hcPEPCK with 25 mM HEPES pH 7.5, 10 mM DTT, 20 mM $MnCl_2$ and 1 mM INSIG peptide) against the PACT, BCS, JCSQ+, MCSG-1 and MCSG-4 crystal

screens from Molecular Dimensions on a 96 well tray (0.3 µl protein solution to 0.3 µl reservoir solution). Four of the best results from the screens were scaled up to the standard 24 well tray using a protein solution of 5 mg/mL S90E hcPEPCK, 25 mM HEPES pH 7.5, 10 mM DTT, 20 mM $MnCl_2$ and 1 mM INSIG peptide. The reservoirs contained 700 µl of the screened crystal conditions, and 2 µl of the reservoir solution was added to 4 µl of the protein solution on a siliconized cover slide for a total drop volume of 6 µl.

The initial PEG 3350 and 100 mM HEPES pH 7.5 crystal condition was revisited using 2.8 mg/mL S90E hcPEPCK and 500 µM INSIG peptide based on the stoichiometric ratio determined from the DSF assay (**see section 2.2.7**).

Data for the WT and S90E hcPEPCK GTP + oxalate complexes, maintained at cryogenic temperatures, were collected at our home source (Rigaku/Cuκα rotating copper anode) diffractometer at the University of Waterloo (Waterloo, ON), and indexed and scaled using HKL-2000.[68] Data for the WT GTP-only complex was collected on the CMCF-ID beamline at the Canadian Light Source (Saskatoon, SK), and indexed and scaled using the DIALS software[69,70] followed by the Aimless program as part of the CCP4 package. All phases were determined by the molecular replacement method using MOLREP in the CCP4 package and a wildtype hcPEPCK structure from the PDB databank (PDB ID: 1KHG).[71,72] Structures were refined with multiple iterations of RefMac5 followed by manual model adjustment and addition of the GTP and oxalate ligands using Coot.[73] Model validation was performed using the Molprobity web server (http://molprobity.biochem.duke.edu). A summary of the crystallographic data statistics is presented in Table E1 (**see Appendix E**).

2.2.6 hcPEPCK kinetic activity and inhibition assays (kinetic replots)

The 1 mL assay mix in the OAA → PEP direction contained 100mM HEPES buffer, pH 7.5, 350µM OAA, 500µM GTP, 1mM ADP, 4mM $MgCl_2$, 100µM $MnCl_2$, 300µM NADH, and 10mM DTT. Coupling enzymes were added, including 10U of PK and 30U of LDH, followed by 2.5µg of PEPCK, then OAA to start the reaction. For the PEP → OAA direction, HEPES, $MgCl_2$, $MnCl_2$, DTT, and NADH concentrations remained the same, omitting the ADP nucleotide while containing 4mM PEP, 1mM GDP, and 10U of MDH, as well as CO_2 in the form of 50mM $KHCO_3$ bubbled with dry ice. The assay in this direction was initiated by adding 2.5µg of PEPCK enzyme. The oxalate inhibition assays for WT and S90E hcPEPCK were performed in the PEP → OAA direction, using the standard assay conditions and increasing concentrations of oxalate from 0 to 400 µM. All assays were performed at 25°C, and PEPCK activity was monitored indirectly via the depletion of NADH at 340 nm for 10 minutes.

The kinetic data from the activity assays were fit non-linearly to the standard Michaelis-Menten equation using SigmaPlot11 to determine kinetic parameters in both directions of catalysis. (**Equation 1**). Apparent V_{max} and K_M parameters were generated for each inhibitor concentration and were re-plotted in a linear relationship against inhibitor concentration to calculate inhibition constants. A competitive K_I for oxalate was determined in both directions of catalysis by plotting K_M/V_{max} vs [oxalate] (**Equation 2**). Similarly, any uncompetitive effects were determined by plotting $1/V_{max}$ vs [oxalate] (**Equation 3**).

Equation 1: $$v = \frac{V_{max} \times [S]}{K_M + [S]}$$

Equation 2: $$\frac{K_{Mapp}}{V_{max\,app}} = \frac{K_M}{V_{max} \times K_I} \times [I] + \frac{K_M}{V_{max}}$$

Equation 3: $$\frac{1}{V_{max\,app}} = \frac{1}{V_{max} \times K_I} \times [I] + \frac{1}{V_{max}}$$

Inhibition of the S90E hcPEPCK enzyme by the INSIG peptide was also initially attempted using the standard spectrophotometric PEPCK assay mentioned above. The INSIG 8-mer peptide (FDR***S***RSGF, molecular weight of 971.03 g/mol) was solubilized to create a 100 mM stock INSIG solution stored in 1M HEPES pH 7.5. S90E hcPEPCK activity in the reverse direction (against PEP substrate) was monitored at 340 nm under increasing concentrations of INSIG peptide up to 300 µM. Inhibition of S90E hcPEPCK by the INSIG peptide was analyzed by generating an [incomplete] IC50. An estimated competitive inhibition constant was calculated using the following equation:[74,75]

Equation 4:

$$K_i = \frac{IC50}{1 + \frac{S}{K_M}}$$

The full set of Michaelis-Menten curves needed for the kinetic replots to calculate a more accurate inhibition constant could not be obtained due to limited resources and solubility of the INSIG peptide in millimolar concentrations.

2.2.7 Differential scanning fluorimetry assay for INSIG binding

Differential scanning fluorimetry (DSF) was used as a thermal shift assay for INSIG peptide binding to the phosphomimetic S90E hcPEPCK mutant.

The DSF assay was performed using a QuantStudioPro 6 thermocycler (Figure 16). The thermocycler program was set up for data collection starting at 20°C followed by a 0.05°C increase per second, reaching a top temperature of 90°C for 5 seconds. Afterwards, the thermocycler was left to cool back down to 20°C at increments of 3°C per second.

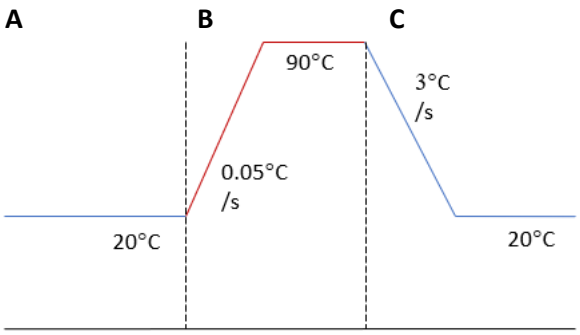


Figure 16: Thermocycler temperature gradient setup for the QuantStudioPro 6.
The thermocycler was allowed to equilibrate at 20°C (A), followed by data collection every 3s for 24 minutes (B) as the temperature incrementally heats up to 90°C. Data collection is finished, and the thermocycler was allowed to cool back down to 20°C (C). Run duration was approximately 36 minutes.

An optimal ratio of 40 µM protein to 20X SYPRO Orange dye was determined previously in our lab. The thermal shift assay was done under the following conditions: 25 mM HEPES pH 7.5, 5 mM $MnCl_2$, 2 mM TCEP and 20X SYPRO Orange in a total reaction volume of 20 µl. Ligand binding was studied by adding INSIG peptide to the assay (0 to 15 mM INSIG), using the preliminary data from the spectrophotometric assays as a starting point. The DSF assay was completed in duplicate for each concentration of INSIG peptide including controls.

Quant Studio Design & Analysis Software 2.6.0 was used to collect raw data and determine the derivative of fluorescence intensity with respect to temperature. The derivatized data was fit to the Boltzmann model using this software and Microsoft Excel and subsequently used to determine the melting point (T_m) of S90E hcPEPCK at each concentration of peptide. Thermal shifts were calculated by subtracting the T_m of the unbound PEPCK protein from the T_m of the INSIG-bound protein. The obtained thermal shift data was then fit to a ligand binding isotherm using SigmaPlot11 to determine a binding constant.

2.3 Results & Discussion

2.3.1 WT and S90E hcPEPCK crystallization

WT and S90E hcPEPCK crystals grew in the presence of the INSIG peptide (Figure 17) using the same crystal conditions, as seen in the 500 μM to 2 mM INSIG crystal screens, however these crystals diffracted poorly (~3Å) and the solved structures were revealed to be in their holoenzyme state, with only an Mn^{2+} cation in the active site and no other ligand bound. Attempting to co-crystallize with a higher concentration of INSIG resulted in heavy precipitate instead of a crystalline product.

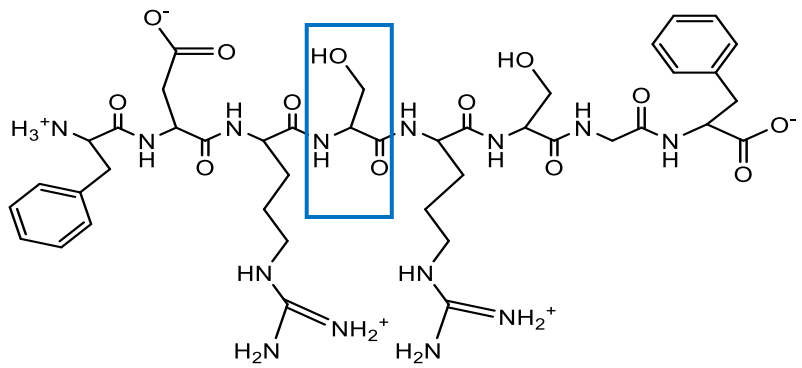


Figure 17: Structure of the INSIG 8-mer peptide (FDRSRSGF) at neutral pH
Sequence of the INSIG-2 mobile loop thought to be phosphorylated (at Ser151 – outlined above) by PEPCK. This diagram was created using ChemDraw.

The wildtype hcPEPCK GTP structure (Figure 18) crystallized over a few days in a cubic morphology. The crystal structure was obtained and refined to 1.7Å. The Ω-loop could not be fully modelled due to intrinsic disorder associated with the WT hcPEPCK GTP structure adopting an open conformation.

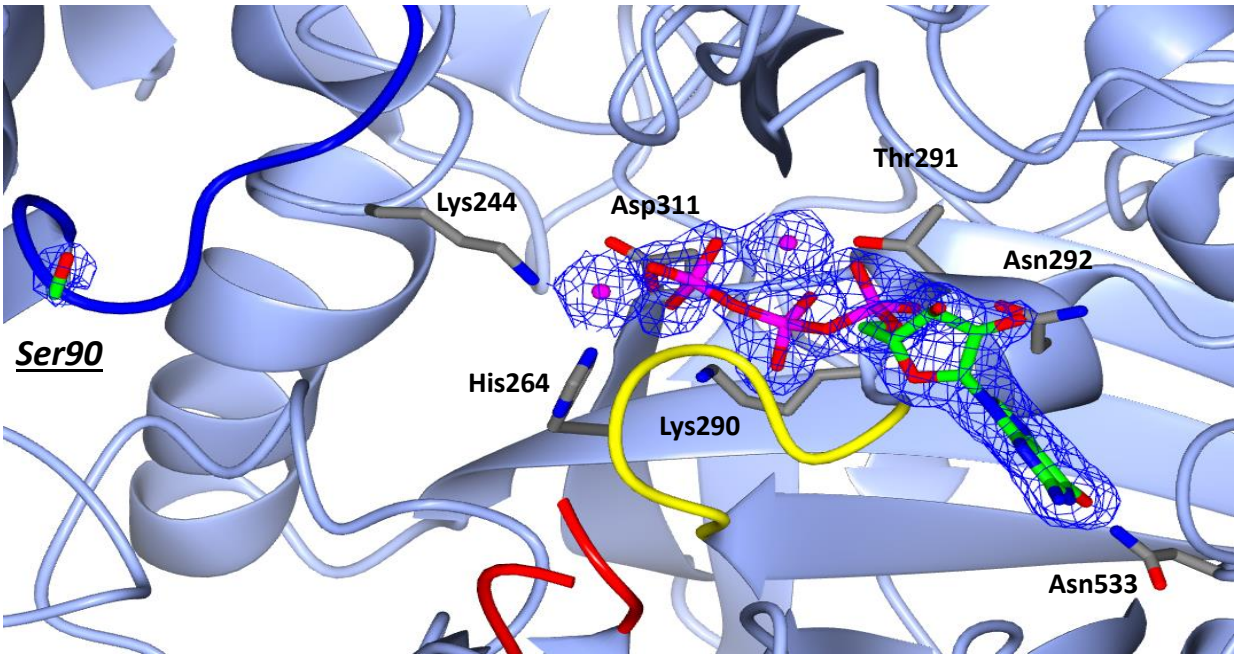


Figure 18: Ribbon structure of the WT hcPEPCK active site with GTP nucleotide
R-loop and P-loops are coloured blue and yellow, respectively, with a portion of the disordered Ω-loop coloured red. Ser90 is shown on the R-loop. Manganese cations are modelled in as pink spheres, representing the M1 and M2 metals. 2Fo-Fc electron density maps are contoured to 1.0 sigma for the GTP nucleotide and 2.5 sigma for the metal ions. The structure was generated using CCP4MG.[26]

The S90E phosphomimetic mutant crystallized in the presence of GTP and $MnCl_2$, but solving the structure yielded the holoenzyme with only Mn^{2+} bound to the active site. It is likely that the S90E mutation has altered the dynamics of the active site, making it significantly more difficult to crystallize the closed enzyme conformation. Superposition of the wildtype and S90E hcPEPCK GTP and oxalate structures (0.24 RMSD) reveals no significant conformational changes in the global fold nor the R-loop where the mutation resides (Figure 19). Focusing on the active sites reveals the only significant difference for the phosphomimetic mutant in that the omega loop lid could not be fully modelled due to the increased intrinsic disorder.

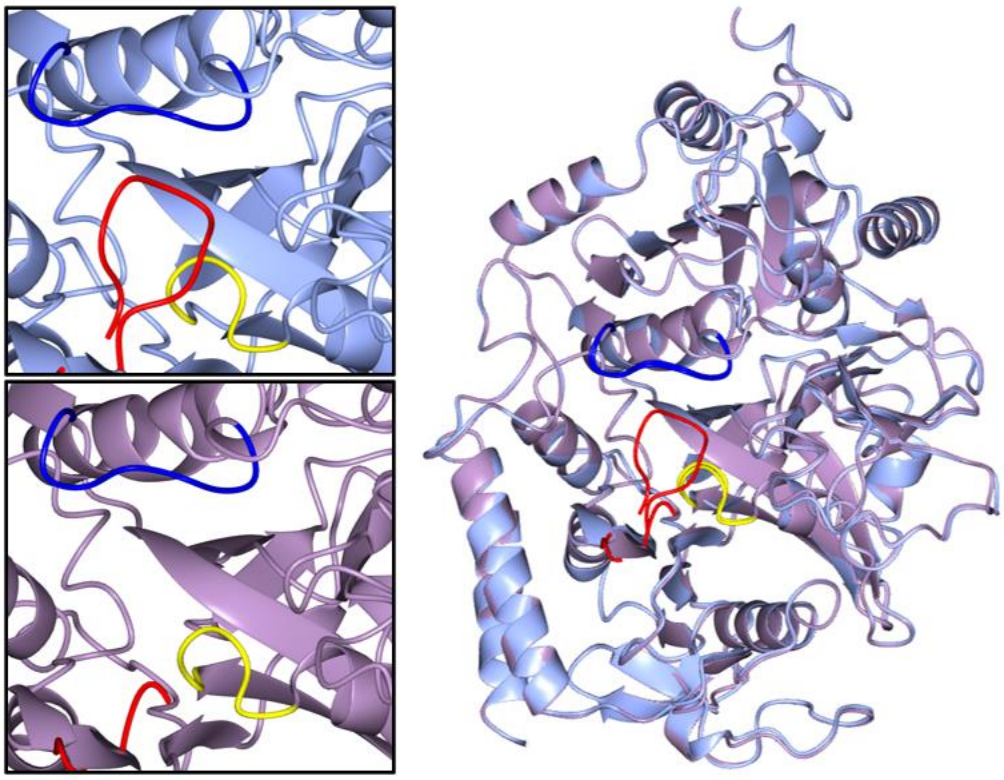


Figure 19: Superposition of the wildtype and S90E structures
The WT hcPEPCK ribbon structure is coloured light blue, and the S90E mutant structure is coloured lilac. The R-loop is coloured dark blue, the P-loop yellow, and the N- and C- terminal ends of the Ω-loop are coloured red. A close-up view of the wildtype (top left) and S90E (bottom left) active sites are provided. This figure was generated using CCP4MG.[26]

Adding oxalate in addition to GTP nucleotide greatly enhanced crystallization of both the wildtype and S90E enzymes. The presence of the oxalate inhibitor stabilized the closed conformation of the enzyme, allowing for both the WT and S90E co-crystal structures with GTP nucleotide and oxalate inhibitor bound to be obtained. Due to oxalate being an analog of the enolate intermediate, the oxalate GTP complex provides an accurate stereo-electronic mimic of the intermediate state of the PEPCK enzyme associated with the catalyzed reaction in either direction.

Structurally, there is no significant difference between the wildtype and S90E hcPEPCK co-crystal structures with oxalate and GTP nucleotide, aside from the difficulty in crystallizing them. Their common global protein structure is unsurprisingly identical, and the active site remains largely similar between them. Minor differences can be found in the integrity of the enzyme's closed conformation, with the wildtype found in the closed conformation with oxalate and GTP bound (Figure 20), and the phosphomimetic occupying a partially closed conformation with the same ligands (Figure 21). Both structures include two protein chains, but while the chains are nearly identical in the wildtype structure, the phosphomimetic has one chain with a fully modelled omega loop lid, and the other with a partially modelled lid as shown in Figure 21. The two structures differ slightly in crystallographic quality, however. The wildtype co-crystal structure (Figure 20) has a slightly better resolution at 2.1Å, while the phosphomimetic mutant (Figure 21) dataset was refined to 2.45Å. The wildtype dataset has near perfect completeness, while the mutant is slightly lower. Accordingly, the R-values for the wildtype dataset are lower than that of the mutant (Table 1).

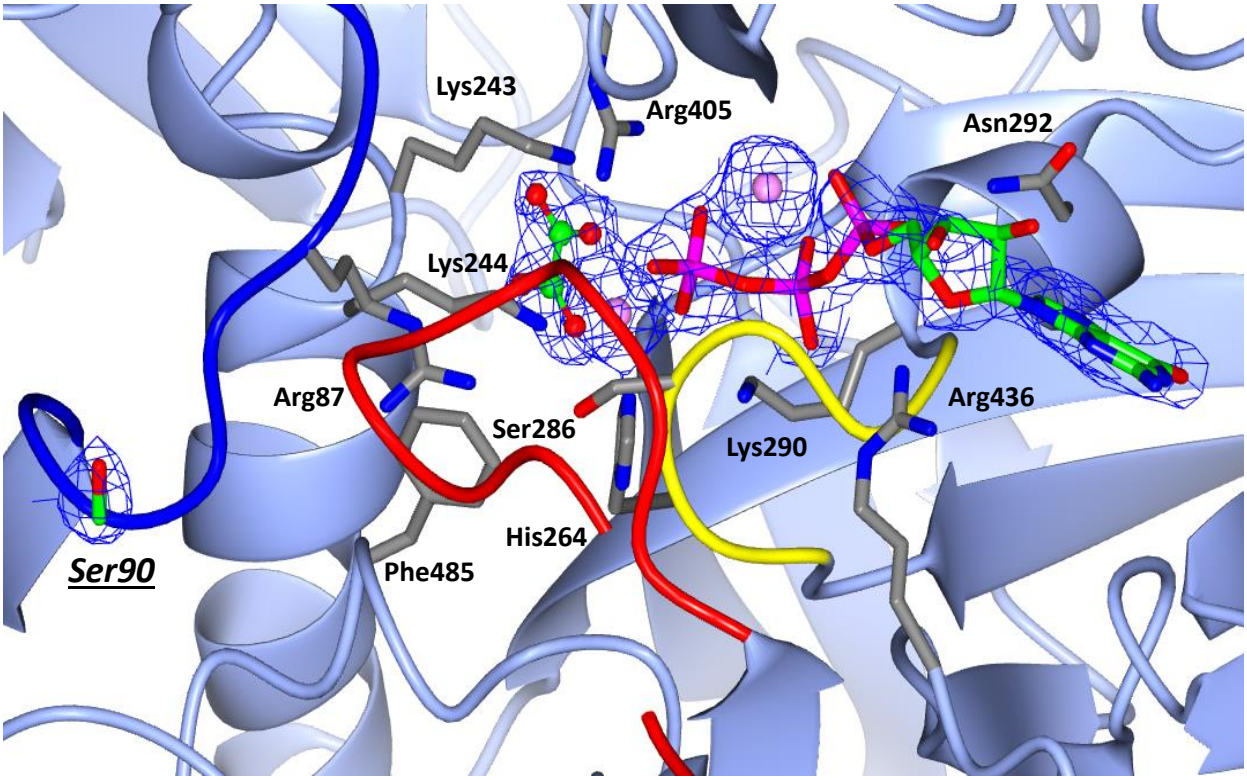


Figure 20: Ribbon structure of WT hcPEPCK with oxalate and GTP
Ribbon structure of wildtype hcPEPCK is coloured ice blue. The R-loop is coloured blue, the P-loop is coloured yellow, and the omega loop is coloured red. Serine at position 90 is highlighted in green, and its corresponding electron density is contoured to 2σ. Both the oxalate ligand and the GTP nucleotide are coloured by atom and rendered as ball-and-stick models. Neighbouring amino acids within 3Å are coloured grey. The 2Fo-Fc electron density map surrounding oxalate and GTP is contoured to 2σ. The 2Fo-Fc map for the metals is contoured to 3σ. Manganese ions are coloured light pink and represent the M1 and M2 metals. This diagram was generated using CCP4MG.[26]

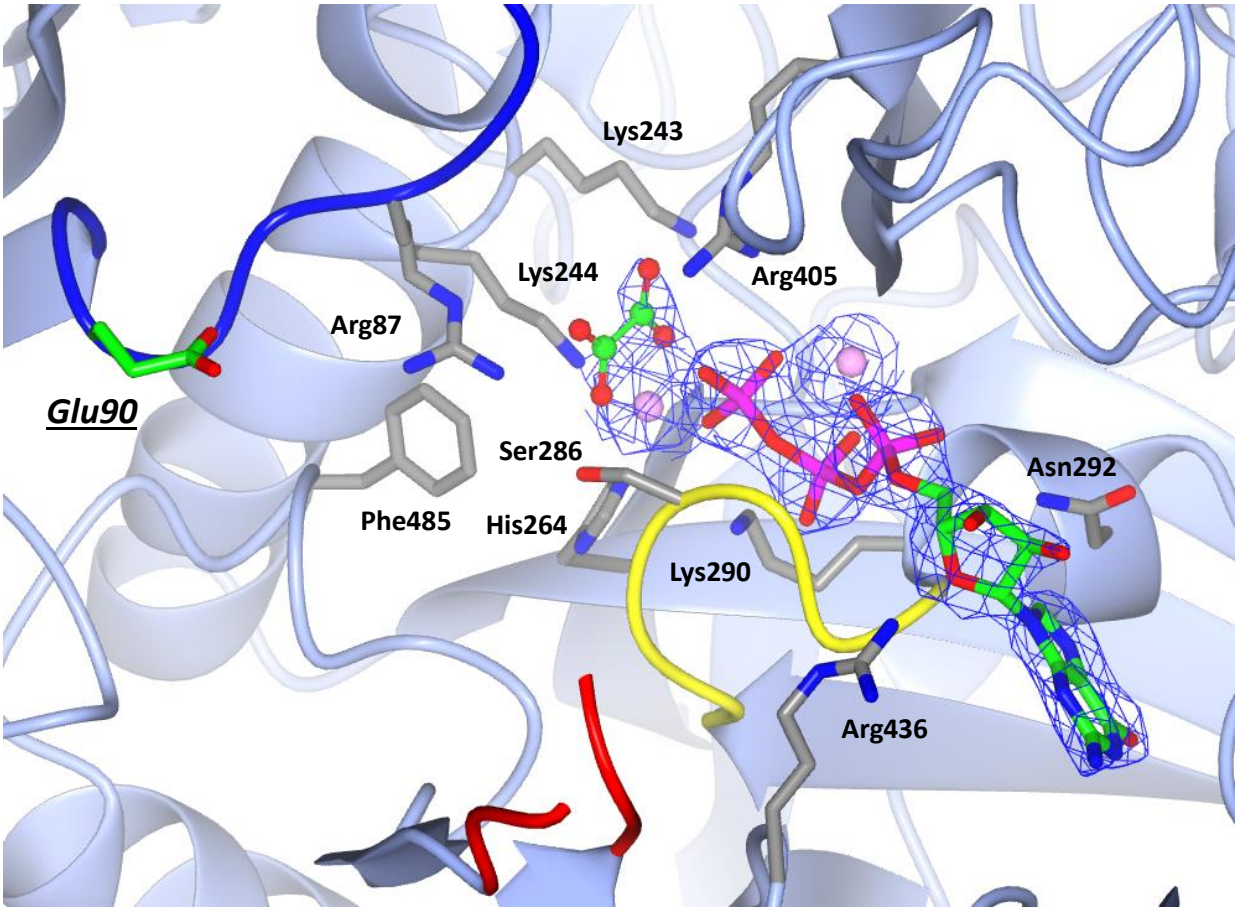


Figure 21: Ribbon structure of S90E hcPEPCK with oxalate and GTP
Ribbon structure of S90E hcPEPCK is coloured ice blue. The R-loop is coloured blue, the P-loop is coloured yellow, and the partial omega loop is coloured red. While the glutamate at position 90 is highlighted in green, the electron density for the sidechain could not be shown. Both the oxalate ligand and the GTP nucleotide are coloured by atom and rendered as ball-and-stick models. Neighbouring amino acids within 3Å are coloured grey. The 2Fo-Fc electron density map surrounding oxalate and GTP is contoured to 2σ. The 2Fo-Fc map for the metals is contoured to 3σ. Manganese ions are coloured light pink and represent the M1 and M2 metals. This diagram was generated using CCP4MG.[26]

Table 1: Simplified WT and S90E hcPEPCK crystallographic data statistics*

	WT hcPEPCK GTP Oxalate	S90E hcPEPCK GTP Oxalate	WT hcPEPCK GTP
Space group	P 1 21 1	P 1 21 1	P 1 21 1
Unit Cell Dimensions	a = 45.125 b = 114.507 c = 114.526 α = γ = 90.00° β = 91.61°	a = 44.803 b = 114.001 c = 114.271 α = γ = 90.00° β = 91.76°	a = 45.120 b = 112.766 c = 60.336 α = γ = 90.00 β = 107.74°
Resolution Limits	35.43 – 2.10	34.14 – 2.45	57.47 – 1.70
Completeness (%)	98.81 (96.65)	92.38 (91.69)	99.55 (96.41)
I/sigma	9.3 (1.6)	7.77 (1.45)	16.1 (0.8)
CC1/2	0.987 (0.638)	0.984 (0.537)	0.998 (0.659)
Rpim	0.071 (0.498)	0.087 (0.500)	0.030 (0.448)
Rmerge	0.125 (0.657)	0.153 (0.865)	0.072 (0.956)
Rfree	0.267 (0.332)	0.266 (0.411)	0.229 (0.291)
Rwork	0.212 (0.307)	0.219 (0.340)	0.188 (0.249)

**Values in parentheses represent the highest resolution shell*

2.3.2 WT and S90E hcPEPCK kinetics & inhibition

When comparing wildtype and S90E hcPEPCK, it became clear that just like the structural approach, there is not a very significant difference between the wildtype and S90E mutant kinetic activities. Both enzymes appear to be equally active in the gluconeogenic direction when at saturating substrates concentrations (Table 2). The only apparent difference in this direction of catalysis is the respective K_M values for OAA and GTP. The phosphomimetic mutant has higher K_M value for its substrates in the gluconeogenic direction by a factor of 2 than that of the wildtype enzyme. A more significant observation can be made in the opposite direction, that of carbon fixation. The specific activity of the mutant is half that of the wildtype. The respective K_M values for PEP and GDP are not significantly different.

Table 2: Kinetic characterization of WT and S90E hcPEPCK using the standard PEPCK assay

	a) OAA + GTP → PEP + GDP + CO_2					
	K_M (μM)		**k_{cat} (s^{-1})**		**k_{cat}/K_M (M^{-1} s^{-1})**	
	OAA	GTP	OAA	GTP	OAA	GTP
WT	33.8 ± 4.6	69.8 ± 10.9	33.4 ± 1.0	29.1 ± 0.96	9.9 x 10^5	4.2 x 10^5
S90E	74.0 ± 11.5	177.4 ± 20.2	28.9 ± 1.3	28.1 ± 0.79	3.9 x 10^5	1.6 x 10^5
	b) PEP + GDP + CO_2 → OAA + GTP					
	K_M (μM)		**k_{cat} (s^{-1})**		**k_{cat}/K_M (M^{-1} s^{-1})**	
	PEP	GDP	PEP	GDP	PEP	GDP
WT	298.0 ± 39.5	29.3 ± 8.9	16.1 ± 0.60	14.1 ± 0.61	5.4 x 10^4	4.8x 10^5
S90E	252.1 ± 51.6	19.6 ± 4.5	5.9 ± 0.31	8.8 ± 0.29	2.3 x 10^4	4.5 x 10^5

Another kinetic analysis that was done included the determination of PEPCK inhibition using oxalate inhibitor. Oxalate is a well characterized competitive inhibitor of GTP-dependent PEPCK with its K_I in the micromolar range.[76] Oxalate is a useful analog of the reaction intermediate (enolate of pyruvate), allowing for the analysis to be done in either direction of catalysis. The resulting information gained from the kinetic replots (the ratio of K_M/specific activity) shows that as a competitive inhibitor of PEP, oxalate binds with a slightly lower affinity to the S90E PEPCK mutant than that of the wildtype (Table 3). There is no significant difference between the S90 wildtype and S90E phosphomimetic mutant enzyme. This correlates to the earlier information gleaned from the co-crystal structures and kinetic characterization.

Table 3: Inhibition of WT and S90E hcPEPCK by oxalate (intermediate analog) against PEP

	a) WT hcPEPCK	
	K_M/V_{max} K_I (μM)	**$1/V_{max}$ K_I (μM)**
Oxalate	83.62 ± 25.1	1257 ± 536.4
	b) S90E hcPEPCK	
	K_M/V_{max} K_I (μM)	**$1/V_{max}$ K_I (μM)**
Oxalate	121.008 ± 18.7	956.0 ± 212.0

Lastly, an IC_{50} analysis of INSIG binding to the S90E enzyme was attempted. Due to weak binding, a full IC_{50} dose-response curve could not be generated, however data was obtained to demonstrate a loss of 30% activity when the S90E mutant was exposed to 300 μM of the INSIG 8-mer peptide. This was not enough data to generate a sigmoidal dose-response curve characteristic of a complete IC_{50} analysis. The incomplete data was plotted as a dose-response curve (Figure 22) using GraphPad Prism to obtain an estimated IC_{50} of 675 ± 280.4 μM. This estimate was used as a starting point for subsequent analysis.

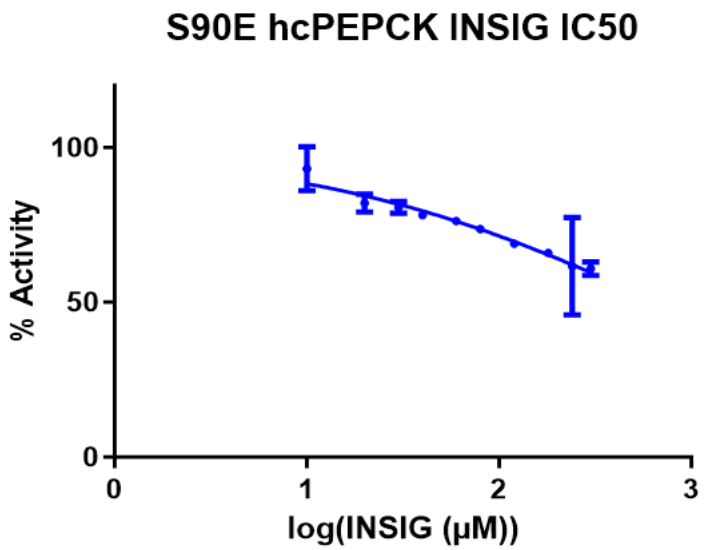


Figure 22: INSIG IC50 of S90E hcPEPCK
S90E hcPEPCK activity data at increasing concentrations of INSIG peptide was plotted on GraphPad and fit to a dose-response curve to determine IC50.

2.3.3 S90E hcPEPCK differential scanning fluorimetry & INSIG binding

Faced with inconclusive binding of INSIG, another type of binding assay was needed. Differential Scanning Fluorimetry (DSF) was used to examine PEPCK protein stability when increasing amounts of INSIG peptide were added. DSF involves gradually heating up a protein in a solution with a hydrophobic fluorescent dye (SYPRO orange), and monitoring fluorescence

of that dye binding when the protein unfolds and exposes hydrophobic residues. The thermocycler used for the DSF assays cycled through temperatures ranging from 20 to 90°C, then a cooling period gradually back down to 20°C again, allowing for the unfolding and refolding of the S90E hcPEPCK protein. Adding a ligand to the DSF assay can increase a protein's stability; in this case the ligand is the INSIG peptide. A protein with a more stable conformation has a higher melting temperature because it takes more energy to unfold. Therefore, when comparing the melting point of the S90E enzyme with and without INSIG present, an obvious T_M shift can be observed (Figure 23).

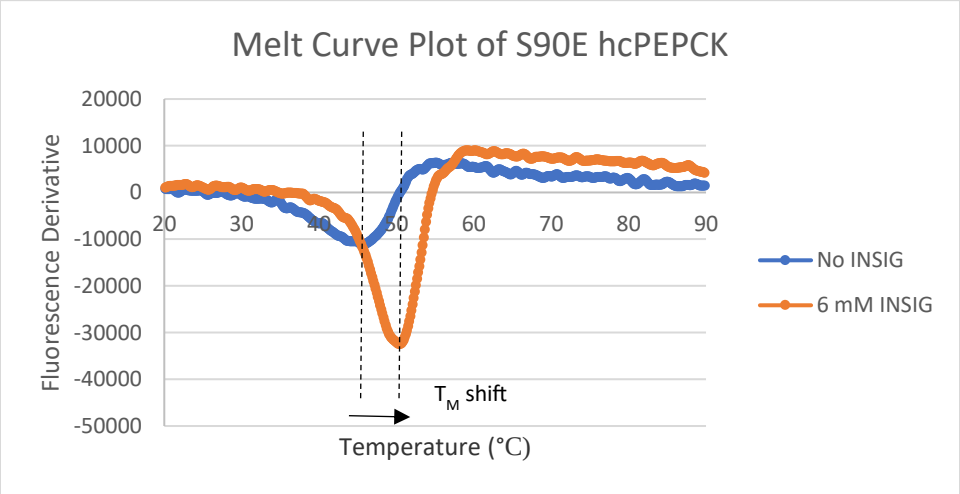


Figure 23: Derivative plot showing T_M shift of S90E hcPEPCK upon adding INSIG
Two superposed fluorescence derivative melt curves visualizing a shift in melting point upon adding a high concentration (6mM) of INSIG peptide. Resulting TM shifts were calculated for each concentration of INSIG peptide added, and used to generate a binding isotherm.

Plotting the fluorescence derivative data versus temperature allows for the generation of melting curves for each of the samples (see Appendix C). The melting point (T_M) can be determined from the melting curve plots where the fluorescence derivative is lowest. Subtracting the T_M of S90E hcPEPCK without INSIG from the TM of S90E hcPEPCK with INSIG yields a T_M shift. The average of these T_M shifts at each duplicate INSIG concentration is listed in Table 4, and used to

generate an INSIG binding isotherm (Figure 24). The data was fit to a ligand-binding dose-response non-linear regression model in SigmaPlot11, with a calculated binding constant (K_D) of 294.8 ± 0.09 μM INSIG, suggesting tighter binding than the K_D determined from the previous spectrophotometric assay listed earlier.

Table 4: List of melting temperatures under increasing INSIG peptide concentration and their resulting thermal shifts

[INSIG] (mM)	T_m	T_m	T_m shift	T_m shift
0	45.738	45.015	0.0000	0.0000
0.5	49.065	47.763	3.6882	2.3865
1	49.387	49.242	4.0108	3.8661
2	50.834	50.400	5.4580	5.0239
4	50.072	49.059	4.6957	3.6835
6	50.361	-	4.9849	-
10	50.369	50.514	4.9924	5.1371
15	50.658	49.935	5.2818	4.5584

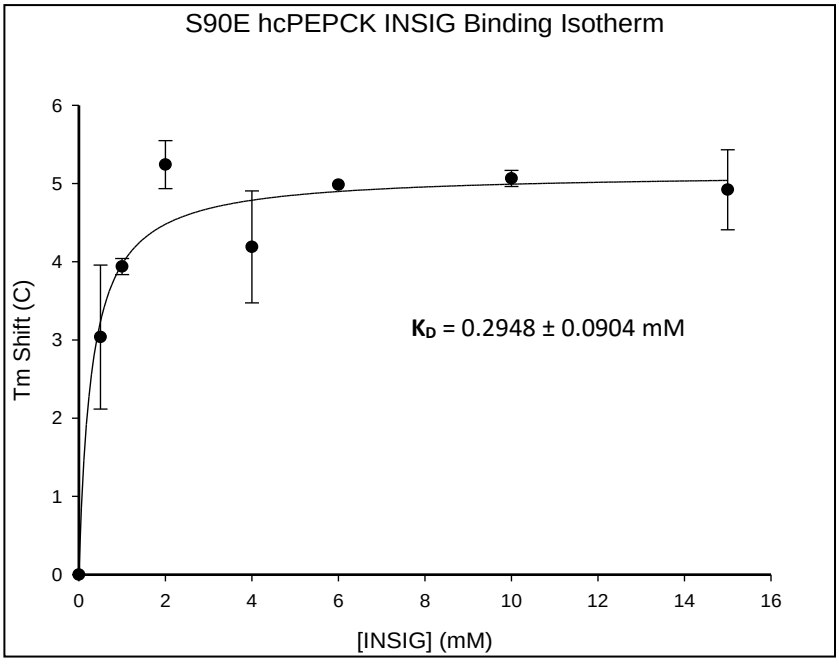


Figure 24: S90E hcPEPCK INSIG Binding Isotherm
The binding isotherm was generated by plotting Tm shift in duplicate vs INSIG concentration and fitting the data to the ligand bonding one-site saturation nonlinear regression model in the Enzyme Kinetics extension in SigmaPlot11.

2.4 Conclusions

The crystallographic data was inconclusive for the binding of INSIG to the phosphomimetic mutant. Further analysis of INSIG binding might only be possible with the full INSIG protein, as the small INSIG-2 peptide does not bind very tightly, which could be in part due to truncating the sequence to only the loop 8-mer for easier crystallization. It should also be noted that the original study by Xu et al. used a larger INSIG-1 peptide 15 amino acids long (LWWTFDRSRSGLGLG), and that the extra amino acids might be required for binding. Trying to crystallize the entire INSIG protein might prove difficult, however, due to the large majority of it being encompassed in the ER membrane. Special considerations for transmembrane proteins need to be made, but crystallizing these hydrophobic proteins is not impossible. In addition, if the protein kinase activity is GTP-dependent, then INSIG binding may also require the presence of GTP or GDP nucleotides.

From the structural data gathered, using the assumption that the S90E mutation is an adequate mimic for the phosphorylation of Ser90, it appears that hcPEPCK does not undergo any significant conformational change upon phosphorylation. Ser90 is located on the R-loop, one of three catalytically active mobile loops of the enzyme's active site that coordinate substrate binding. The R-loop in particular coordinates the binding of the OAA/PEP substrate. Although the phosphomimetic mutation does not distinctly affect the structural conformation of said R-loop, it likely affects the structural stability and increases disorder of the loop, as can be seen by the partially closed conformation of the S90E structure. This can also be seen from the kinetic assays by observing and comparing enzyme turnover for both the wildtype and mutant in both directions of catalysis – turnover is significantly decreased in the reverse direction, and the catalytic efficiency is thus an order of magnitude less. The increased disorder of the R-loop in

the phosphomimetic mutant likely also contributed to difficulties in crystallization. The only S90E hcPEPCK structure able to be obtained was the GTP + oxalate co-crystal structure, where the presence of the bound ligand and nucleotide was sufficient enough to stabilize the enzyme enough to crystallize. Conversely, the S90E mutant crystal with only GTP nucleotide bound was unable to be obtained, unlike the wildtype enzyme, which corresponds to the lower K_M for GTP as determined via the kinetic assays. Also, it is important to point out that although both wildtype and mutant PEPCK were attempted to be co-crystallized with INSIG, the majority of INSIG binding including the DSF assays were performed using the S90E mutant only, and INSIG binding with the wildtype enzyme was never reinvestigated.

Previous research has suggested that phosphorylation of Ser90 signals for PEPCK to be translocated to the ER.[46] This change in cellular location allows for hcPEPCK to moonlight as a protein kinase without a significant conformational change. Since the majority of gluconeogenesis occurs in the cytoplasm and the INSIG protein is located in the membrane of the ER, this separation allows for the two separate and distinct enzymatic functions of hcPEPCK (gluconeogenic carboxy-lyase vs lipogenic protein kinase). Therefore, instead of considering phosphorylation of Ser90 to be an ON/OFF switch of canonical PEPCK activity, a more appropriate analogy would be a dimmer switch, and even then, the effect on enzyme activity is unidirectional, only observed in the reverse direction. This appears to be in direct contrast to the research performed by Xu et al., where they state that AKT-mediated phosphorylation of the wildtype or introduction of the S90E mutation reduces the binding affinity of OAA and its corresponding PEPCK enzyme activity.[46] However, a closer look into the supplementary material and a comparison of kinetic assays reveals that the relative decrease in enzyme activity of the S90E mutant as reported by Xu et al. was assayed in the reverse direction (OAA synthesis)

using a commercially available PEPCK activity kit (BioVision, K359).[46] Their results (Xu et al., Extended Data Fig.2u & Fig.2v) also show not a complete abolishment of S90E hcPEPCK activity, but rather a significant decrease (~75%) compared to the wildtype, which corresponds to our own research on the S90E phosphomimetic mutant (~63% decrease in k_{cat}). To my knowledge, all *in vitro* PEPCK activity in their study was performed in the reverse direction of catalysis using the PEPCK activity kit. Conversely, they assayed wildtype and S90E binding affinity for OAA by incubation of 200 μM OAA with the wildtype and S90E constructs bound to Ni-NTA resin for 30 minutes. Binding affinity for OAA was determined as a percentage of OAA found in the supernatant fraction after centrifugation. As per our own research where a small observed increase in the K_M for OAA was observed for the phosphomimetic mutant, they determined a lower relative binding affinity for OAA for S90E hcPEPCK. This can be explained due to the location of the phosphomimetic serine 90 on the R-loop, the catalytically active mobile loop that coordinates binding of the OAA substrate. The unidirectional decrease in PEPCK activity of the phosphomimetic mutant again can be attributed to the difference in Gibbs free energy needed to overcome the activation barrier of the reverse direction (OAA synthesis) compared to the forward direction of gluconeogenesis (PEP synthesis).

In summary, the research performed here provides additional evidence to the study postulating a secondary role for hcPEPCK as a protein kinase. Most importantly, expanding upon previous research we suggest that the phosphorylation of serine 90 does not structurally change the conformation of hcPEPCK, and that the lipogenic activity is a result of compartmentalization to the ER. There is no apparent evolutionary drive to completely abolish PEPCK activity when P-Ser90 results in translocation to the ER and gluconeogenesis occurs primarily in the cytosol. In

addition to gluconeogenesis, this moonlighting role for PEPCK as a lipogenic enzyme adds to the reasoning as to why cytosolic PEPCK is upregulated in cancer cells.[47,77]

Chapter 3: A complete characterization of ATP-dependent *E. coli* PEPCK and its differences compared with GTP-dependent PEPCK enzymes

3.1 Introduction

Although the first PEPCK crystal structure to be determined was that of an ATP-dependent PEPCK, much of today's research in the PEPCK enzyme family is done on the GTP-dependent PEPCKs found in humans and other higher eukaryotic species. Logically, this makes sense as biological research tends to focus on human metabolism, and PEPCK is one of the key gluconeogenic enzymes central to regulating glucose homeostasis in the human body. It is largely established that for higher eukaryotes, GTP-dependent PEPCK enzymes catalyze the reaction in the gluconeogenic direction, that is in the direction of PEP production (Figure 25).

Gluconeogenesis (forward)

$$OAA + GTP \rightarrow PEP + GDP + CO_2$$

Carbon Fixation (reverse)

$$PEP + GDP + CO_2 \rightarrow OAA + GTP$$

Figure 25: PEPCK-catalyzed reactions and their role *in vivo*
Directional schematic of the forward and reverse reactions catalyzed by GTP-dependent PEPCK enzymes.

However, this reaction, while for the most part is kinetically unidirectional *in vivo*, is actually thermodynamically reversible, as can be seen in several different enzyme assays *in vitro*. There is only a slight preference in the kinetic parameters for the forward direction of PEP synthesis, and in fact it has been demonstrated that the free energy difference for the PEPCK-catalyzed reaction is close to zero.[67] Previous research on both the rat and human cytosolic enzymes, both GTP-dependent members of the PEPCK family, shows that *in vitro* enzyme catalysis is higher in the forward direction of gluconeogenesis (PEP synthesis) than the reverse direction, though this

is while being subjected to the saturating substrate concentrations of the *in vitro* assay, and is likely not how the enzyme operates in a biological system where the directional activity is largely determined by substrate availability and compartmentalization.[67]

ATP-dependent PEPCK enzymes likewise are thought to preferentially catalyze the reaction in the direction of gluconeogenesis. ATP-PEPCK has been studied extensively in Bakers' yeast, *Saccharomyces cerevisiae*, all the way back in the 1970's.[39,78] This study on Bakers' yeast was the first to acknowledge that ATP-PEPCK had an absolute requirement for a divalent cation for its enzyme activity, and that there are two distinct binding sites for the metal: one for the cation itself, and one for a metal-ATP complex. Six different divalent metals were assayed for enzyme activity: Mn^{2+}, Mg^{2+}, Cd^{2+}, Co^{2+}, Zn^{2+}, Ca^{2+}. Manganese was found to have the highest activity, with magnesium having only about 20% as much activity. The other metals tested have negligible activity. They noted that although manganese (Mn^{2+}) could fulfill both roles, there was a synergistic effect when magnesium (Mg^{2+}) or cadmium (Cd^{2+}) was added (although Cd^{2+}'s effect appears to be pH dependent).[39] This synergistic effect between Mn^{2+} and Mg^{2+} has also been seen in GTP-PEPCK isolated from chicken liver mitochondria[79,80] and cytosol of rat liver[81], suggesting that this two metal requirement is a common characteristic of PEPCKs from several different species, as well as across the entire superfamily of PEPCK enzymes.

The study on Bakers' yeast was concluded by stating that yeast PEPCK catalysis is influenced by the ATP:ADP ratio in the cell. It was suggested that yeast PEPCK was activated by free ATP and inhibited by high ADP concentrations. When the ATP:ADP ratio is high, gluconeogenesis can occur, while in reverse where there's a high-energy demand, inhibition sets in. There is evidence that PEPCK and pyruvate carboxylase, the two enzymes located at the integral pyruvate-OAA-

PEP node that together convert pyruvate to PEP, are both inhibited by ADP.[82,83] This is also supported by the fact that during high-energy demand when ATP concentrations are low and glucose concentrations are high, PEPCK gene expression is repressed, suggesting several levels of PEPCK regulation.

Interestingly, there are a few species of ATP-PEPCK that are capable of catalyzing the reverse reaction *in vivo* for carbon fixation. It was initially believed that microorganisms use PEPCK only for gluconeogenesis, and phosphoenolpyruvate carboxylase (PEPC) for carbon fixation. The PEPC-catalyzed reaction is irreversible, and doesn't use a phosphate acceptor as a cofactor, but rather produces inorganic phosphate instead.

1) Phosphoenolpyruvate carboxylase (PEPC)

$PEP + HCO_3^- \longrightarrow OAA + Pi$

2) Phosphoenolpyruvate carboxykinase (PEPCK)

$PEP + ADP + CO_2 \rightleftharpoons OAA + ATP$

3) Pyruvate carboxylase (PC)

$Pyruvate + HCO_3^- + ATP \longrightarrow OAA + ADP + Pi$

4) Pyruvate kinase (PK)

$PEP + ADP \longrightarrow pyruvate + ATP$

Figure 26: Important enzymes in central carbon metabolism
A list of four metabolic enzymes and their respective catalyzed reactions integral to the PEP-pyruvate-OAA node and central carbon metabolism.

Initial research focusing on other CO_2-fixing enzymes in *Pseudomonas fluorescens* accidentally discovered PEPCK activity when trying to monitor CO_2 fixation into OAA using cell extracts.[41] They had identified PEPC enzyme in the cell extract, but upon addition of ADP to the reaction mixture, OAA production had significantly increased. Since PEPC does not use a nucleotide cofactor, they attributed this increase in CO_2 fixation into OAA to the presence of PEPCK in the cell extracts.[41] Later in 1977, and then in 1986, CO_2-fixing activity of GTP-PEPCKs was discovered in various eukaryotic parasites.[84–86]

In 1993, researchers characterized a CO_2-fixing PEPCK from the obligate anaerobe *Anaerobiospirillum succiniproducens*, a rare cause of human bacteremia.[40] This anaerobic microorganism produces high amounts of succinate as a by-product of glucose fermentation. It was suggested that PEPCK catalyzes the conversion of PEP to OAA as the first step of the fermentation pathway leading to succinate production. Lastly, more recent research provides evidence that PEPCK isolated from the aerobic bacterium, *B. subtilus*, is capable of catalyzing the reaction in the direction of carbon fixation when other, more appropriate CO_2-fixing enzymes have their gene expression knocked out.[38] It was demonstrated that when pyruvate kinase (PK) is knocked out, instead of PEP/ADP being converted to pyruvate/ATP, PEPCK is able to convert PEP into OAA instead, producing the ATP required by the cell. Understandably, due to the more unfavourable thermodynamics of the PEPCK-catalyzed reaction, the PK-knockout mutants grew more slowly.[38]

Historically, the ATP-PEPCK that is the most well characterized from a kinetic standpoint is the enzyme isolated from Bakers' yeast, or *Saccharomyces cerevisiae*.[42] Yeast PEPCK is a bit of an exception, since yeast is a eukaryotic organism and its PEPCK is ATP-dependent, unlike other eukaryotes higher up in the phylogenetic tree. Unlike the yeast's tetrameric enzyme, ATP-

PEPCK from *E. coli* is monomeric (EcPEPCK).[21] Despite being the first PEPCK to have its structure determined and being the best structurally characterized ATP-PEPCK enzyme, the *E. coli* enzyme remains poorly characterized from a kinetic standpoint. Many of the conclusions reached in the structure-function relationship for the ATP-dependent enzymes are drawn from structural data obtained using GTP-PEPCKs as a surrogate. It was assumed that EcPEPCK behaved in a similar fashion to the monomeric GTP-PEPCKs found in higher eukaryotes that our laboratory is well versed in. A comprehensive analysis of the rat cytosolic PEPCK, and to some extent the human cytosolic enzyme, has elucidated the structure-function relationship of the GTP-dependent class of PEPCKs, but this is lacking in the ATP-dependent class. Using EcPEPCK as a proxy for ATP-PEPCK, and the standard kinetic and structural approaches used previously in our laboratory to study PEPCK, I have completed a full kinetic characterization and obtained various structural data to further our understanding of the ATP-dependent class of PEPCK enzymes that is currently lacking.

3.2 Materials & Methods

3.2.1 Materials

Nickel-NTA resin was from UBP-Bio and P-6DG resin was from Bio-Rad. PEP potassium salt, NADH disodium salt (>96% purity), TCEP and DTT were from ChemImpex, ADP disodium dihydrate was from BioShop, ultrapure ATP disodium trihydrate was from GoldBio, PGA lithium salt (>95% purity) and oxalic acid (98% purity) were from Sigma Aldrich, and OAA (>97% purity) was from Millipore Sigma (CalBioChem). The reagents for the sulfopyruvate synthesis, sodium sulfite (>98% purity) and bromopyruvic acid (97% purity), were purchased from ChemImpex and Fisher Scientific respectively.

3.2.2 Enzymes

The coupling enzymes pyruvate kinase (PK), lactate dehydrogenase (LDH) and malate dehydrogenase (MDH) were purchased from Calzyme Laboratories. The plasmid expressing His6-SUMO protease was a gift from C. Lima (Sloan-Kettering Institute, New York, NY) and was expressed and purified as previously described.[24]

3.2.3 β-sulfopyruvate (BSP) chemical synthesis

Following the protocol by Weinstein and Griffith[87], BSP was chemically synthesized via direct reaction of sulfite with bromopyruvic acid in aqueous solution. 2.84g of sodium sulfite (22.5 mmol) was dissolved in 12.5mL of water, and 1.72g of bromopyruvic acid (10 mmol, 10 mL) was added dropwise to the mixture over a 30-minute period with vigorous stirring. The chemical

reaction took place in a small round-bottom flask kept at 80°C using a hot oil bath heated on top of a hot plate.

The crude mixture was dissolved in water and applied to a Dowex I column. The sample was then eluted using a linear gradient (formed between 750 mL water and 750 mL 2M HCl) followed by an additional 200 mL of 2M HCl. Eluted fractions were collected and monitored for BSP at 280 nm (**Appendix F**). Fractions with A_{280} values greater than 0.3 were assayed for BSP using malate dehydrogenase (MDH) enzyme (Figure 27). Fractions found to contain BSP were pooled, concentrated by rotary evaporation, and washed free of HCl. The resulting solution was neutralized to pH 7 using dilute lithium hydroxide, yielding the BSP product as a dilithium salt. Total yield of the BSP salt was ~80%.

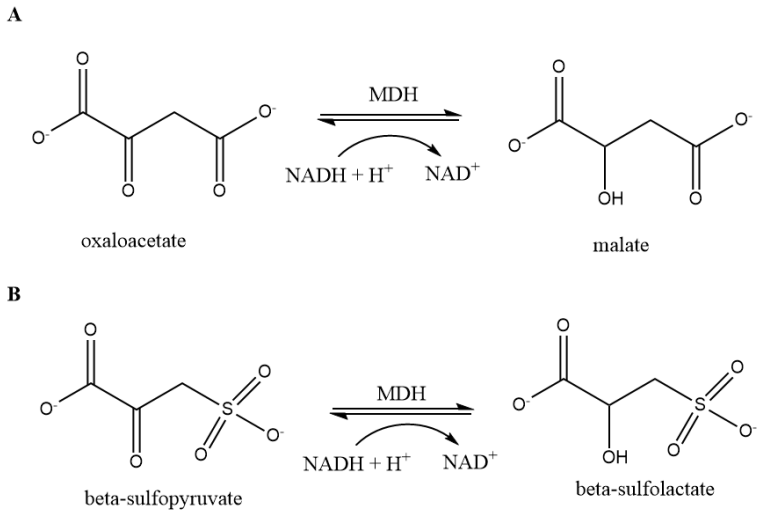


Figure 27: Reactions catalyzed by malate dehydrogenase (MDH) using NADH as a cofactor. Reaction A is the typical reaction catalyzed by MDH used in the PEPCK assay. Reaction B shows the ability of MDH to use BSP as an alternative substrate.[87]

The concentration of the purified BSP product was determined enzymatically using MDH. The Li_2BSP product was dissolved to create a 100 mM stock solution (molecular weight of 182.012 g/mol) in 1M HEPES pH 7.5. The MDH assay conditions included 100 mM HEPES pH 7.5, 300 µM NADH, 10 U of MDH and 100 µM of Li_2BSP. The spectrophotometric assay was performed in a similar manner to the PEPCK assay, using a 1 mL reaction volume, and a 340 nm wavelength at 25°C. Kinetic replicates were averaged and the absorbance was converted to determine a precise concentration of BSP using the Beer-Lambert Law, assuming a 1:1 stoichiometric ratio of BSP to NADH (ε of NADH at 340 nm is 6.22 $mM^{-1}cm^{-1}$).

3.2.4 Cloning & transformation of EcPEPCK

The gene sequence of *E. coli* PEPCK in a pE-SUMO-star (Kan) vector (LifeSensors PE-1101-0020) was synthesized by GenScript. The resulting plasmid allows for the recombinant expression of EcPEPCK as a N-terminal HIS_6-SUMO fusion protein. The plasmid was transformed into BL21(DE3) *Escherichia coli* cells as previously described (**see section 2.2.3**), and glycerol cell stocks of EcPEPCK were stored at -80°C.

3.2.5 EcPEPCK protein expression and purification

EcPEPCK was expressed and purified as previously described for human PEPCK (**see section 2.2.4**). In brief, EcPEPCK was expressed at 20°C for 20-24 hours in 6L of ZYP-5052 autoinduction media supplemented with 50 µg/mL of kanamycin.[66] After expression, cells were lysed via two passes through a French press cylinder at 1000 psi and the soluble lysate fraction was passed through a Ni-NTA column to bind the His_6-SUMO tagged EcPEPCK. The column

was washed, the tagged EcPEPCK was eluted off the column and the tag was cleaved via an overnight digest with SUMO protease. The protein was passed through the Ni-NTA column once more to remove the cleaved tag from the sample. The purified EcPEPCK was concentrated to ~30 mg/mL via 30 kDa filter centrifugation, and the final protein concentration was determined using a NanoDrop spectrophotometer (A_{280} of 0.1%: 1.3 mg/mL^{-1}). The concentrated protein was then flash-frozen into 30 µl aliquots and stored at -80°C.

3.2.6 EcPEPCK crystallization

Wildtype EcPEPCK crystals were grown using the hanging drop vapour diffusion method at ambient temperature. Initial attempts to grow EcPEPCK crystals were done using a reservoir solution of 0.1M sodium acetate pH 4.5 and 16 – 30% PEG 3350, adding water to adjust the final volume of the wells to 700 µl. The protein solution contained 10 mg/mL EcPEPCK, 25 mM HEPES pH 7.5, 10 mM DTT and 25 mM $MnCl_2$. Drops were set up on a siliconized cover slide using 3 µl protein and 3 µl of the reservoir solution. All of the EcPEPCK constructs crystallized using this condition, however the crystals were of poor quality, requiring seeding. 10% acetone was later included in the wells as an additive to improve crystal morphology and diffraction quality, but only to a minimal extent.

The PGA co-crystals were grown using the above acetate condition (0.1M sodium acetate pH 4.5 and 26 – 30% PEG 3350). The protein solution contained 15 mg/mL EcPEPCK, 25 mM HEPES pH 7.5, 10 mM DTT, 2 mM ADP and 25 mM $MnCl_2$. Crystals were setup using 3 µl protein solution, 3 µl mother liquor and 0.5 µl 100 mM PGA added directly to the drops on a siliconized cover slide. The crystal tray was seeded the next day with concentrated seed stock made using seed beads (Hampton Research, HR2-320). Amorphous crystals grew two days after seeding, and

were harvested by soaking for 5 minutes in a cryoprotectant containing 30% PEG 3350, 10% PEG 400, 0.1M sodium acetate pH 4.5, 25 mM $MnCl_2$, 10 mM DTT and 2 mM ADP, before being cryo-cooled by immersion in liquid nitrogen.

Another crystal condition was used to produce larger, higher quality crystals. The new reservoir solution contained 0.1M BisTris pH 6.5, 24-28% PEG 3350 and 10 mM $MnCl_2$, with water added to make up a final volume of 700 µl. The protein solution for the ATP oxalate structure contained 10 mg/mL EcPEPCK, 25 mM HEPES pH 7.5, 10 mM DTT, 1 mM oxalate and 1 mM ATP, while the protein solution for the ATP BSP structure contained 15 mg/mL EcPEPCK, 25 mM HEPES pH 7.5, 10 mM DTT, 100 mM BSP and 1 mM ATP. Drops were set up using 3 µl of protein solution and 3 µl of the reservoir solution on a siliconized cover slide. Both crystals grew within one to two days at ambient temperature. The ATP oxalate crystals were harvested and soaked for 5 minutes in a cryoprotectant containing 30% PEG 3350, 10% PEG 400, 0.1M BisTris pH 6.5, 10 mM DTT, 10 mM $MnCl_2$, 1 mM oxalate and 1 mM ATP. The ATP BSP crystals were soaked for 1 hour in a similar cryoprotectant containing 1 mM ATP and 100 mM BSP to form the EcPEPCK ATP BSP complex. The crystals were then cryo-cooled by immersion into liquid nitrogen.

Diffraction data on the EcPEPCK oxalate complex was collected at our home source (CuKα = 1.54Å) Xray diffractometer at the University of Waterloo. The BSP and PGA crystals were diffracted at the CHESS synchrotron (beamline 7B2) at Cornell University. The resulting datasets were indexed and scaled using HKL2000 and converted into mtz files using CCP4.[71] The phases were determined through molecular replacement (MOLREP) using a wildtype EcPEPCK structure found on the PDB database (PDB ID: 1OS1). Refinement was also done through CCP4, through several rounds of RefMac5.[71]

3.2.7 EcPEPCK kinetic assays

Kinetic assays in both directions of catalysis were performed as previously described for the GTP-dependent enzyme with minor changes (**see section 2.2.6**). For the OAA → PEP direction assay, 500 µM ATP was substituted instead of the GTP nucleotide used in the human cytosolic assay. ADP was not added directly to the assay in this direction as a cofactor for the PK coupling enzyme, due to EcPEPCK producing ADP via the catalyzed reaction. Additional PK enzyme was added for a total of 100U of PK, 30U of LDH and 2.5 µg of EcPEPCK, followed by OAA to start the reaction. For the PEP → OAA direction, all components remained the same as that of the GTP-PEPCK assay, with the exception of using 1 mM ADP instead of GDP. The assay in this direction was initiated by adding 2.5 µg of EcPEPCK enzyme. All assays were performed at 25°C, and PEPCK activity was monitored at 340 nm for 10 minutes.

Divalent cation assays were performed in the PEP → OAA direction using the standard assay mix above, analyzing Mn^{2+}, Fe^{2+}, Ca^{2+}, Co^{2+}, Zn^{2+} and Ni^{2+} in the presence of 4 mM excess $MgCl_2$ (M2 metal). EcPEPCK activity was assayed using increasing concentrations of divalent cation (0, 1, 2, 5, 10, 15, 20, 40, 60 µM) to determine which cation activated EcPEPCK the most as the M1 metal. Activity was measured using 4 mM saturating $MgCl_2$ for reference. All cations were also assayed in the absence of $MgCl_2$.

The oxalate inhibition assays were performed using the standard assay mix above, and in both directions of catalysis EcPEPCK activity was assayed using increasing concentrations of oxalate from 0 to 80 µM. The BSP inhibition assays were performed in the OAA → PEP direction using the standard assay conditions. The PGA IC50 curve was performed in the PEP → OAA direction

using the standard assay conditions, with the exception of using a lower concentration of PEP, closer to its K_M (1 mM).

The kinetic data from the activity assays were fit non-linearly to the standard Michaelis-Menten equation using SigmaPlot11to determine kinetic parameters in both directions of catalysis. (**Equation 1**). Apparent V_{max} and K_M parameters were generated for each oxalate and BSP concentration and were re-plotted in a linear relationship against inhibitor concentration to calculate inhibition constants. A competitive K_I for BSP was determined in the forward direction, and a competitive K_I for oxalate was determined in both directions of catalysis by plotting K_M/V_{max} vs [inhibitor] (**Equation 2**). Similarly, any uncompetitive effects were determined by plotting $1/V_{max}$ vs [inhibitor] (**Equation 3**).

Equation 1: $$v = \frac{V_{max} \times [S]}{K_M + [S]}$$

Equation 2: $$\frac{K_{Mapp}}{V_{max\,app}} = \frac{K_M}{V_{max} \times K_I} \times [I] + \frac{K_M}{V_{max}}$$

Equation 3: $$\frac{1}{V_{\max\,app}} = \frac{1}{V_{max} \times K_I} \times [I] + \frac{1}{V_{max}}$$

Due to limited resources and poor inhibition, an incomplete IC50 curve was generated for the PGA inhibitor. The full set of Michaelis-Menten curves needed for the kinetic replots to calculate a more accurate inhibition constant could not be obtained due to limited resources and solubility of the inhibitor in millimolar concentrations. An estimated competitive inhibition constant was calculated using the following equation:[74,75]

Equation 4:

$$K_i = \frac{IC50}{1 + \frac{S}{K_M}}$$

3.3 Results & Discussion

One of the first things considered for the *E. coli* PEPCK enzyme (EcPEPCK) was the determination of its basic kinetic parameters in both directions of catalysis. Upon using the standard kinetic assay optimized for the rat enzyme, with the exception of using ATP or ADP nucleotide, there is an immediate preference for the reverse direction that can be observed, instead of the gluconeogenic direction (Table 5). These results are observed under saturating substrate conditions *in vitro*, and likely do not represent how the enzyme operates in a physiological system, where intracellular concentrations of OAA versus PEP result in gluconeogenic activity in the form of PEP synthesis. Understandably, the apparent K_M for OAA is significantly lower than that of PEP, demonstrating a preference for the forward direction substrate, which likely translates well *in vivo*. It should be noted that the K_M for ATP could not be accurately determined because pyruvate kinase, the subsequent enzyme in the coupled-enzyme assay, uses the ADP nucleotide product as its substrate. This is not the case in the reverse direction assays where ADP itself is being monitored, as only MDH is required for the coupled-enzyme assay in this direction, and MDH is not nucleotide-dependent like pyruvate kinase.

Table 5: Comparison of kinetic parameters between WT EcPEPCK, hcPEPCK and rcPEPCK

a) OAA + GTP → PEP + GDP + CO_2						
	K_M (μM)		**k_{cat} (s^{-1})**		**k_{cat}/K_M ($M^{-1} s^{-1}$)**	
	OAA	NTP	OAA	NTP	OAA	NTP
E. coli	204.5 ± 27.5	*	36.4 ± 1.9	*	1.8 x 10^5	*
Human	33.8 ± 4.6	69.8 ± 10.9	33.4 ± 1.0	29.1 ± 0.96	9.9 x 10^5	4.2 x 10^5
Rat[a]	39.6 ± 4.8	54.6 ± 6.0	22.3 ± 0.7	n/a	5.6 x 10^5	4.1 x 10^5
b) PEP + GDP + CO_2 → OAA + GTP						
	K_M (μM)		**k_{cat} (s^{-1})**		**k_{cat}/K_M ($M^{-1} s^{-1}$)**	
	PEP	NDP	PEP	NDP	PEP	NDP
E. coli	789.7 ± 89.3	93.3 ± 11.3	94.7 ± 2.6	92.7 ± 2.3	1.2 x 10^5	9.9 x 10^5
Human	298.0 ± 39.5	29.3 ± 8.9	16.1 ± 0.60	14.1 ± 0.61	5.4 x 10^4	4.8 x 10^5
Rat[a]	206.3 ± 17.2	85.9 ± 9.3	15.0 ± 0.3	n/a	7.3 x 10^4	1.8 x 10^5

**Kinetic parameters for ATP were unable to be determined because the coupled-enzyme assay was incompatible.*
[a]Rat cytosolic data is from ref 14

When compared to GTP-PEPCKs like rat or human cytosolic PEPCK, the EcPEPCK enzyme is significantly more dynamic. Enzyme turnover is higher than that of the studied GTP-PEPCKs, even in the direction of carbon fixation, in which lower enzyme activity is usually observed for GTP-PEPCKs. The apparent K_M for the different substrates of EcPEPCK is higher than their GTP-PEPCK counterparts, despite demonstrating greater enzyme turnover. However, due to the E. coli enzyme having the higher turnover (k_{cat}), EcPEPCK has a catalytic efficiency (k_{cat}/K_M) on par to that of the GTP-dependent PEPCK enzymes. The mechanism that allows for greater enzyme turnover must involve the transition between open and closed conformations in some way. Analysis of the Ω-loop lid latching mechanism reveals fewer inter-residue contacts than when compared to the GTP-dependent rat enzyme (Figure 28).

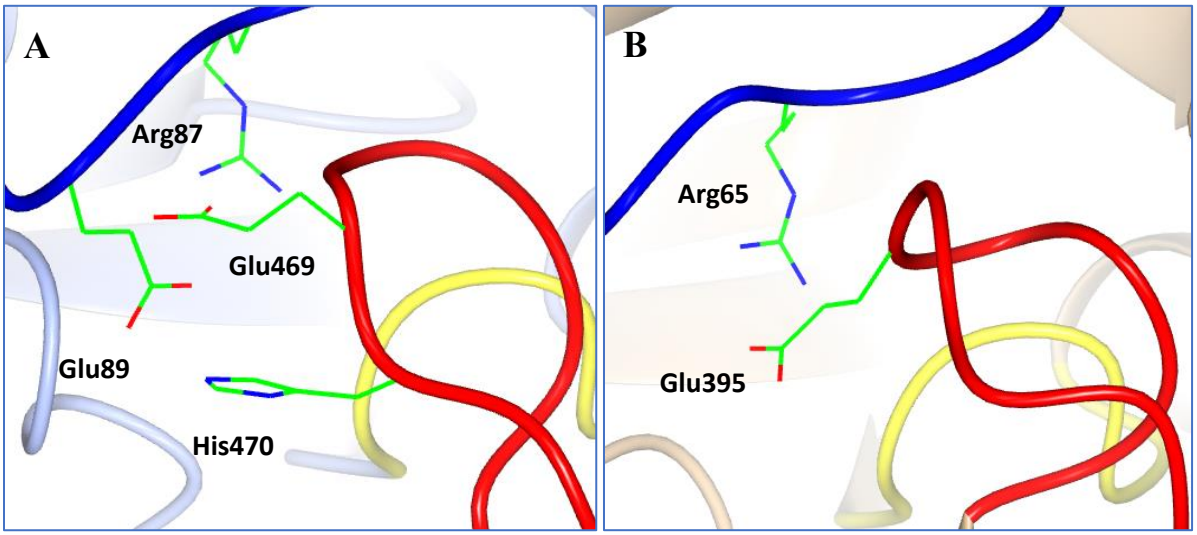


Figure 28: Comparisons of the Ω-loop lid mechanism in human and *E. coli* PEPCK
A) Structure of hcPEPCK and **B**) structure of EcPEPCK. The R-loop is coloured blue, the P-loop is coloured yellow, and the omega loop is coloured red. Amino acid side chains that make up the latch mechanism are shown in green. This diagram was generated using CCP4MG.[26]

This could suggest that EcPEPCK is less stable in the closed conformation when compared to GTP-PEPCKs such as hcPEPCK. To test this hypothesis, non-productive production of pyruvate was assayed, that is the non-enzymatic conversion of the reaction intermediate (enolate of pyruvate) to pyruvate caused by premature omega loop lid opening (Table 6). Interestingly, EcPEPCK demonstrates non-productive pyruvate production to a similar extent as the rat cytosolic enzyme, but in the opposite direction of catalysis.[24,50] This follows the same trend seen for EcPEPCK product turnover, in that the PEP → OAA direction is preferred, but does not explain why product turnover is higher than GTP-PEPCK.

Table 6: Non-productive production of pyruvate by PEPCK enzymes

a) Pyruvate production in the direction of OAA → PEP		
	K_M (μM)	**kcat (s^{-1})**
EcPEPCK	n.d.	n.d.
**rcPEPCK*	53 ± 13	0.40 ± 0.02
b) Pyruvate production in the direction of PEP → OAA		
	K_M (μM)	**kcat (s^{-1})**
EcPEPCK	356 ± 113	1.0 ± 0.1
**rcPEPCK*	n.d.	n.d.

*n.d. – no pyruvate production detected. *rcPEPCK data is from ref 25*

EcPEPCK also uses a two-metal system similar to rat cytosolic PEPCK (rcPEPCK), in which two *different* divalent cations show synergistic behaviour for optimal enzyme activity.[88] The M1 metal binds first to the apoenzyme with a characteristic octahedral coordination geometry, followed by the M2 metal in the form of a metal-nucleotide complex.[24] Although initial research on ATP-PEPCK in Bakers' yeast states that Mn^{2+} is the M1 metal, more recent structural studies on the *E. coli* enzyme suggest Mn^{2+} and Ca^{2+} are interchangeable, since they both occupy the active site where the M1 metal binds.[89] Prior research on GTP-dependent rcPEPCK shows Mn^{2+}

is the preferred M1 metal, whereas Ca^{2+} shows no significant activation of enzyme activity.[14] A full characterization of EcPEPCK was performed using several physiologically relevant divalent cations: manganese, iron, calcium, zinc, cobalt and nickel. These metals were present in micromolar quantities in the form of chloride salts, either as the sole metal or in the presence of saturating $MgCl_2$. Mg^{2+} was chosen as the proxy for the M2 metal due to its physiological relevance and high abundance inside the cell.[90]

EcPEPCK shows the same preference as hcPEPCK for Mn^{2+} as its M1 metal, followed by Fe^{2+} and Mg^{2+} to a lesser extent (Table 7). This is evident when only Mn^{2+} or Fe^{2+} are used in the kinetic assays, as the enzyme is still active even when these two cations serve as both the M1 and M2 metals. The other divalent cations, including Ca^{2+}, show no significant activation as the sole M1 metal, and appear to have an antagonistic effect when in the presence of saturating $MgCl_2$. They likely outcompete Mg^{2+} for the M1 metal, lowering the enzyme activity below the rate contributed by saturating $MgCl_2$ alone. This is problematic as previously published structural data for the EcPEPCK enzyme use Ca^{2+} as the ligand[89] despite the kinetic data demonstrating that Ca^{2+} does not have an activating effect as the M1 metal (Table 7). Interestingly, Fe^{2+}, the divalent cation with the second highest affinity for the M1 site, also happens to be the preferred M1 metal for the lesser-known family of PP_i-PEPCKs.[31] It should also be noted that in the original study on Bakers' yeast, Fe^{2+} was never tested for its activation of the ATP-dependent PEPCK enzyme.[39]

Table 7: Kinetic characterization of different divalent cations against PEP to determine the highest activating M1 metal of EcPEPCK.

Two-metal divalent cation activation (100 µM M1 + 4 mM Mg)		
	K_M (µM)	**Specific Activity (µmol/min/mg)**
Mn^{2+}	1.1 ± 0.2	56.1 ± 1.9
Fe^{2+}	3.0 ± 0.6	40.4 ± 1.7
Ca^{2+}	-	18.0 ± 5.6
Co^{2+}	-	13.0 ± 2.1
Zn^{2+}	-	9.6 ± 1.9
Ni^{2+}	-	6.4 ± 0.6
4 mM Mg^{2+}	-	**20.7 ± 1.99**
One-metal divalent cation activation (100 µM M1 metal only)		
	K_M (µM)	**Specific Activity (µmol/min/mg)**
Mn^{2+}	2.3 ± 0.9	36.0 ± 2.8
Fe^{2+}	9.5 ± 5.8	20.7 ± 3.8
Mg^{2+}	2.5 ± 0.7	4.8 ± 0.3

All cations were added to the assay in the form of chloride salts. 4 mM excess magnesium was added as a proxy for the M2 metal in a two-metal system. K_M could not be determined for Ca, Co, Zn or Ni as they did not activate EcPEPCK.

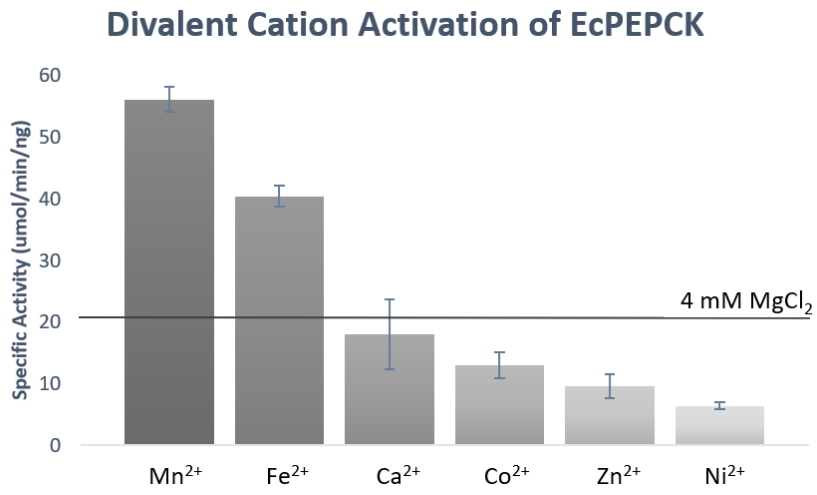


Figure 29: Divalent cation activation of EcPEPCK via a two-metal system
Visual representation of divalent cation activation of EcPEPCK via a two-metal system. Six different divalent metal cations were assayed for their role as the M1 activating metal in the presence of saturating $MgCl_2$ as the M2 metal. The activity attributed to 4mM $MgCl_2$ is denoted as a line on the graph.

This two-metal system is near-identical to that of rcPEPCK, with even the relative affinities for Mn^{2+} followed closely by Fe^{2+} staying the same, whereas Ca^{2+} does not appear to show any significant activation for either EcPEPCK or rcPEPCK. Since both rcPEPCK and EcPEPCK share the exact same active site residues (M1 binding residues: lysine, aspartic acid, histidine), this was not unexpected, despite the low sequence identity between the two classes of PEPCKs.

Thus far the similarities between the ATP- and GTP-dependent enzymes were expected, and due to the shared active site residues between them, inhibition by PEPCK reaction analogs was also expected to show similar effects. Oxalate, and analog of the enolate intermediate, was co-crystallized with EcPEPCK and the ATP nucleotide, and was demonstrated to bind to EcPEPCK. Oxalate binds to the OAA/PEP subsite, and ATP is bound to the nucleotide subsite (Figure 30). This co-crystal structure is an example of the closed conformation of PEPCK, wherein the omega loop lid is closed over to protect the active site during catalysis.

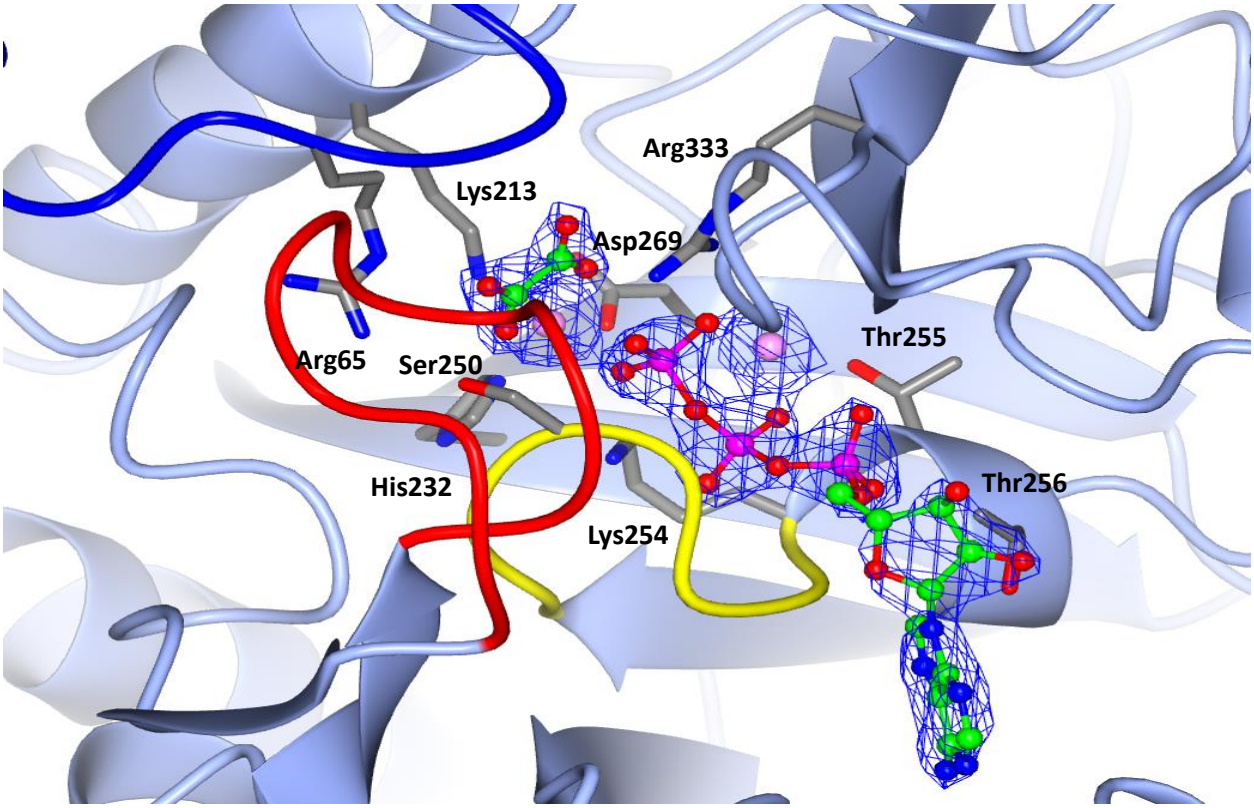


Figure 30: EcPEPCK co-crystal structure with ATP and oxalate
Ribbon structure of EcPEPCK is coloured ice blue. The R-loop is coloured blue, the P-loop is coloured yellow, and the omega loop is coloured red. Both the oxalate ligand and the ATP nucleotide are coloured by atom and rendered as ball-and-stick models. Neighbouring amino acids within 3.5Å are coloured grey. The 2Fo-Fc electron density map surrounding oxalate and ATP is contoured to 2σ. The 2Fo-Fc map for the metals is contoured to 3σ. Manganese ions are coloured light pink and represent the M1 and M2 metals. This diagram was generated using CCP4MG.[26]

However, interesting differences were observed when running the kinetic inhibition assays using oxalate. Oxalate shows stronger inhibition in the reverse direction against PEP, in contrast to rcPEPCK, where oxalate inhibition is more potent against OAA.[76] In addition, oxalate shows mixed inhibition with an equal contribution from both competitive and uncompetitive inhibition in the reverse direction against PEP, as seen by the kinetic replots (Table 8). This would suggest that oxalate binds to two separate sites, one competitively at the OAA/PEP subsite and another at either the nucleotide subsite, or an allosteric site. However, this putative double oxalate-bound crystal structure could not be obtained after several lengthy attempts at crystallization.

Table 8: Inhibition of EcPEPCK by Oxalate (analog of reaction intermediate)

	c) OAA + ATP -> PEP + ADP + CO_2	
	K_M/V_{max} K_I (µM)	**$1/V_{max}$ K_I (µM)**
EcPEPCK	71.7 ± 29.0	358 ± 91.2
rcPEPCK	9 ± 1	n/d
	d) PEP + ADP + CO_2 -> OAA + ATP	
	K_M/V_{max} K_I (µM)	**$1/V_{max}$ K_I (µM)**
EcPEPCK	12.4 ± 3.2	15.5 ± 0.6
rcPEPCK	106 ± 6	n/d

n/d - uncompetitive inhibition constant for rcPEPCK was not determined.

The next inhibitor looked at was β-sulfopyruvate (BSP), an analog of OAA, the substrate in the forward direction. BSP is well-known as an inhibitor of the rat enzyme (GTP-PEPCK).[76] A co-crystal structure of BSP and ATP (Figure 31) was obtained after some difficulty. Initial attempts at co-crystallization yielded a structure where BSP only had partial occupancy, however through a combination of co-crystallization and soaking, full occupancy was observed. Harvesting a crystal grown at 100 mM BSP and soaking in 150 mM BSP for 10 minutes yielded the final EcPEPCK ATP BSP co-crystal structure. Difficulty in obtaining a co-crystal structure with BSP fully occupying the active site suggested that it was not a potent inhibitor of EcPEPCK, in contrast to the GTP-dependent rat cytosolic enzyme. This structure had a highly flexible omega loop lid, making it difficult to model initially, but after several rounds of refinement the conformation of the omega loop was able to be determined. The ATP + BSP co-crystal dataset had the best resolution of the three EcPEPCK structures, refined to 1.45Å, and also had a different spacegroup (C2) from the others ($P2_1$) (Table 11).

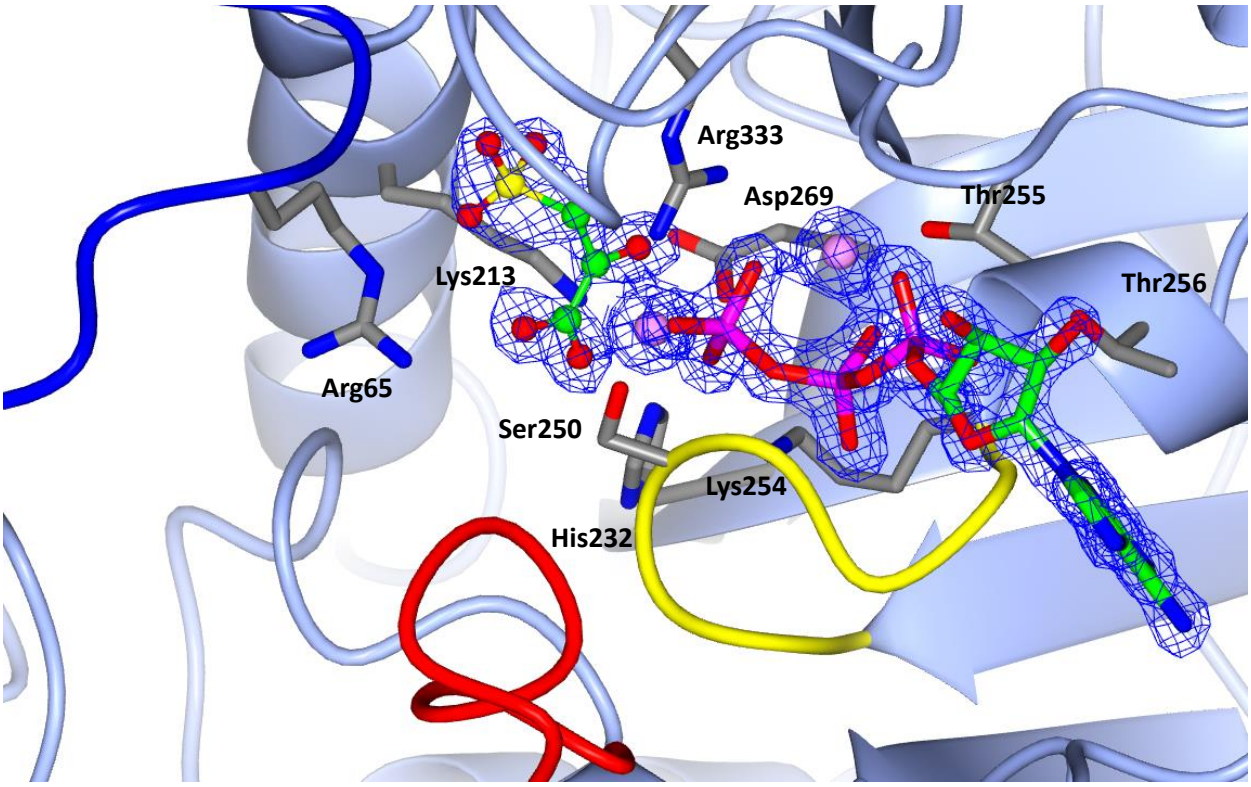


Figure 31: EcPEPCK co-crystal structure with ATP and β-sulfopyruvate
Ribbon structure of EcPEPCK is coloured ice blue. The R-loop is coloured blue, the P-loop is coloured yellow and the omega loop is coloured red. Both the BSP ligand and the ATP nucleotide are coloured by atom and rendered as ball-and-stick models. Neighbouring amino acids within 3.5Å are coloured grey. The 2Fo-Fc electron density map surrounding BSP and ATP is contoured to 1.5 and 2σ, respectively. The 2Fo-Fc map for the metals is contoured to 3.5σ. Manganese ions are coloured light pink and represent the M1 and M2 metals. This diagram was generated using CCP4MG.[26]

Inhibition assays revealed what was expected based on the structural data, that BSP does not bind very tightly to EcPEPCK. BSP is known to be a tightly binding inhibitor of GTP-PEPCK[76] (K_I ~25 μM), however BSP does not appear to bind nearly as tightly to EcPEPCK, with a competitive inhibition constant of about 8 mM (Table 9). This explains why it was difficult to obtain a crystal structure with BSP bound at full occupancy. This inhibitor was assayed against OAA only due to the limitations of the assay – malate dehydrogenase, the coupling enzyme used in the opposite direction of catalysis, is capable of catabolizing BSP as a substrate.[87]

Table 9: Inhibition of EcPEPCK by β-sulfopyruvate (analog of the reaction substrate)

OAA + ATP → PEP + ADP + CO_2	K_M/V_{max} K_I (mM)	$1/V_{max}$ K_I (mM)
EcPEPCK	7.91 ± 1.9	61.5 ± 7.9
**rcPEPCK*	0.025 ± 0.006	n.d.

**rat cytosolic data is from ref 24*

The last of the three analogs used to study enzyme inhibition was phosphoglycolate (PGA). PGA is an analog of the PEPCK reaction product, phosphoenolpyruvate (PEP). Unlike the other two aforementioned co-crystal structures, this PGA and ADP structure (Figure 32) was obtained using a different crystal condition (0.1M sodium acetate pH 4.5), and also required seeding, eventually yielding amorphous protein crystals significantly different in morphology from the previous crystals. In attempting to co-crystallize EcPEPCK with PGA and ADP using the initial crystal condition (0.1M BisTris pH 6.5) PGA did not bind, and instead ATP nucleotide was bound to the active site. After ruling out potential ATP contamination in the crystallization reagents, it was concluded that this is a case of product turn-over, such that EcPEPCK can transfer a phosphate group from phosphoglycolate to ADP, creating ATP (see **Chapter 4**). This explains why the PGA and ADP structure was only obtained using the sodium acetate conditions, since EcPEPCK is likely unable to catalyze the phosphate transfer at a lower pH. Going from an approximately neutral pH to an acidic pH of 4.5 could affect the amino acid residues in the active site that coordinate M1 metal binding, or one of the other residues that are directly involved in catalysis. Previous research on GTP-PEPCK demonstrates the importance of pH dependence for PEPCK catalysis, namely for the deprotonated histidine and lysine sidechains that coordinate the M1 metal.[91] It was determined that the pKa's were 6.5 and 7.0 for histidine and lysine, respectively. Since the M1 coordinating residues are conserved in all PEPCKs, it can be assumed that the corresponding residues in EcPEPCK, His232 and Lys213, have similar pKa values. Therefore, lowering the pH below 6.5 would render EcPEPCK unable to catalyze the reaction.

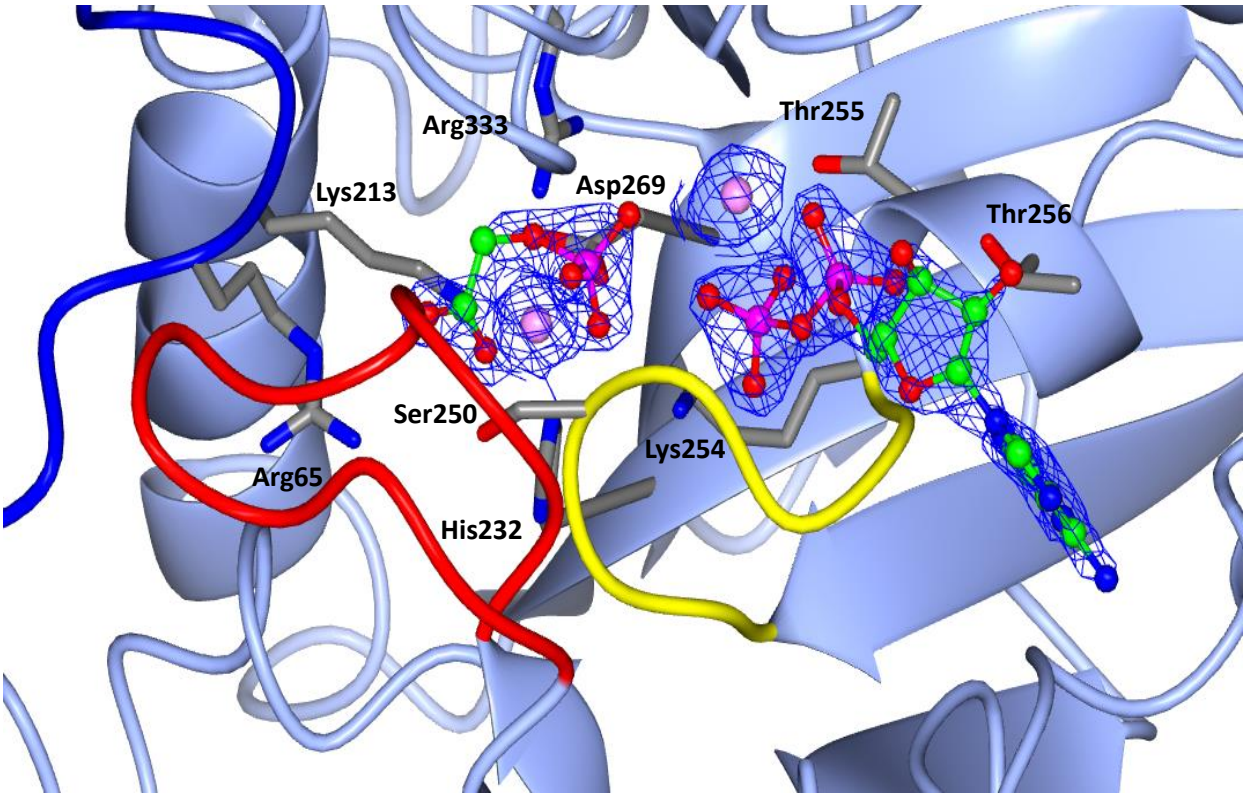


Figure 32: EcPEPCK co-crystal structure with ADP and phosphoglycolate
Ribbon structure of EcPEPCK is coloured ice blue. The R-loop is coloured blue, the P-loop is coloured yellow, and the omega loop is coloured red. Both the PGA ligand and the ADP nucleotide are coloured by atom and rendered as ball-and-stick models. Neighbouring amino acids within 3.5Å are coloured grey. The 2Fo-Fc electron density map surrounding PGA and ADP is contoured to 1.5 and 2σ, respectively. The 2Fo-Fc map for the metals is contoured to 3.5σ. Manganese ions are coloured light pink and represent the M1 and M2 metals. This diagram was generated using CCP4MG.[26]

Determining phosphoglycolate inhibition of EcPEPCK was difficult to achieve with the prior knowledge that PGA is known as a poor inhibitor of rat cytosolic PEPCK (a reported K_I of 2.53 mM)[24], and the potential that EcPEPCK can use it as a substrate. Neither a full set of inhibition curves nor a complete IC50 could be obtained with the resources in our laboratory, however 25% loss of enzyme activity was observed at 5 mM phosphoglycolate. Fitting the available data to a dose-response curve using GraphPadPrism yields an estimated IC50 over 100 mM (Figure 33). Extrapolating from the IC50 value, the K_I for PGA is about 74 mM (see **section 3.2.7**).

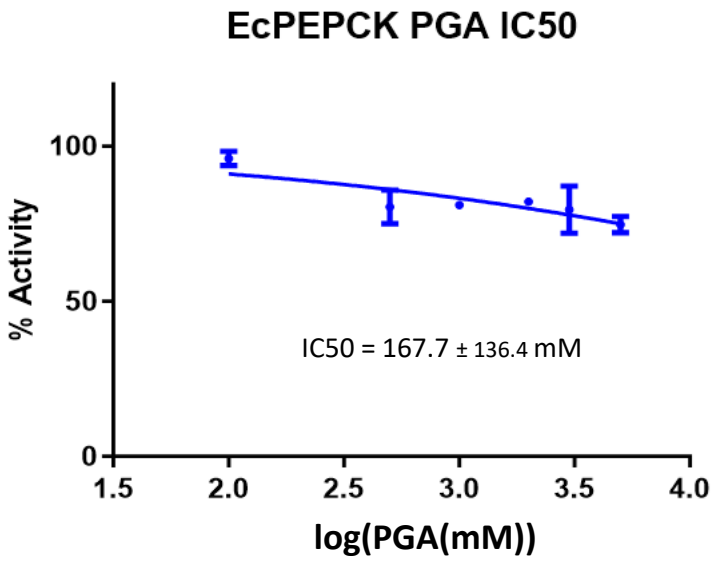


Figure 33: Phosphoglycolate (PGA) IC50 EcPEPCK
EcPEPCK activity data at increasing concentrations of PGA was plotted on GraphPad and fit to a dose-response curve to determine IC50.

Table 10: PGA inhibition for EcPEPCK against PEP

	PEP + ADP + CO_2 → OAA + ATP	
	Calculated IC50 (mM)	**Estimated K_I (mM)**
EcPEPCK	167.7 ± 136.4	~74
**rcPEPCK*	-	2.5 ± 0.01

**rat cytosolic data is from ref 24*

Table 11: Simplified EcPEPCK crystallographic statistics data*

	EcPEPCK ATP Oxalate	**EcPEPCK ATP BSP**	**EcPEPCK ADP PGA**
Space group	P 1 21 1	C1 2 1	P1 21 1
Unit Cell Dimensions	a = 54.236 b = 194.846 c = 54.162 α = γ = 90.00° β = 114.51°	a = 123.998 b = 93.219 c = 46.054 α = γ = 90.00° β = 95.89°	a = 54.527 b = 195.183 c = 54.572 α = γ = 90.00° β = 115.09°
Resolution Limits	39.29-2.00	74.37-1.45	49.47-1.73
Completeness (%)	99.36 (91.77)	98.11 (80.74)	98.22 (93.03)
I/sigma	11.88 (1.68)	8.73 (3.18)	13.84 (1.93)
CC1/2	0.988 (0.758)	0.998 (0.944)	0.991 (0.750)
Rpim	0.060 (0.354)	0.033 (0.154)	0.044 (0.381)
Rmerge	0.137 (0.486)	0.134 (0.243)	0.096 (0.777)
Rfree	0.204 (0.242)	0.188 (0.275)	0.225 (0.424)
Rwork	0.173 (0.190)	0.169 (0.259)	0.187 (0.358)

**Values in parentheses represent the highest resolution shell*

Lastly, with the new EcPEPCK crystal structures obtained, a direct comparison of the nucleotide binding sites of ATP and GTP PEPCKs was achieved (Figure 34). GTP is surrounded by several aromatic residues and the amino group of the Asn533 sidechain is in direct contact (3Å) with the carbonyl of GTP. On the other hand, the ATP binding site is lacking aromatic residues, but has several threonine residues contributing to the coordination of ATP via their hydroxyl side chains.

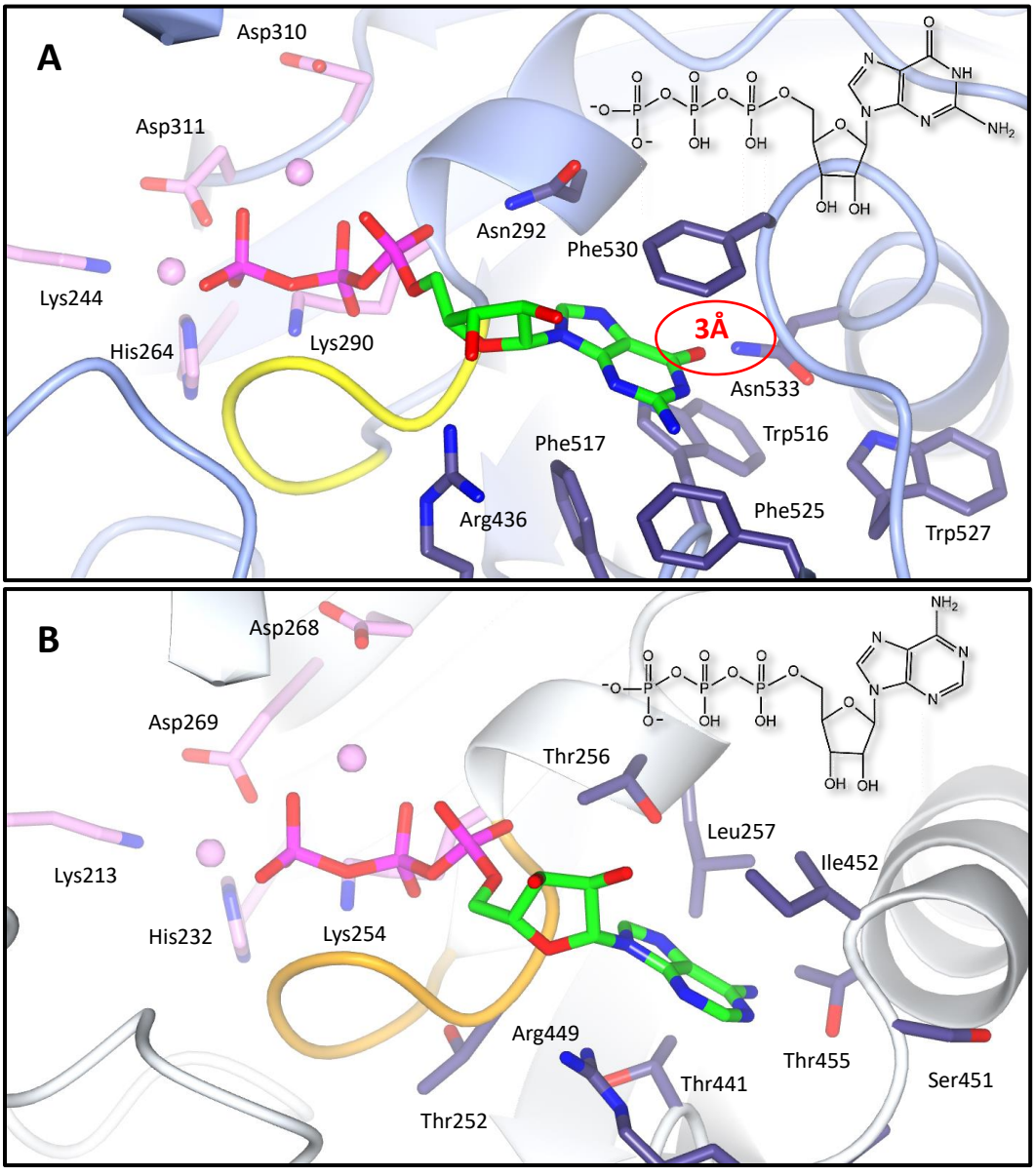


Figure 34: Comparison of the nucleotide binding sites of hcPEPCK and EcPEPCK

Ribbon structures of **A)** hcPEPCK is coloured ice blue and **B)** EcPEPCK is coloured white. The P-loop is coloured yellow and orange for hcPEPCK and EcPEPCK, respectively. The GTP and ATP nucleotides are coloured by atom. Amino acids within 3.5Å of the nitrogenous base are coloured dark purple. Neighbouring amino acids of the metals and phosphate groups are conserved in both structures and are coloured pink. Manganese ions are coloured light pink and represent the M1 and M2 metals. This diagram was generated using CCP4MG.[26] See **Figure 20** and **Figure 30** for the complete active site structures.

3.4 Conclusions

Despite the fact that the first PEPCK crystal structure to be uploaded to the PDB databank was that of *E. coli* enzyme, an ATP-dependent PEPCK, the majority of the kinetic data and characterization of PEPCK enzymes have been done on GTP-dependent PEPCKs. Information gained over the years on the GTP-dependent class of enzymes, including the rat and the human cytosolic enzymes, have been assumed to apply to the family as a whole. Here I present the importance of directly comparing the two nucleotide-dependent classes of PEPCK using the same experimental methods, using the human and *E. coli* enzymes as a proxy for GTP and ATP-PEPCK, respectively.

One of the most apparent observations when comparing the EcPEPCK enzyme to previously studied GTP-dependent PEPCKs, based on the optimized kinetic assay *in vitro*, is that EcPEPCK has a faster rate of product turnover in both directions of catalysis. However, the GTP-PEPCKs have since evolved to have a lower K_M for their substrates in comparison, thus achieving a catalytic efficiency more on par with the ATP-dependent class, despite having slower turnover. It is likely that these enzymes have evolved to account for differences in substrate concentration in their own unique cellular environments. Both classes of PEPCK enzymes share the same absolute requirement for two different divalent cations in their active site. Based on metal concentrations in a biological system, the identity of the M2 metal is Mg^{2+}, in the form of the Mg:ATP complex. The M1 metal can be assumed to be Mn^{2+}, based on the apparent K_M when

compared to that of the other divalent cations tested. Fe^{2+} and Mg^{2+} cations also activate the EcPEPCK enzyme, though not to the same extent as Mn^{2+}, and when present as the sole divalent cations, Mn^{2+}, Fe^{2+} and Mg^{2+} can act as both the M1 and M2 metals simultaneously, although PEPCK activity is greatly diminished. Despite the low sequence identity (~20%), the fact that the two classes of PEPCK enzymes share the same preference for the M1 metal is unsurprising due to conserved metal binding residues in the active site (Lys213, His232, Asp269, and Lys244, His264, Asp311 for ATP and GTP-PEPCKs, respectively).

Well-known inhibitors of GTP-PEPCK include BSP, oxalate and PGA. Likely related to the higher K_M of their analogous substrates, BSP and PGA do not bind as tightly to EcPEPCK when compared to GTP-PEPCKs like the rat and human enzymes. BSP inhibits in the micromolar range for GTP-PEPCK, and the millimolar range for ATP-PEPCK. PGA, as an analog of PEP, is projected to have a K_I of ~74 mM, with structural data suggesting that EcPEPCK is able to use PGA as a substrate. Oxalate, the inhibitor analogous to the reaction intermediate, still proves to be a strong micromolar inhibitor of EcPEPCK. Oxalate demonstrates inhibition in both directions of catalysis, with increased inhibition in the PEP → OAA direction, in direct contrast to rcPEPCK, where inhibition is stronger in the OAA → PEP direction. In addition, for EcPEPCK in the PEP → OAA direction oxalate inhibition can be attributed to both competitive and uncompetitive effects, suggesting a secondary binding site. These differences in inhibition between EcPEPCK and hcPEPCK are despite the fact that all the residues that are directly in contact with the inhibitors/substrates and the bond conformations they adopt are conserved between ATP/GTP-PEPCKs. This result was unanticipated based on the structural homology of the active site shared between these two enzymes, suggesting that there is a difference in the coupling of the amount of free energy required for the conformational changes necessary for

catalytic turnover, as the main active site sequence differences between the two enzymes are in the omega loop lid and hinge regions.

Comparisons of the nucleotide binding sites between hcPEPCK and EcPEPCK, members of the GTP-dependent and ATP-dependent classes of PEPCK, demonstrate a difference in structural selectivity. The ATP binding site in EcPEPCK is composed of polar, electronegative amino acids including threonine to stabilize the amine groups of the adenosine ring. Conversely, the GTP binding site seems to have evolved to incorporate the rarer amino acids with large aromatic side chains like phenylalanine and tryptophan. The presence of multiple aromatic groups likely indicates pi-stacking as a form of coordinating the guanosine ring in the nucleotide binding site, resulting in the K_M of GTP/GDP being lower than that of ATP/ADP. In addition, the extra carbonyl group on the GTP nucleotide is stabilized by the amino group on the asparagine sidechain nearby, which is not present in the EcPEPCK structure.

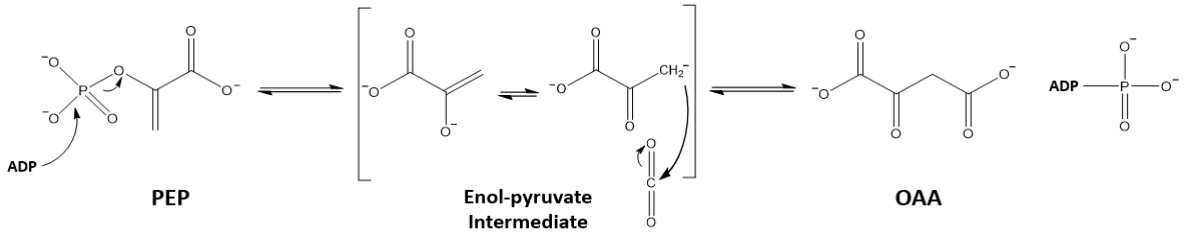
With the assumption that the *E. coli* enzyme is an appropriate proxy for the class of ATP-dependent PEPCKs, the kinetic and structural data summarized here aims to fill in the gaps in the characterization of ATP-PEPCK. This information provides insight into the differences in behaviour between ATP and GTP-PEPCKs that might not have been discovered otherwise, and demonstrates subtle changes in catalytic activity as PEPCK enzymes evolved.

Chapter 4: Using ^{31}P NMR to assay for pyruvate kinase-like activity in *Escherichia coli* PEPCK

4.1 Introduction

Yeast PEPCK, also belonging to the ATP-dependent class of enzymes, has been known to generate pyruvate from PEP, catalyzing a "pyruvate kinase-like" (PK-like) reaction where the phosphate group of PEP is transferred to ADP, resulting in the production of both pyruvate and ATP (Figure 35). This phenomenon was first observed in 1962 in the ATP-PEPCK isolated from Bakers' yeast.[42] Initially, they called the enzyme phosphopyruvate carboxylase, before labelling it by its correct name, phosphoenolpyruvate carboxykinase, a few publications later, as per the Enzyme Commission of the International Union of Biochemistry and Molecular Biology.[39] When characterizing the carbon fixation activity of the yeast PEPCK enzyme, they noted slight ATP production in the absence of CO_2, which was attributed to contaminating pyruvate kinase in the enzyme preparation.[42]

CO_2-fixation reaction[21]



Pyruvate kinase-like reaction[92]

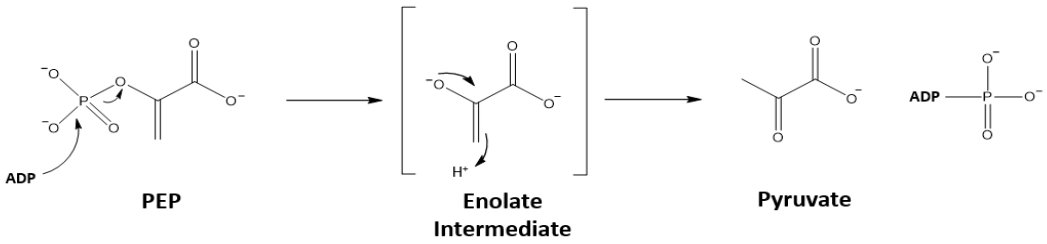


Figure 35: PEPCK-catalyzed reactions using PEP and ADP to generate ATP

Later in 1993, the enzyme from the obligate anaerobe, *Anaerobiospirillum succiniproducens,* was reported to be the first CO_2-fixing PEPCK enzyme derived from a bacterium.[40] This PEPCK enzyme was determined to catalyze the nucleotide-dependent conversion of PEP to OAA as the first step in a fermentation pathway that generates high amounts of succinate, as is characteristic of this microorganism.[40] Shortly thereafter, the pyruvate kinase-like reaction catalyzed by ATP-dependent PEPCKs was characterized in both *Saccharomyces cerevisiae* and *Anaerobiospirillum succiniproducens,* and it was noted that these two enzymes catalyzed the PEPCK reaction in physiologically opposite directions (gluconeogenesis and CO_2 fixation in *S. cerevisiae* and *A. succiniproducens,* respectively).[93] Pyruvate kinase contamination was ruled out by expressing *A. succiniproducens* PEPCK in a PB25 strain of *E. coli,* where the chromosomal *pykA* and *pykF* genes were interrupted by antibiotic resistance markers. It was concluded that PK-like activity of these enzymes was 1.2% and 0.13% of the CO_2-fixation reaction for the yeast and bacterial enzyme, respectively and that the K_M for PEP was similar for both reactions.[93] Since this PK-like reaction has been observed in *S. cerevisiae* and *A. succiniproducens,* both of which have ATP-dependent PEPCK, so it follows that the *E. coli* enzyme would have the potential to demonstrate PK-like activity as well, although thus far it has not been characterized to my knowledge.

Beyond the putative PK-like activity of PEPCK, initial crystallographic data suggests that EcPEPCK may also have the ability to catabolize phosphoglycolate (PGA), as hypothesized in the previous chapter. Early attempts to crystallize the EcPEPCK ADP + PGA structure showed neither of these ligands bound, but rather ATP, suggesting product turnover. The synthesis of the high energy phosphoanhydride bond in ATP from ADP and P_i requires a high amount of energy ($\Delta G° = 30$ kJ/mol), and the hydrolysis of a phosphate group from PEP, the enzyme's typical substrate, is enough to provide this energy ($\Delta G° = -55$ kJ/mol).[94] The use of PGA as a substrate

appears unlikely, as the standard free energy of the hydrolysis of PGA (ΔG° = -31 kJ/mol) is almost half that of the hydrolysis from PEP, and is only barely enough energy to generate the high-energy phosphate bond of ATP.[94] If the structural data demonstrates the production of ATP from PGA and ADP, then the binding of the substrates to EcPEPCK must raise the free energy of hydrolysis for PGA sufficiently high enough to allow for this chemistry to occur. The standard spectrophotometric assay can be used to determine PK-like activity using PEP, but to address the question of EcPEPCK using PGA as a substrate, a new assay was needed. Since this putative activity of EcPEPCK also involves a phosphoryl transfer, ^{31}P NMR was used to design a novel qualitative PEPCK assay.

4.2 Materials & Methods

4.2.1 Materials

All materials needed for EcPEPCK protein expression, purification, kinetic analysis and crystallization were as previously described (**Chapter 3.2.1**).

4.2.2 Enzymes

The coupling enzymes pyruvate kinase (PK), lactate dehydrogenase (LDH) and malate dehydrogenase (MDH) were purchased from Calzyme Laboratories. The plasmid expressing His6-SUMO protease was a gift from C. Lima (Sloan-Kettering Institute, New York, NY) and was expressed and purified as previously described.[24]

4.2.3 EcPEPCK protein expression and purification

The EcPEPCK protein was recombinantly expressed as a His6-SUMO fusion protein and purified as previously described (**Chapter 3.2.5**).

4.2.4 EcPEPCK kinetic assays

The PK-like reaction was assayed using the same conditions as the PEPCK assay in the PEP → OAA direction, with the exception that $KHCO_3$ was left out of the reaction mixture and LDH was used as the coupling enzyme instead of MDH to indirectly monitor pyruvate formation. In addition, a stoichiometric ratio of 4 mM ADP nucleotide to 4 mM $MgCl_2$ was used for both the singular metal assay and the two-metal assay (in which 100 μM $MnCl_2$ was added).

4.2.5 HPLC determination of ADP purity

To determine ADP purity, a 100 mM stock of adenosine-5-diphosphate disodium dihydrate sourced from BioShop Canada Inc (Cat No. ADP006) was made in 1M HEPES pH 7.5 buffer. A diluted 2 mM sample of ADP in 100 mM HEPES pH 7.5 was added to a Higgins Analytical Targa C18 5uM column (TS-2510-C185) with a flow rate of 2 mL/min. The separation methods by Stocchi et al. were followed with a few alterations.[95–97] The first buffer (15 mM ammonium acetate pH 6) was used for 8 minutes, followed by a linear gradient up to 100% of the second buffer (15 mM ammonium acetate pH 6, 10% methanol) for 18 minutes. The gradient was returned to the first buffer for 2 minutes, and continued for 12 minutes eluting in the first buffer. The same protocol was repeated with a second sample containing 2 mM ADP and 500 μM ATP.

4.2.6 EcPEPCK NMR assay design

^{31}P-NMR assays were performed on the Bruker Ascend 300 MHz spectrometer at 25°C. The spectrometer was set for 1D pulse sequence ^{1}H decoupling with 30° flip angle (default waltz16 sequence). Relaxation or recycle delay (D1) was set to 2 seconds between scans. The number of scans collected for the control samples was 16, while the total scans collected over a 2-hour period for the experimental assay spectra was 2458. Phosphoric acid (85% in D_2O) was used as the reference standard.

A 1 mL sample was created for the NMR assay, then 700 µl of that sample was added to a 500MHz NMR sample tube. Standards were created of the substrates and expected products in the presence of reaction buffer, DTT and $MgCl_2$ and D_2O for locking purposes (see **Table 12**). Both samples and standards were chelated before the addition of magnesium. Samples and standards were incubated at room temperature and kept within a locked drawer to prevent exposure to light before being analyzed on the NMR spectrometer for subsequent timepoints.

Table 12: NMR sample composition for standards and reaction assays*

P_i control	**ATP control**	**AMP control**	**PEP ADP control**	**PGA ADP control**	**PGA reaction**	**'PK-like' reaction**
25mM HEPES	25mM HEPES	25mM HEPES	25mM HEPES	25mM HEPES	25mM HEPES	25mM HEPES
1mM Na_3PO_4	1mM ATP	1mM AMP	1mM PEP	1mM PGA	1mM PGA	1mM PEP
4mM $MgCl_2$	4mM $MgCl_2$	4mM MgCl2	1mM ADP	1mM ADP	1mM ADP	1mM ADP
10mM DTT	10mM DTT	10mM DTT	4mM $MgCl_2$	4mM $MgCl_2$	4mM $MgCl_2$	4mM $MgCl_2$
			10mM DTT	10mM DTT	10mM DTT	10mM DTT
					1mg EcPEPCK	**100ug EcPEPCK**

**The pH of the HEPES buffer for all samples was 7.5*

4.3 Results & Discussion

To investigate whether the *E. coli* PEPCK enzyme could catalyze the PK-like reaction, the standard kinetic assay was setup in the reverse direction, omitting the bicarbonate used as a source of CO_2 to ensure the canonical PEPCK-catalyzed reaction would not occur. There was minimal activity detected, indicating that EcPEPCK can catalyze this PK-like reaction in the absence of CO_2, albeit to a much lesser extent than the canonical PEPCK-catalyzed reaction.

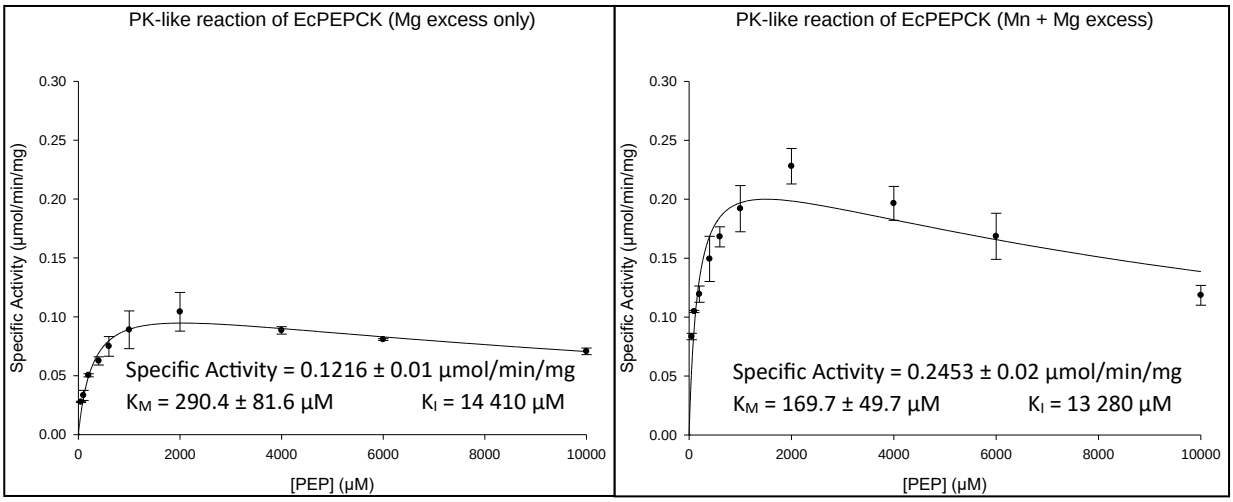


Figure 36: Michaelis-Menten curves showing pyruvate formation by EcPEPCK
Pk-like activity of EcPEPCK using Mg^{2+} (left) and Mn^{2+} + Mg^{2+} (right) both fit to a substrate inhibition model.

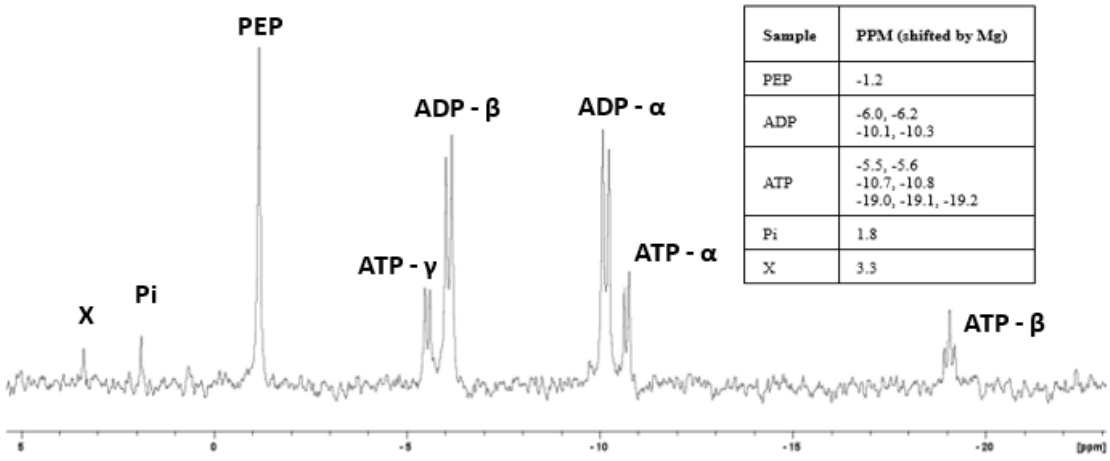
The PK-like reaction was observed in the presence of excess $MgCl_2$, and to a somewhat greater extent by adding 100 µM $MnCl_2$ (Figure 36). The presence of Mn^{2+} increases enzyme activity by a factor of two, and results in a lower K_M for the PEP substrate (Table 13). Although activity is minimal compared to that of PEPCK's true carbon fixation activity, there is evidence that the PK-like reaction is occurring, and that the two divalent metal system is still required for optimal activity. While it is interesting that apparent substrate inhibition was observed when PEP was added to the assay at concentrations greater than 2 mM, the origins of this phenomenon were not further investigated in this study.

Table 13: Pyruvate kinase-like reaction catalyzed by EcPEPCK

	PEP + ADP → Pyruvate + ATP		
	K_M (μM)	**k_{cat} (s^{-1})**	**k_{cat}/K_M ($M^{-1} s^{-1}$)**
Mg^{2+} only	290.4 ± 81.6	0.1209 ± 0.013	4.2×10^2
Mg^{2+} + Mn^{2+}	169.7 ± 49.7	0.2438 ± 0.024	1.4×10^3

Based on the catalytic efficiency, having both Mg^{2+} and Mn^{2+} present results in activity that is an order of magnitude more efficient than what is observed with just Mg^{2+} present. The PK-like reaction is understandably two orders of magnitude less than that of the carbon fixation reaction normally catalyzed by PEPCK. With the kinetic data obtained thus far, pyruvate formation can be observed, but whether this is due to true PK-like activity, or simply the hydrolysis of PEP and generation of P_i, remains to be seen. To give further credence to this putative PK-like reaction, I designed a qualitative assay using phosphorus NMR (^{31}P NMR) to analyze the production of ATP from PEP and ADP. To more easily visualize the PK-like reaction occurring, 100ug of EcPEPCK was used for the NMR assay compared to the 2.5 ug used for the standard spectrophotometric assay. Due to the paramagnetic nature of manganese, Mg^{2+} was used as the sole divalent metal cation in the PK-like reaction.

After a 4-hour incubation of the assay mixture at room temperature, the doublet and triplet peaks corresponding to the γ, α and β ATP phosphates are observed at -5.5/-5.6, -10.7/-10.8 and -19.0/-19.1/-19.2 ppm, respectively (Figure 37). A small peak corresponding to inorganic phosphate can be seen at 1.85 ppm, indicating that EcPEPCK is also hydrolyzing PEP as a side reaction. Control reactions were set up in the absence of enzyme to ensure that there was no product degradation over the same assay timepoints (see **Appendix G**).



Sample	PPM (shifted by Mg)
PEP	-1.2
ADP	-6.0, -6.2 -10.1, -10.3
ATP	-5.5, -5.6 -10.7, -10.8 -19.0, -19.1, -19.2
Pi	1.8
X	3.3

Figure 37: ^{31}P NMR spectra for the PK-like reaction catalyzed by EcPEPCK after 4 hours
^{31}P NMR spectrum for the PK-like reaction after 4 hours of incubating the reaction mix with 100 ug EcPEPCK. The 300 MHz NMR spectrometer was run for 2 hours to obtain this signal-to-noise ratio.

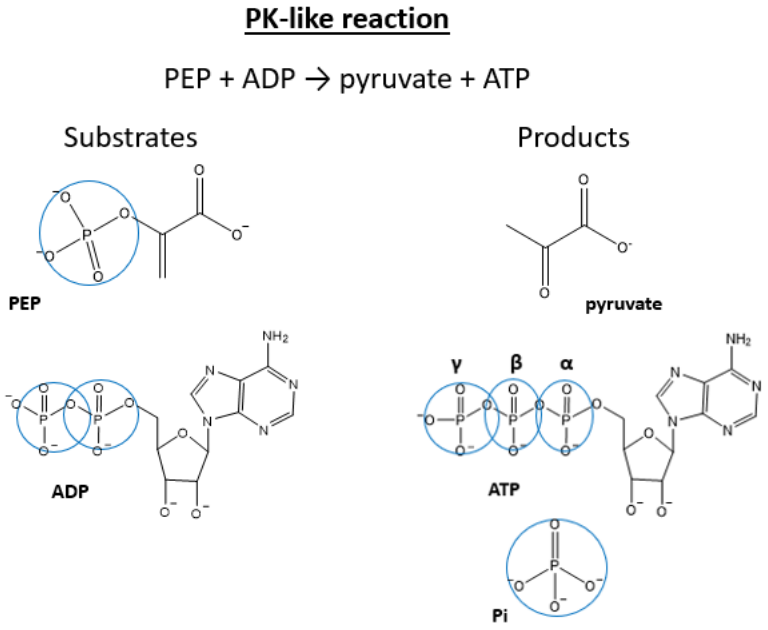


Figure 38: Substrates and products of the PK-like reaction and their visibility in ^{31}P NMR
Chemical structures of the substrates and products involved in the PK-like reaction. Phosphate groups highlighted in blue represent the peaks observed in the ^{31}P-NMR spectra.

The PK-like reaction has a much lower turnover rate and catalytic efficiency when compared to the canonical carbon fixation reaction of PEPCK, even though both reactions use PEP and ADP as substrates. It was reported that the PK-like reaction is 1.2% and 0.13% of the carbon fixation reaction for yeast and *A. succiniproducens* PEPCK, respectively[93]. and the PK-like reaction of EcPEPCK is on par with its PK-like reaction roughly 0.25% of its normal activity. PK-like activity in EcPEPCK at roughly 0.25% of its normal activity puts it more on par with *A. succiniproducens*, an anaerobic bacterium. To support this conclusion, the ^{31}P-NMR assay demonstrates the presence of ATP nucleotide that is produced after the phosphoryl transfer. There also appears to a side reaction of PEP hydrolysis, resulting in an accumulation of inorganic phosphate.

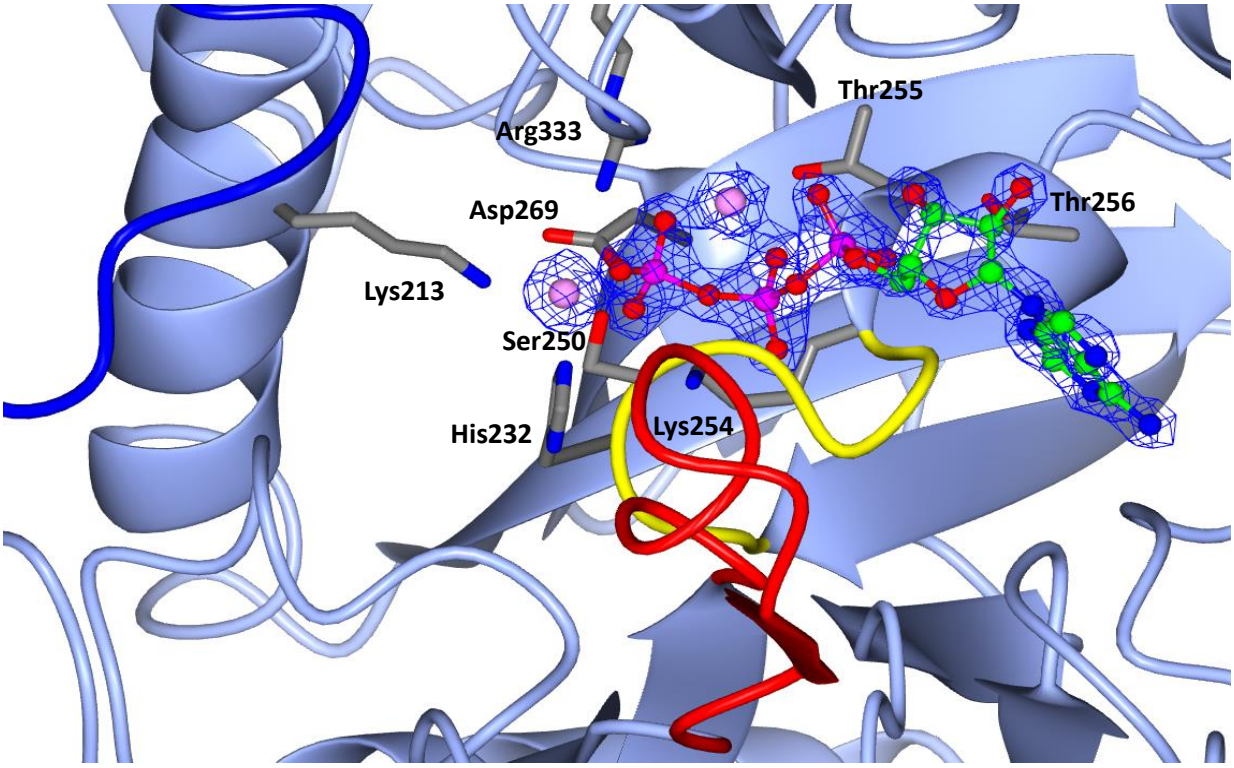


Figure 39: Ribbon structure of the EcPEPCK-ATP cocrystal complex generated when EcPEPCK is crystallized under ADP and PGA crystallographic conditions at neutral pH

Ribbon structure of EcPEPCK is coloured ice blue. The R-loop is coloured blue, the P-loop is coloured yellow, and the omega loop is coloured red. The ATP nucleotide is coloured by atom-type and rendered as ball-and-stick models. Neighbouring amino acids within 3.5Å are coloured grey. The 2Fo-Fc electron density map surrounding ATP is contoured to 2σ. The 2Fo-Fc map for the metals is contoured to 3.5σ. Manganese ions are coloured light pink and represent the M1 and M2 metals. This diagram was generated using CCP4MG.[26]

Initial attempts to obtain a co-crystal structure of EcPEPCK with PGA + ADP (see **section 3.3**) under crystal conditions approaching a neutral pH (pH 6.5) yielded an interesting phenomenon of product turnover, where ATP was found bound to the EcPEPCK active site (Figure 39). Utilizing the same NMR-based assay, I also endeavoured to test the hypothesis that EcPEPCK is able to use the phosphate group from phosphoglycolate (PGA) to make ATP, as observed in the structural data. This phosphorus-based assay would determine if EcPEPCK was simply hydrolyzing PGA into glycolate and inorganic phosphate, or if it was indeed transferring the phosphate group onto ADP. Due to the fact that PGA is not the preferred substrate of EcPEPCK, this reaction took place quite slowly, with only minimal results observed after 24 hours. Small peaks with low signals can be seen at the expected chemical shifts for ATP (two pairs of doublets on either side of the ADP peaks, and one triplet peak further downfield). After 1 week of incubation at RT with control samples running in parallel, as the ADP doublets decreased in signal, ATP peaks were observed at -5.5, -10.7 and -19.1 ppm (Figure 40). This supports EcPEPCK being able to use PGA instead of PEP in the phosphoryl transfer reaction, forming ATP and presumably glycolate. This also likely corresponds to why PGA is an even worse inhibitor of the *E. coli* enzyme when compared to rcPEPCK, however the bigger detriment to PGA as an inhibitor of EcPEPCK is the ability of the enzyme to hydrolyze the phosphate group off of PGA, making inorganic phosphate and glycolate, as seen in the large P_i peak in the spectra.

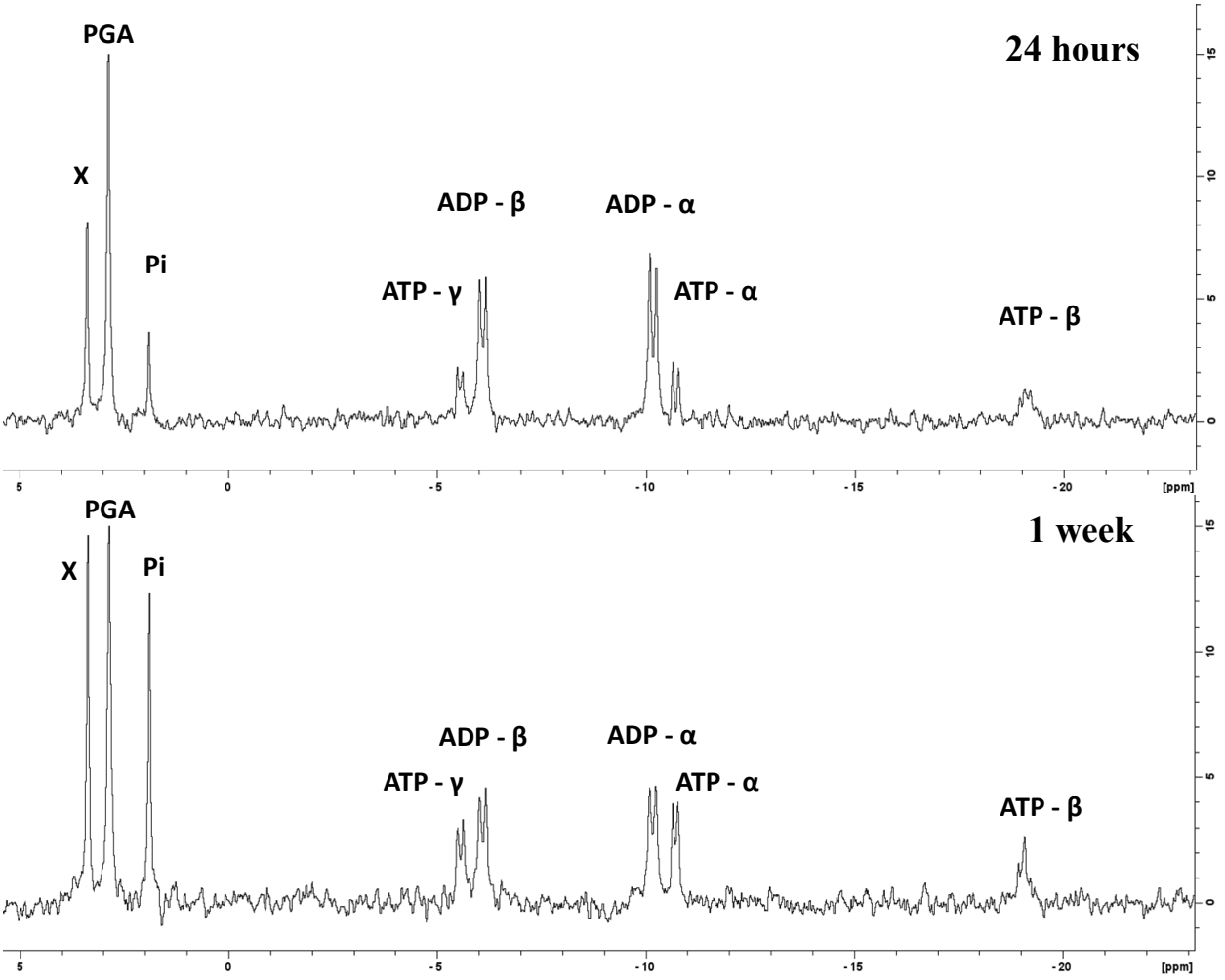


Figure 40: ^{31}P NMR spectra for the catalysis of phosphoglycolate by EcPEPCK

^{31}P NMR spectrum for the catalysis of phosphoglycolate after 24 hours (top) and 1 week (bottom) of incubating the reaction mix with 1mg EcPEPCK. The 300 MHz NMR spectrometer was run for 2 hours to obtain this signal-to-noise ratio.

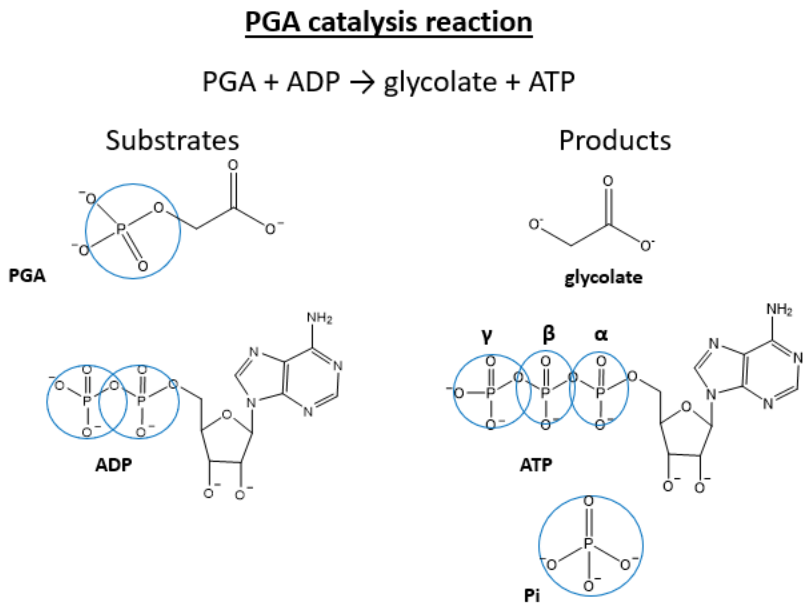


Figure 41: Substrates and products of the PGA catalysis reaction and their visibility in ^{31}P NMR
Chemical structures of the substrates and products involved in the catalysis of PGA. Phosphate groups highlighted in blue represent the peaks observed in the ^{31}P-NMR spectra.

However, in addition to the formation of ATP, there is also a large singlet peak corresponding to inorganic phosphate (P_i) that can be seen after 24 hours and also increases in signal over a week time period. There is also an unknown peak around 3.5 ppm that also increases in signal over the time course of the assay. This peak labelled 'X' increases faster than that of inorganic phosphate, and also appears in the aforementioned PK-like reaction. Neither the singlet for inorganic phosphate or peak X appeared for the non-enzymatic control sample after 1 week. (**Figure G6**). Both the P_i and unknown singlet peaks increase at a faster rate than the peaks corresponding to ATP. This peak could not correspond to an intermediate, as it would eventually decrease as the reaction continued over the week-long time course. Based on the chemical shift (δ = 1 - 4 ppm for AMP depending on pH and buffer composition)[98–100], it is likely that this peak, that appears in both the PK-like reaction and the catalysis of PGA, corresponds to the production of AMP as another side reaction. The mechanism is unclear, but it appears that EcPEPCK might be cleaving

a phosphate group from the ADP nucleotide. This would also generate P_i on top of what is generated from the incomplete transfer of phosphate from PGA to ATP, explaining why the P_i peak at 1.8 ppm increases so rapidly. The adenosine 5'-monophosphate (AMP) control (see **Appendix G3**) that was run using the same buffer and metal conditions shows a peak at 3.46 ppm, further providing evidence that the identity of the unknown peak in both the PK-like reaction and the catalysis of PGA is likely AMP.

4.4 Conclusions

The combination of the kinetic assay for the production of pyruvate and the ^{31}P-NMR assay for the production of ATP provides sufficient evidence that EcPEPCK is capable of catalyzing a pyruvate kinase-like (or PK-like) reaction. Although hinted at via the activities of the ATP-dependent *S. cerevisiae* and *A. succiniproducens* PEPCK enzymes, this is the first time that PK-like activity of the *E. coli* PEPCK has been observed and characterized. The data presented here suggests that PK-like activity is an inherent property of ATP-dependent PEPCKs. The question as to whether it can be applied to the entire family of PEPCK enzymes remains unclear. To our knowledge the PK-like dephosphorylation of PEP to generate pyruvate has not been observed in GTP-PEPCK, suggesting that this activity is a distinguishing feature between ATP and GTP-PEPCK enzymes. Interestingly, the lesser understood PP_i-PEPCKs are also able to catalyze a PK-like reaction, albeit using P_i and forming PP_i instead.[31] This could reinforce the idea that PP_i-PEPCKs indeed arose from a common ancestor, and perhaps existed as the earliest form of PEPCK before adapting and gaining the ability to use ATP nucleotide instead of PP_i. Conversely, evolving the ability to use GTP as a phosphoryl donor and getting farther away from the mechanistic similarities to pyruvate kinase could explain why GTP-dependent PEPCKs no longer have the ability to catalyze a PK-like reaction.

The secondary phenomenon of phosphoglycolate (PGA) catalysis is slightly more ambiguous. Initial crystallographic data of what was supposed to be a co-crystal structure of ADP and PGA yielded only ATP bound to the EcPEPCK active site. After ruling out potential contamination, it could only be concluded that this was a case of enzyme turnover, where PGA was being used as an alternate substrate. The likely scenario is that at the near neutral pH conditions of the crystal buffer, EcPEPCK was able to catalyze the transfer of the phosphate group from PGA to ADP, thereby producing ATP and what was presumably glycolate. The NMR assay was able to verify what was seen from the structural data, showing evidence of doublet and triplet peaks corresponding to the ATP product. The ability of EcPEPCK to use PGA as a substrate can be assumed to be a property of all ATP-PEPCKs. Revisiting the similarities of ATP-PEPCK and pyruvate kinase reveals that this use of PGA as an alternate substrate is not surprising since pyruvate kinase is also known to catalyze the reverse of this reaction, that is the phosphorylation of glycolate via ATP. [101] Interestingly, like pyruvate kinase, there is some evidence that GTP-PEPCKs can also catalyze what we assume is the reverse of the PGA catalysis reaction, that is the phosphorylation of glycolate (and other α-substituted carboxylic acids) via GTP.[102] Whether GTP-PEPCK can also catalyze the reverse reaction and use PGA as a substrate was not investigated, as the structural data already published (e.g. PDB ID: 3DTB) from the rat cytosolic enzyme demonstrates we can capture the GDP + PGA complex at neutral pH.[103,104]

Chapter 5: Future Directions for Phosphoenolpyruvate carboxykinases

5.1 GTP-dependent hcPEPCK's role as a protein kinase

To properly study the role of hcPEPCK as a protein kinase, obtaining a source of AKT1 enzyme would be ideal. Previous research used AKT1 and PEPCK in an *in vitro* phosphorylation assay along with LC/MS-MS to demonstrate that AKT1 phosphorylated PEPCK at Ser90.[46] AKT1 could be purchased from a commercial source, then used to obtain a source of phosphorylated PEPCK that could be used in downstream kinetic assays for characterization or INSIG binding. Alternatively, if AKT1 was recombinantly expressed as a tagged fusion protein, then it could be isolated and removed, yielding a pure sample of phosphorylated PEPCK that could potentially be used for crystallography. Conversely, the sequence of the R-loop that Ser90 is located on could be used to synthesize a short peptide derivative of PEPCK that would be useful in analyzing how AKT1 binds to PEPCK. A more ambitious study would involve attempting to obtain a co-crystal structure of AKT1 and PEPCK at the point of phosphorylation.

Due to the lack of structural information on the PEPCK-INSIG interaction, further research needs to be done to provide more insight. A crystal structure of hcPEPCK with the small INSIG peptide bound was unable to be obtained. Both the wildtype and S90E enzymes were only able to grow in the presence of ligands to stabilize the closed conformation of the active site. If phosphorylation of INSIG is dependent upon GTP, then future attempts should include GTP or GDP nucleotides in the crystal buffer. In addition, the longer 15-mer INSIG peptide used in the initial kinetic study should also be tried, since the extra amino acids might be necessary for binding. Beyond those initial steps, further crystallographic conditions should be screened for PEPCK in the presence of INSIG peptide. Unfortunately, the full length INSIG protein might be necessary, as some important protein-protein interactions might be missing otherwise. This of

course will lead to some difficulty in crystallizing the full PEPCK-INSIG co-crystal structure, on top of the fact that INSIG in a transmembrane protein, which will require additional measures to crystallize.

Further research into the potential moonlighting abilities of PEPCK is definitely needed to elucidate the question of how hcPEPCK is involved in lipogenesis, and its implications in cancer cells. Preliminary research already suggests that PEPCK phosphorylates INSIG in HCC cancer cells, but a complete characterization of this moonlighting activity is still required. The research summarized here demonstrates how the phosphorylation of hcPEPCK does not significantly affect canonical gluconeogenic activity, but rather the upregulation of lipogenic pathway can be attributed to the compartmentalization and translocation of PEPCK to the ER. This change in subcellular location where the substrates of the gluconeogenic reaction are not available allows for hcPEPCK to act as a protein kinase. It is already a well-known concept that PEPCK expression is upregulated in tumour cells for cell proliferation, so having a role in lipogenesis would logically not be much of a stretch. PEPCK is a ubiquitous enzyme that lies at the very center of carbon metabolism, and this would be yet another role for GTP-PEPCK.

5.2 Comparison of ATP-dependent and GTP-dependent PEPCK enzymes

When PEPCK was first discovered, it was not clear from the early kinetic data that the enzymes that had different nucleotide preferences belonged to two separate classes and were indeed different enzymes. From an evolutionary standpoint, it can be summarized that in general, bacteria have ATP-dependent PEPCK and eukaryotes have GTP-dependent PEPCK enzymes, but there are countless exceptions. An obvious exception is the ATP-dependent PEPCK from

Saccharomyces cerevisiae or Bakers' yeast, a single celled eukaryote. The prevailing hypothesis is that ATP-dependent and GTP-dependent PEPCKs are an example of divergent evolution, and that GTP-PEPCKs likely evolved from ATP-PEPCKs. This can be seen in their similar global folds, and the structural homology of the PEPCK active site. The two classes of PEPCK enzymes have a relatively low sequence identity, yet the integral residues that are directly involved in catalysis are entirely conserved.

Using *E. coli* PEPCK as a proxy for the ATP-dependent class of PEPCK enzymes and human cytosolic PEPCK as a proxy for the GTP-dependent enzymes is a good first step. Both of these two PEPCKs are monomeric *in vivo*, and they both preferentially catalyze the gluconeogenic direction, as is reflected by their substrate and nucleotide affinities. Unfortunately, the standard spectrophotometric assay using NADH can not be adapted perfectly to ATP-dependent PEPCKs, due to the coupling enzyme in the forward direction (pyruvate kinase) requiring both PEP and ADP, which are both products of the PEPCK-catalyzed reaction. This makes the PK-catalyzed reaction the rate-determining step instead of PEPCK. I've attempted to overcome this problem by using ten times the amount of PK in the EcPEPCK kinetic assays, and this has allowed me to shed some light onto the gluconeogenic activity of EcPEPCK while understanding that the kinetic parameters in the forward direction may not represent the intrinsic values. Finding another assay more compatible with the EcPEPCK catalyzed reaction would be ideal for analyzing parameters in the OAA → PEP direction.

From the kinetic assays, one of the initial observations made is that the EcPEPCK (ATP-dependent) enzyme has a higher turnover in both directions of catalysis than that of the hcPEPCK enzyme (GTP-dependent). Interestingly, this higher activity is in spite of having lower

apparent affinity for its substrates and ADP. Considering the main amino acids in the active site that are involved in catalysis are conserved between ATP- and GTP-dependent PEPCK, some other factor must be contributing to this phenomenon. Another factor to be considered is mobile loop dynamics and conformational change. For PEPCK in particular, from substrate binding to product release the enzyme undergoes a conformational change where the N and C-terminal domains enclose around the active site. In addition, the Ω-loop is very flexible and acts as a lid closing over the active site. The interactions between residues on the Ω-loop and the R-loop differ between ATP- and GTP-dependent PEPCK. ATP-PEPCK has fewer interactions between mobile loops in its active site, which could explain how EcPEPCK has greater product turnover. Fewer interactions suggest less energy is required to go from the different conformational states seen during the process of enzyme turnover.

Lastly, the question arises: is EcPEPCK a true proxy for all ATP-dependent PEPCK enzymes? It can be concluded that in choosing a single enzyme to represent the entire subclass, EcPEPCK is as good a choice as any, and is likely better than other options. EcPEPCK is physiologically active as a monomer, just like its GTP-dependent counterpart, hcPEPCK. However, for a more complete analysis of different enzymes from both nucleotide-dependent classes, enzymes derived from a wider variety of species would be beneficial. Especially a better focus on ATP-PEPCK from anaerobic microorganisms previously reported to catalyze CO_2 fixation *in vivo* instead of the more common gluconeogenic reaction.

To add to the confusion of the PEPCK superfamily, a third class of PEPCK enzymes was recently added. This new class includes the non-nucleotide dependent PP_i-PEPCK enzymes which were initially termed PEP carboxytransphosphorylases. This class of enzymes use

pyrophosphate as the phosphoryl donor, as the name suggests. These enzymes are more distantly related to the rest of the PEPCK family, and while it is possible the PEPCK core structure evolved once again, it is more likely that this is also a case of divergent evolution. Further research needs to be done on the PP_i-PEPCK enzyme, and starting with lobe deletion mutants would be a good first step. As of late 2023, there are only two PP_i-PEPCK structures in the PDB databank, both of enzymes derived from bacteria: *Actinomyces israelii* and *Propionibacterium freudenreichii*. Initial studies in our lab have been performed using the latter enzyme, and said structure was published and uploaded to the PDB by a previous lab member in 2023.[31]

5.3 EcPEPCK with pyruvate kinase-like activity

Using a combination of spectrophotometric assays and NMR-based assays, the question of whether EcPEPCK can catalyze the conversion of PEP to pyruvate (that is, a pyruvate kinase-like activity) has been answered. EcPEPCK has PK-like activity that has only been observed before in yeast PEPCK and *A. succiniproducens*.[93]

Accordingly, this PK-like activity is significantly lower than the normal PEPCK activity (~1%), so much so that when it was first observed in Bakers' yeast, it was assumed that there was endogenous pyruvate kinase contaminating the isolated PEPCK sample.[105] Using recombinant protein expression and affinity-tagged purification of EcPEPCK, we have concluded that there is no contamination, and all PK-like activity comes from EcPEPCK. The spectrophotometric assay provides evidence of pyruvate formation, and the corresponding ^{31}P-NMR assay provides the evidence of ATP formation after the phosphate transfer. The ability of EcPEPCK to catalyze a PK-like reaction is presumably an artifact of evolution, due to its similarity to its typical PEPCK-catalyzed reaction. All of the substrates and nucleotides required are the same, excluding the

CO_2. An interesting benefit, however, is in the case of *B. subtillus*, where this PK-like activity was also observed when the gene for pyruvate kinase was knocked out. Not all of the bacteria thrived, but a small proportion were able to grow, albeit very slowly due to its ATP-PEPCK enzyme stepping in to catalyze the formation of pyruvate and ATP in place of pyruvate kinase. It would be interesting if future research can attempt to replicate this phenomenon in *E. coli* and see what effect a PK knockout mutation would have on cell growth if EcPEPCK is recombinantly overexpressed.

The secondary phenomenon encountered via the structural data was the ability of EcPEPCK to use phosphoglycolate (PGA) as an alternate substrate. Presumably in this reaction, EcPEPCK is using ADP as the phosphoryl acceptor and transferring the phosphate group from PGA, forming ATP and glycolate. The same ^{31}P-NMR assay was used to qualitatively examine ATP production over time, revealing that this reaction was even slower than that of the PK-like reaction (less than a tenth of the activity). Logically this makes sense since PGA is not the typical substrate of EcPEPCK, and is in fact usually considered an inhibitor of all PEPCK enzymes, as a structural analog of PEP. Interestingly, it has previously been established that pyruvate kinase can catalyze the phosphorylation of glycolate using ATP, forming PGA and ADP.[101] Knowing now that EcPEPCK is capable of PK-like activity, and pyruvate kinase can also produce PGA, it is not an unreasonable conclusion that EcPEPCK would also be able to use PGA as a substrate. This observation brings up more questions to be answered, including whether the catalysis of this reaction is reversible, and whether a more appropriate assay can be designed to quantitatively monitor this reaction. ^{31}P-NMR is difficult to quantify, due to the properties of phosphorous and a long relaxation delay. A first step would be to use mass spectrometry to confirm that glycolate is being formed, then design a new assay for glycolate formation.

Previous research suggests that both rabbit muscle pyruvate kinase and GTP-PEPCK enzymes (isolated from chicken and rat liver) can catalyze the nucleotide-dependent phosphorylation of α-substituted carboxylic acids including glycolate, essentially the reverse of the PGA dephosphorylation catalyzed by EcPEPCK.[102] The study also states that GTP-PEPCK can catalyze another reaction shared with pyruvate kinase that more closely resembles the gluconeogenic direction typically catalyzed by PEPCK.[102] This reaction is the decarboxylation of oxaloacetate producing pyruvate and carbon dioxide. However, they conclude that GTP-PEPCK can not use pyruvate as a substrate, in contrast to the EcPEPCK data we present here. Future research can include revisiting this assay with EcPEPCK and other ATP-dependent PEPCKs, as well as hcPEPCK for a full comparison. In addition, the reverse of the reaction, aka dephosphorylation of these P-carboxylic acids could be assayed.

These additional catalytic activities are an exciting new discovery for ATP-dependent PEPCK enzymes, and provide more information on the evolutionary relationship between ATP and GTP PEPCK. PEPCK enzymes are characterized for their role in the initial key steps of gluconeogenesis. Both classes of PEPCK have similarities to pyruvate kinase, with the ATP-dependent EcPEPCK being able to mimic the last step of glycolysis as catalyzed by pyruvate kinase. Their similarities and differences also appear to be related to the direction of catalysis of which they preferentially perform. GTP-PEPCK enzymes, more commonly found in higher eukaryotes, are known for their role in gluconeogenesis *in vivo,* and while ATP-PEPCK are also known for the same, they more frequently play a role in carbon fixation in bacterial species. Both PEPCK and pyruvate kinase are integral to central metabolism and it is fascinating to hypothesize how they have evolved from their presumably primitive counterparts to the more selective and specialized enzymes of today.

References

(1) Utter, M. F., and Kurahashi, K. (1954) Purification of Oxalacetic Carboxylase from Chicken Liver. *J. Biol. Chem. 207*, 787–802.

(2) Utter, M. F., Kurahashi, K., and Rose, I. A. (1954) Some properties of oxalacetic carboxylase. *J. Biol. Chem. 207*, 803–819.

(3) Utter, M. F., and Kurahashi, K. (1954) Mechanism of action of oxalacetic carboxylase. *J. Biol. Chem. 207*, 821–841.

(4) Kurahashi, K., Pennington, R. J., and Utter, M. F. (1957) Nucleotide specificity of oxalacetic carboxylase. *J. Biol. Chem. 226*, 1059–1075.

(5) Matte, A., Goldie, H., Sweet, R. M., and Delbaere, L. T. J. (1996) Crystal structure of Escherichia coli phosphoenolpyruvate carboxykinase: A new structural family with the P-loop nucleoside triphosphate hydrolase fold. *J. Mol. Biol. 256*, 126–143.

(6) Matte, A., Tari, L. W., Goldie, H., and Delbaere, L. T. J. (1997) Structure and Mechanism of Phosphoenolpyruvate Carboxykinase. *J. Biol. Chem. 272*, 8105–8108.

(7) Dunten, P., Belunis, C., Crowther, R., Hollfelder, K., Kammlott, U., Levin, W., Michel, H., Ramsey, G. B., Swain, A., Weber, D., and Wertheimer, S. J. (2002) Crystal structure of human cytosolic phosphoenolpyruvate carboxykinase reveals a new GTP-binding site. *J. Mol. Biol. 316*, 257–264.

(8) Chiba, Y., Miyakawa, T., Shimane, Y., Takai, K., Tanokura, M., and Nozaki, T. (2019) Structural comparisons of phosphoenolpyruvate carboxykinases reveal the evolutionary trajectories of these phosphodiester energy conversion enzymes. *J. Biol. Chem. 294*, 19269–19278.

(9) Voet, D., Voet, J. G. (2011) Biochemistry 4th ed. John Wiley & Sons, Hoboken, New Jersey.

(10) Lipmann, F. (1941) Metabolic Generation and Utilization of Phosphate Bond Energy, in *Advances in Enzymology and Related Areas of Molecular Biology*, pp 99–162.

(11) Bowman, E., Mcqueney, M., Barry, R. J., and Dunaway-mariano, D. (1988) Catalysis and Thermodynamics of the Phosphoenolpyruvate/Phosphonopyruvate Rearrangement. Entry into the Phosphonate Class of Naturally OCcuring Organophosphorus Compounds. *J. Am. Chem. Soc. 110*, 5575–5576.

(12) Sauer, U., and Eikmanns, B. J. (2005) The PEP-pyruvate-oxaloacetate node as the switch point for carbon flux distribution in bacteria. *FEMS Microbiol. Rev. 29*, 765–794.

(13) Koendjbiharie, J. G., Van Kranenburg, R., and Kengen, S. W. M. (2021) The PEP-pyruvate-oxaloacetate node: variation at the heart of metabolism. *FEMS Microbiol. Rev. 45*, 1–19.

(14) Barwell, S. (2019) Structural and kinetic characterization of mechanisms of regulation of phosphoenolpyruvate carboxykinase by anions. *UW Sp. - Thesis Repos.* University of Waterloo.

(15) Stark, R., Guebre-Egziabher, F., Zhao, X., Feriod, C., Dong, J., Alves, T. C., Ioja, S., Pongratz, R. L., Bhanot, S., Roden, M., Cline, G. W., Shulman, G. I., and Kibbey, R. G. (2014) A role for mitochondrial phosphoenolpyruvate carboxykinase (PEPCK-M) in the regulation of hepatic gluconeogenesis. *J. Biol. Chem. 289*, 7257–7263.

(16) Stark, R., and Kibbey, R. G. (2014) The mitochondrial isoform of phosphoenolpyruvate

carboxykinase (PEPCK-M) and glucose homeostasis: has it been overloooked? *Biochim. Biophys. Acta - Gen. Subj. 1840*, 1313–1330.

(17) Yang, J., Kalhan, S. C., and Hanson, R. W. (2009) What Is the Metabolic Role of Phosphoenolpyruvate Carboxykinase? *J. Biol. Chem. 284*, 27025–27029.

(18) Hanson, R. W. (2009) Thematic minireview series: A perspective on the biology of phosphoenolpyruvate carboxykinase 55 years after its discovery. *J. Biol. Chem. 284*, 27021–27023.

(19) Modaressi, S., Brechtel, K., Christ, B., and Jungermann, K. (1998) Human mitochondrial phosphoenolpyruvate carboxykinase 2 gene: Structure, chromosomal localization and tissue-specific expression. *Biochem. J. 333*, 359–366.

(20) Weldon, S. L., Rando, A., Matathias, A. S., Hod, Y., Kalonick, P. A., Savon, S., Cook, J. S., and Hanson, R. W. (1990) Mitochondrial phosphoenolpyruvate carboxykinase from the chicken: Comparison of the cDNA and protein sequences with the cytosolic isozyme. *J. Biol. Chem. 265*, 7308–7317.

(21) McLeod, M. J., and Holyoak, T. (2021) Enyzmes | Phosphoenolpyruvate carboxykinases, in *Encyclopedia of Biological Chemistry: Third Edition*, pp 400–412.

(22) McLeod, M. (2020) Functional consequences of changing structure, dynamics, and free-energy landscapes of phosphoenolpyruvate carboxykinases. University of Waterloo.

(23) Chiba, Y., Kamikawa, R., Nakada-Tsukui, K., Saito-Nakano, Y., and Nozaki, T. (2015) Discovery of PP-type phosphoenolpyruvate carboxykinase genes in Eukaryotes and Bacteria. *J. Biol. Chem. 290*, 23960–23970.

(24) Troy A. Johnson, T. H. (2012) The Ω-Loop Lid Domain of Phosphoenolpyruvate Carboxykinase Is Essential for Catalytic Function. *Biochemistry 51*, 9547–9559.

(25) Johnson, T. A., McLeod, M. J., and Holyoak, T. (2016) Utilization of Substrate Intrinsic Binding Energy for Conformational Change and Catalytic Function in Phosphoenolpyruvate Carboxykinase. *Biochemistry 55*, 575–587.

(26) McNicholas, S., Potterton, E., Wilson, K. S., and Noble, M. E. M. (2011) Presenting your structures: The CCP4mg molecular-graphics software. *Acta Crystallogr. Sect. D Biol. Crystallogr. 67*, 386–394.

(27) Needleman, S. B., and Wunsch, C. D. (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol. 48*, 443–453.

(28) Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A. R. N., Potter, S. C., Finn, R. D., and Lopez, R. (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res. 47*, W636–W641.

(29) Zhang, Y., and Skolnick, J. (2005) TM-align: A protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res. 33*, 2302–2309.

(30) Burnell, J. N. (1986) Purification and properties of phosphoenolpyruvate carboxykinase from C4 Plants. *Aust. J. Plant Physiol. 13*, 577–587.

(31) McLeod, M. J., and Holyoak, T. (2023) Biochemical, structural, and kinetic characterization of PPi-dependent phosphoenolpyruvate carboxykinase from Propionibacterium freudenreichii. *Proteins Struct. Funct. Bioinforma. 91*, 1261–1275.

(32) Ramnanan, C. J., Edgerton, D. S., and Cherrington, A. D. (2009) The role of insulin in the regulation of PEPCK and gluconeogenesis in vivo. *US Endocrinol. 5*, 34–39.

(33) SHRAGO, E., LARDY, H. A., NORDLIE, R. C., and FOSTER, D. O. (1963) Metabolic and Hormonal Control of Phosphoenolpyruvate Carboxykinase and. *J. Biol. Chem. 238*, 3188–3192.

(34) Shrago, E., Shug, A. L. (1969) Regulation of Phosphoenolpyruvate Carboxykinase in Escherichia coli by Carbohydrate Metabolism. *Arch. Biochem. Biophys. 130*, 393–398.

(35) Yang, J., Kalhan, S. C., and Hanson, R. W. (2009, October 2) What is the metabolic role of phosphoenolpyruvate carboxykinase? *J. Biol. Chem.*

(36) Forest, C., Tordjman, J., Glorian, M., Duplus, E., Chauvet, G., Quette, J., Beale, E. G., and Antoine, B. (2003) Fatty acid recycling in adipocytes: A role for glyceroneogenesis and phosphoenolpyruvate carboxykinase. *Biochem. Soc. Trans. 31*, 1125–1129.

(37) Feliu, J. E., Hue, L., and Hers, H. G. (1976) Hormonal control of pyruvate kinase activity and of gluconeogenesis in isolated hepatocytes. *Proc. Natl. Acad. Sci. USA 73*, 2762–2766.

(38) Zamboni, N., Maaheimo, H., Szyperski, T., Hohmann, H. P., and Sauer, U. (2004) The phosphoenolpyruvate carboxykinase also catalyzes C3 carboxylation at the interface of glycolysis and the TCA cycle of Bacillus subtilis. *Metab. Eng. 6*, 277–284.

(39) Cannata, J. J. B., and de Flombaum, M. A. C. (1974) Phosphoenolpyruvate Carboxykinase from Bakers' Yeast. Kinetics of Phosphoenolpyruvate Formation. *J. Biol. Chem. 249*, 3356–3365.

(40) Podkovyrov, S. M., and Zeikus, J. G. (1993) Purification and characterization of phosphoenolpyruvate carboxykinase, a catabolic CO2-fixing enzyme, from Anaerobiospirillum succiniciproducens. *J. Gen. Microbiol. 139*, 223–228.

(41) Higa, A. I., Milrad de Forchetti, S. R., and Cazzulo, J. J. (1976) C02-fixing Enzymes in Pseudomonas fluorescens. *Jounral Gen. Microbiol. 93*, 69–74.

(42) Cannata, J.J.B., Stoppani, A. O. M. (1963) Phosphopyruvate Carboxylase from Bakers' Yeast. I. Isolation, Purification, and Characterization. *J. Biol. Chem. 238*, 1196–1207.

(43) Burgess, S. C., He, T. T., Yan, Z., Lindner, J., Sherry, A. D., Malloy, C. R., Browning, J. D., and Magnuson, M. A. (2007) Cytosolic Phosphoenolpyruvate Carboxykinase Does Not Solely Control the Rate of Hepatic Gluconeogenesis in the Intact Mouse Liver. *Cell Metab. 5*, 313–320.

(44) DeBerardinis, R. J., Lum, J. J., Hatzivassiliou, G., and Thompson, C. B. (2008) The Biology of Cancer: Metabolic Reprogramming Fuels Cell Growth and Proliferation. *Cell Metab. 7*, 11–20.

(45) Balsa-Martinez, E., and Puigserver, P. (2015) Cancer Cells Hijack Gluconeogenic Enzymes to Fuel Cell Growth. *Mol. Cell 60*, 509–511.

(46) Xu, D., Wang, Z., Xia, Y., Shao, F., Xia, W., Wei, Y., Li, X., Qian, X., Lee, J. H., Du, L., Zheng, Y., Lv, G., Leu, J. shiun, Wang, H., Xing, D., Liang, T., Hung, M. C., and Lu, Z. (2020) The gluconeogenic enzyme PCK1 phosphorylates INSIG1/2 for lipogenesis. *Nature 580*, 530–535.

(47) Jiang, H., Zhu, L., Xu, D., and Lu, Z. (2020) A newly discovered role of metabolic enzyme

PCK1 as a protein kinase to promote cancer lipogenesis. *Cancer Commun. 40*, 389–394.

(48) McLeod, M. J., Krismanich, A. P., Assoud, A., Dmitrienko, G. I., and Holyoak, T. (2019) Characterization of 3-[(Carboxymethyl)thio]picolinic Acid: A Novel Inhibitor of Phosphoenolpyruvate Carboxykinase. *Biochemistry 58*, 3918–3926.

(49) Stiffin, R. M., Sullivan, S. M., Carlson, G. M., and Holyoak, T. (2008) Differential inhibition of cytosolic PEPCK by substrate analogues. Kinetic and structural characterization of inhibitor recognition. *Biochemistry 47*, 2099–2109.

(50) Troy A. Johnson, T. H. (2010) Increasing the conformational entropy of the Ω-loop lid domain in PEPCK impairs catalysis and decreases catalytic fidelity. *Biochemistry 49*, 5176–5187.

(51) Cousins, K. (1993) ChemOffice Plus: A Package of Programs for Chemists. *J. Chem. Inf. Comput. Sci. 33*, 788–789.

(52) Cousins, K. R. (2011) Computer review of ChemDraw ultra 12.0. *J. Am. Chem. Soc. 133*, 8388.

(53) Robinson, B. H., and Oei, J. (1975) 3-Mercaptopicolinic acid, a preferential inhibitor of the cytosolic phosphoenolpyruvate carboxykinase. *FEBS Lett. 58*, 12–15.

(54) Makinen, A. L., and Nowak, T. (1983) 3-Mercaptopicolinate. A reversible active site inhibitor of avian liver phosphoenolpyruvate carboxykinase. *J. Biol. Chem. 258*, 11654–11662.

(55) Makinen, A. L., and Nowak, T. (2011) 3-Mercaptopicolinate. *J. Biol. Chem. 258*, 11654–11662.

(56) Dong, X. Y., Tang, S. Q., and Chen, J. D. (2012) Dual functions of Insig proteins in cholesterol homeostasis. *Lipids Health Dis. 11*, 1.

(57) Horton, J. D., Goldstein, J. L., and Brown, M. S. (2002) SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. *J. Clin Invest. 109*, 1125–1131.

(58) Brown, M. S., Ye, J., Rawson, R. B., and Goldstein, J. L. (2000) Regulated intramembrane proteolysis: a control mechanism conserved from bacteria to humans. *Cell 100*, 391–398.

(59) Yabe, D., Brown, M. S., and Goldstein, J. L. (2002) Insig-2, a second endoplasmic reticulum protein that binds SCAP and blocks export of sterol regulatory element-binding proteins. *PNAS 99*, 12753–12758.

(60) Yang, T., Espenshade, P. J., Wright, M. E., Yabe, D., Gong, Y., Aebersold, R., Goldstein, J. L., and Brown, M. S. (2002) Crucial step in cholesterol homeostasis: Sterols promote binding of SCAP to INSIG-1, a membrane protein that facilitates retention of SREBPs in ER. *Cell 110*, 489–500.

(61) DeBose-Boyd, R. A., and Ye, J. (2018) SREBPs in Lipid Metabolism, Insulin Signaling, and Beyond. *Trends Biochem Sci. 43*, 358–368.

(62) Jiang, T., Zhang, G., and Lou, Z. (2020) Role of the Sterol Regulatory Element Binding Protein Pathway in Tumorigenesis. *Front. Oncol. 10*, 1–15.

(63) Sun, L. P., Li, L., Goldstein, J. L., and Brown, M. S. (2005) Insig required for sterol-mediated inhibition of Scap/SREBP binding to COPII proteins in vitro. *J. Biol. Chem. 280*, 26483–26490.

(64) Yan, R., Cao, P., Song, W., Qian, H., Du, X., Coates, H. W., Zhao, X., Li, Y., Gao, S., Gong, X., Liu, X., Sui, J., Lei, J., Yang, H., Brown, A. J., Zhou, Q., Yan, C., and Yan, N. (2021) A structure of human scap bound to insig-2 suggests how their interaction is regulated by sterols. *Science (80-.). 371*.

(65) Feramisco, J. D., Goldstein, J. L., and Brown, M. S. (2004) Membrane Topology of Human Insig-1, a Protein Regulator of Lipid Synthesis. *J. Biol. Chem. 279*, 8487–8496.

(66) Studier, F. W. (2005) Protein Production by Auto-Induction in High-Density Shaking Cultures. *Protein Expr. Purif. 41*, 207–234.

(67) Barwell, S. A. E., Duman, R., Wagner, A., and Holyoak, T. (2022) Directional regulation of cytosolic PEPCK catalysis is mediated by competitive binding of anions. *Biochem. Biophys. Res. Commun. 637*, 218–223.

(68) Otwinowski, Z., Minor, W. (1997) Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzym. 276*, 307–326.

(69) Winter, G., Waterman, D. G., Parkhurst, J. M., Brewster, A. S., Gildea, R. J., Gerstel, M., Fuentes-Montero, L., Vollmar, M., Michels-Clark, T., Young, I. D., Sauter, N. K., and Evans, G. (2018) DIALS: Implementation and evaluation of a new integration package. *Acta Crystallogr. Sect. D Struct. Biol. 74*, 85–97.

(70) Winter, G., Beilsten-Edmands, J., Devenish, N., Gerstel, M., Gildea, R. J., McDonagh, D., Pascal, E., Waterman, D. G., Williams, B. H., and Evans, G. (2022) DIALS as a toolkit. *Protein Sci. 31*, 232–250.

(71) 4, C. C. P. N. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. Sect. D Biol. Crystallogr. 50*, 760–763.

(72) Agirre, J., Atanasova, M., Bagdonas, H., Ballard, C. B., Baslé, A., Beilsten-Edmands, J., Borges, R. J., Brown, D. G., Burgos-Mármol, J. J., Berrisford, J. M., Bond, P. S., Caballero, I., Catapano, L., Chojnowski, G., Cook, A. G., Cowtan, K. D., Croll, T. I., Debreczeni, J., Devenish, N. E., Dodson, E. J., Drevon, T. R., Emsley, P., Evans, G., Evans, P. R., Fando, M., Foadi, J., Fuentes-Montero, L., Garman, E. F., Gerstel, M., Gildea, R. J., Hatti, K., Hekkelman, M. L., Heuser, P., Hoh, S. W., Hough, M. A., Jenkins, H. T., Jiménez, E., Joosten, R. P., Keegan, R. M., Keep, N., Krissinel, E. B., Kolenko, P., Kovalevskiy, O., Lamzin, V. S., Lawson, D. M., Lebedev, A. A., Leslie, A. G. W., Lohkamp, B., Long, F., Malý, M., McCoy, A. J., McNicholas, S. J., Medina, A., Millán, C., Murray, J. W., Murshudov, G. N., Nicholls, R. A., Noble, M. E. M., Oeffner, R., Pannu, N. S., Parkhurst, J. M., Pearce, N., Pereira, J., Perrakis, A., Powell, H. R., Read, R. J., Rigden, D. J., Rochira, W., Sammito, M., Sánchez Rodríguez, F., Sheldrick, G. M., Shelley, K. L., Simkovic, F., Simpkin, A. J., Skubak, P., Sobolev, E., Steiner, R. A., Stevenson, K., Tews, I., Thomas, J. M. H., Thorn, A., Valls, J. T., Uski, V., Usón, I., Vagin, A., Velankar, S., Vollmar, M., Walden, H., Waterman, D., Wilson, K. S., Winn, M. D., Winter, G., Wojdyr, M., and Yamashita, K. (2023) The CCP4 suite: integrative software for macromolecular crystallography. *Acta Crystallogr. Sect. D, Struct. Biol. 79*, 449–461.

(73) Emsley, P., Lohkamp, B., Scott, W. G., Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr. 66*, 486–501.

(74) Yung-Chi, C., and Prusoff, W. H. (1973) Relationship between the inhibition constant (KI) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. *Biochem. Pharmacol. 22*, 3099–3108.

(75) Cheng, H. C. (2002) The power issue: Determination of KB or Ki from IC50 - A closer look at the Cheng-Prusoff equation, the Schild plot and related power equations. *J. Pharmacol. Toxicol. Methods 46*, 61–71.

(76) Cui, D. S., Broom, A., McLeod, M. J., Meiering, E. M., and Holyoak, T. (2017) Asymmetric anchoring is required for efficient Ω-loop opening and closing in cytosolic phosphoenolpyruvate carboxykinase. *Biochemistry 56*, 2106–2115.

(77) Grasmann, G., Smolle, E., Olschewski, H., and Leithner, K. (2019) Gluconeogenesis in cancer cells – Repurposing of a starvation-induced metabolic pathway? *Biochim. Biophys. Acta - Rev. Cancer 1872*, 24–36.

(78) Cannata, J. J. B. (1970) Phosphoenolpyruvate Carboxykinase from Bakers' Yeast. Isolation of the Enzyme and Study of Its Physical Properties. *J. Biol. Chem. 245*, 792–798.

(79) Chang, H.-C., and Lane, M. D. (1966) The Enzymatic Carboxylation of Phosphoenolpyruvate II. Purification and Properties of Liver Mitochondrial Phosphoenolpyruvate Carboxykinase. *J. Biol. Chem. 241*, 2413–2420.

(80) Chang, H.-C., Maruyama, H., Miller, R. S., and Lane, M. D. (1966) The Enzymatic Carboxylation of Phosphoenolpyruvate. III. Investigation of the Kinetics and Mechanism of the Mitochondrial Phosphoenolpyruvate Carboxykinase-Catalyzed Reaction. *J. Biol. Chem. 241*, 2421–2430.

(81) Foster, D. O., Lardy, H. A., Ray, P. D., and Johnston, J. B. (1967) Alteration of Rat Liver Phosphoenolpyruvate Carboxykinase Activity by L-Tryptophan in Vivo and Metals in Vitro. *Biochemistry 6*, 2120–2128.

(82) Cazzulo, J J, Stoppani, A. O. M. (1969) Effects of Adenosine Phosphates and Nicotinamide Nucleotides on Pyruvate Carboxylase from Baker's Yeast *112*, 755–762.

(83) Utter, M. F., and Kolenbrander, H. M. (1972) Formation of Oxalacetate by CO2 Fixation on Phosphoenolpyruvate, in *The Enzymes*, pp 117–168. Academic Press Inc.

(84) de Flombaum, M. A., Cannata, J. J. B., Cazzulo, J. J., and Segura, E. L. (1977) Co2-fixing enzymes in Trypanosoma cruzi. *Comp. Biochem. Physiol. 58*, 67–69.

(85) Rohrer, S. P., Saz, H. J., and Nowak, T. (1986) Purification and characterization of phosphoenolpyruvate carboxykinase from the parasitic helminth Ascaris suum. *J. Biol. Chem. 261*, 13049–13055.

(86) Cymeryng, C., Cazzulo, J. J., and Cannata, J. J. B. (1995) Phosphoenolpyruvate carboxykinase from Trypanosoma cruzi. Purification and physicochemical and kinetic properties. *Mol. Biochem. Parasitol. 73*, 91–101.

(87) Weinstein, C. L., and Griffith, O. W. (1986) β-sulfopyruvate: Chemical and enzymatic syntheses and enzymatic assay. *Anal. Biochem. 156*, 154–160.

(88) Colombo, G., Carlson, G. M., and Lardy, H. A. (1981) Phosphoenolpyruvate Carboxykinase (Guanosine 5'-Triphosphate) from Rat Liver Cytosol. Dual-Cation Requirement for the Carboxylation Reaction. *Biochemistry 20*, 2749–2757.

(89) Sudom, A., Walters, R., Pastushok, L., Goldie, D., Prasad, L., Delbaere, L. T. J., and Goldie, H. (2003) Mechanisms of activation of phosphoenolpyruvate carboxykinase from Escherichia coli by Ca2+ and of desensitization by trypsin. *J. Bacteriol. 185*, 4233–4242.

(90) Ryan, M. F. (1991) The role of magnesium in clinical biochemistry: An overview. *Ann. Clin. Biochem. 28*, 19–26.

(91) Holyoak, T., and Nowak, T. (2004) pH dependence of the reaction catalyzed by avian mitochondrial phosphoenolpyruvate carboxykinase. *Biochemistry 43*, 7054–7065.

(92) Dombrauckas, J. D., Santarsiero, B. D., and Mesecar, A. D. (2005) Structural Basis for Tumor Pyruvate Kinase M2 Allosteric Regulation and Catalysis. *Biochemistry 44*, 9417–9429.

(93) Jabalquinto, A. M., Laivenieks, M., Zeikus, J. G., and Cardemil, E. (1999) Characterization of the oxaloacetate decarboxylase and pyruvate kinase- like activities of Saccharomyces cerevisiae and Anaerobiospirillum succiniciproducens phosphoenolpyruvate carboxykinases. *J. Protein Chem. 18*, 659–664.

(94) Beber, M. E., Gollub, M. G., Mozaffari, D., Shebek, K. M., Flamholz, A. I., Milo, R., and Noor, E. (2022) EQuilibrator 3.0: A database solution for thermodynamic constant estimation. *Nucleic Acids Res. 50*, D603–D609.

(95) Stocchi, V., Magnani, M., Cucchiarini, L., Novelli, G., and Dallapiccola, B. (1985) Red blood cell adenine nucleotides abnormalities in Down syndrome. *Am J Med Genet. 20*, 131–5.

(96) Crescentini, G., and Stocchi, V. (1984) Fast reversed-phase high-performance liquid chromatographic determination of nucleotides in red blood cells. *J. Chromatogr. 290*, 393–399.

(97) Stocchi, V., Cucchiarini, L., Magnani, M., Chiarantini, L., Palma, P., and Crescentini, G. (1985) Simultaneous extraction and reverse-phase high-performance liquid chromatographic determination of adenine and pyridine nucleotides in human red blood cells. *Anal. Biochem. 146*, 118–124.

(98) Arranz-Mascarós, P., Bazzicalupi, C., Bianchi, A., Giorgi, C., Godino-Salido, M. L., Gutiérrez-Valero, M. D., Lopez-Garzón, R., and Valtancoli, B. (2011) Binding and recognition of AMP, ADP, ATP and related inorganic phosphate anions by a tren-based ligand containing a pyrimidine functionality. *New J. Chem. 35*, 1883–1891.

(99) Brown, T. R., and Ogawa, S. (1977) 31P nuclear magnetic resonance kinetic measurements on adenylate kinase. *Proc. Natl. Acad. Sci. U. S. A. 74*, 3627–3631.

(100) Friebolin, H. (1998) Basic One-and Two-Dimensional NMR Spectroscopy.

(101) Kayne, F. J. (1974) Pyruvate Kinase Catalyzed Phosphorylation of Glycolate. *Biochem. Biophys. Res. Commun. 59*, 8–13.

(102) Ash, D. E., Emig, F. A., Chowdhury, S. A., Satoh, Y., and Schramm, V. L. (1990) Mammalian and Avian Liver Phosphoenolpyruvate Carboxykinase. Alternate Substrates and Inhibition by Analogues of Oxaloacetate. *J. Biol. Chem. 265*, 7377–7384.

(103) Sullivan, S. M., and Holyoak, T. (2008) Enzymes with lid-gated active sites must operate by an induced fit mechanism instead of conformational selection. *Proc. Natl. Acad. Sci. U. S. A. 105*, 13829–13834.

(104) Sullivan, S. M., and Holyoak, T. (2007) Structures of rat cytosolic PEPCK: Insight into the mechanism of phosphorylation and decarboxylation of oxaloacetic acid. *Biochemistry 46*, 10078–10088.

(105) Cannata, J. J. B., and Stoppani, A. O. M. (1963) Phosphopyruvate Carboxylase from Bakers' Yeast. II. Properties of Enzyme. *J. Biol. Chem. 238*, 1208–1212.

Appendices

Appendix A – PEPCK Expression Vector Map

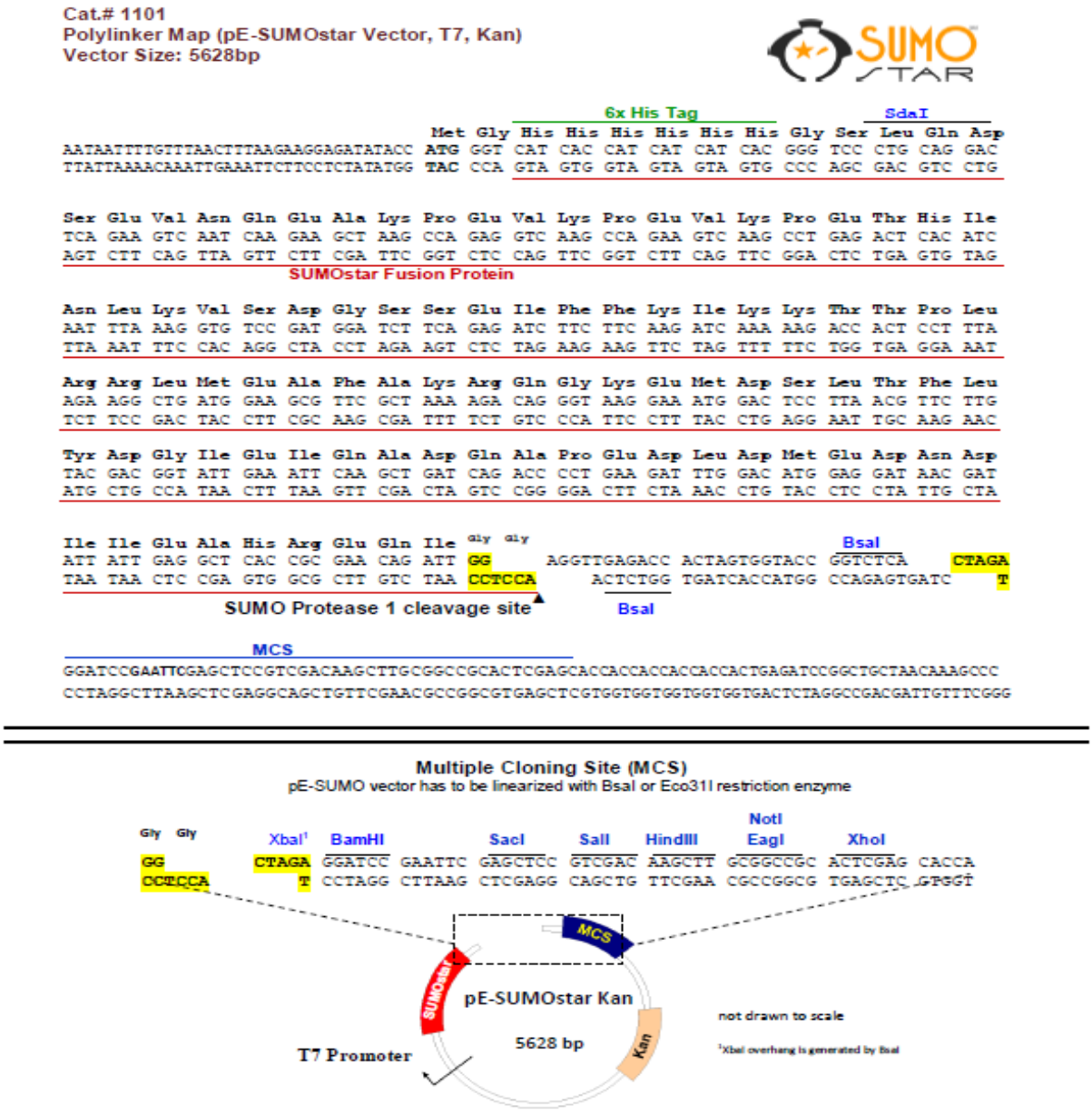


Figure A1: Polylinker map for the pE-SUMOstar expression vector. The hcPEPCK and EcPEPCK genes were cloned into the pE-SUMOstar vector as described previously, transformed into the BL21(DE3) *E. coli* protein expression cell line and stored at -80°C as a common cell stock.

Appendix B – Raw kinetic data for hcPEPCK

WT and S90E hcPEPCK Kinetic Characterization of Enzyme Activity and Substrate Affinity

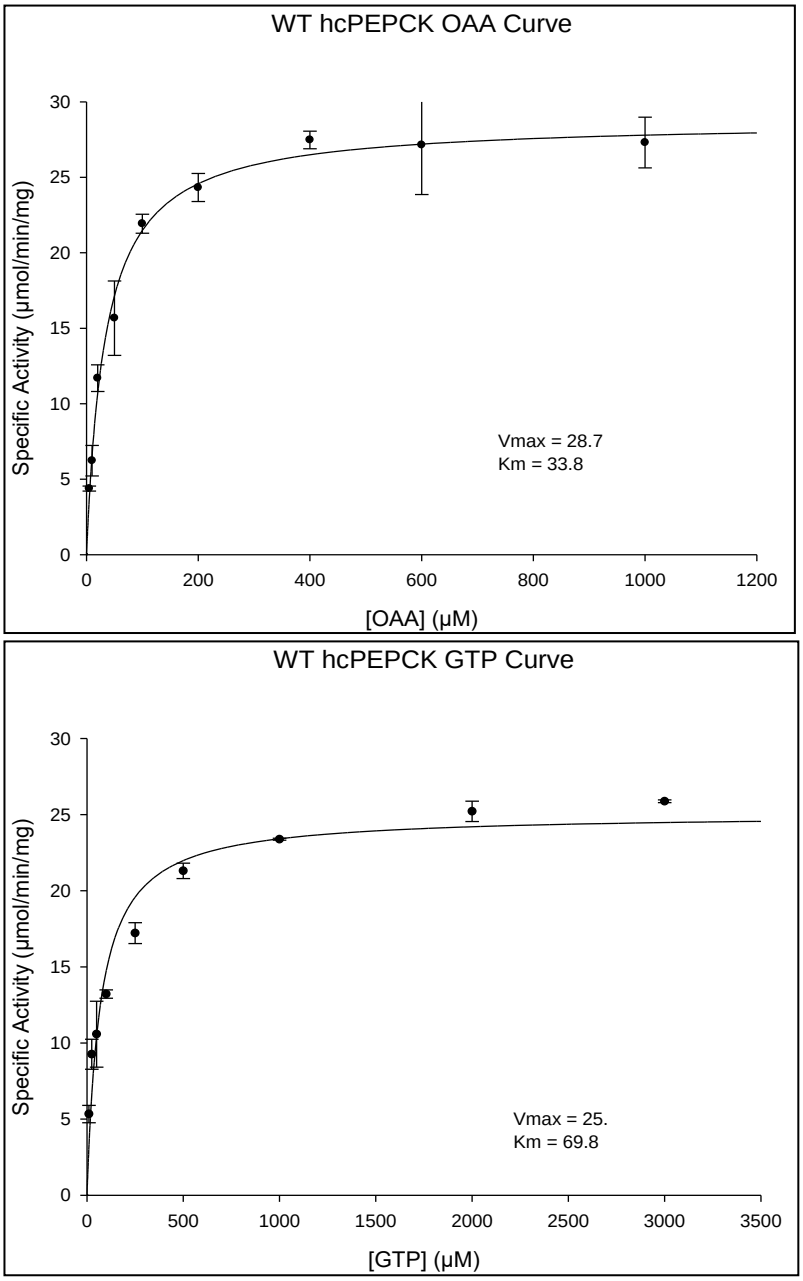
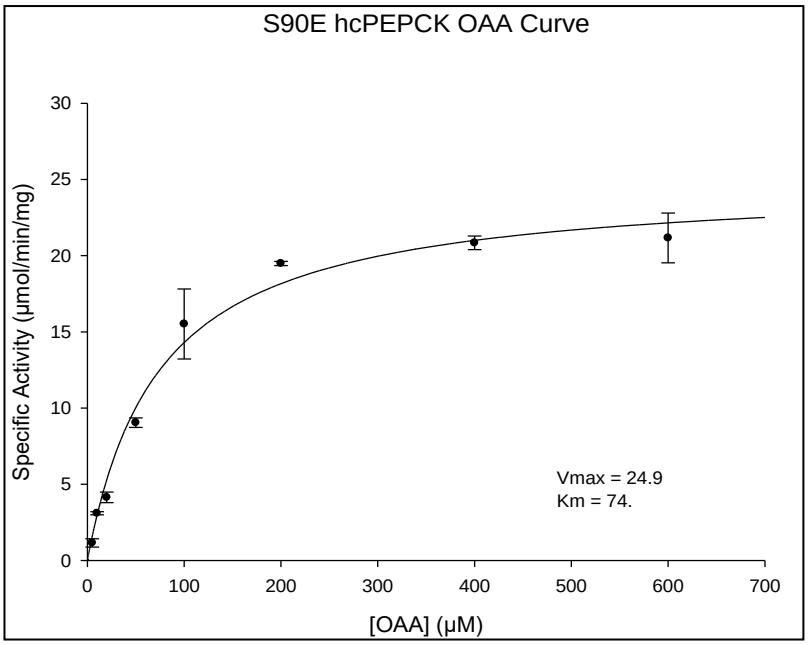


Figure B1: Michaelis-Menten curves for the characterization of WT hcPEPCK activity in the OAA to PEP direction.



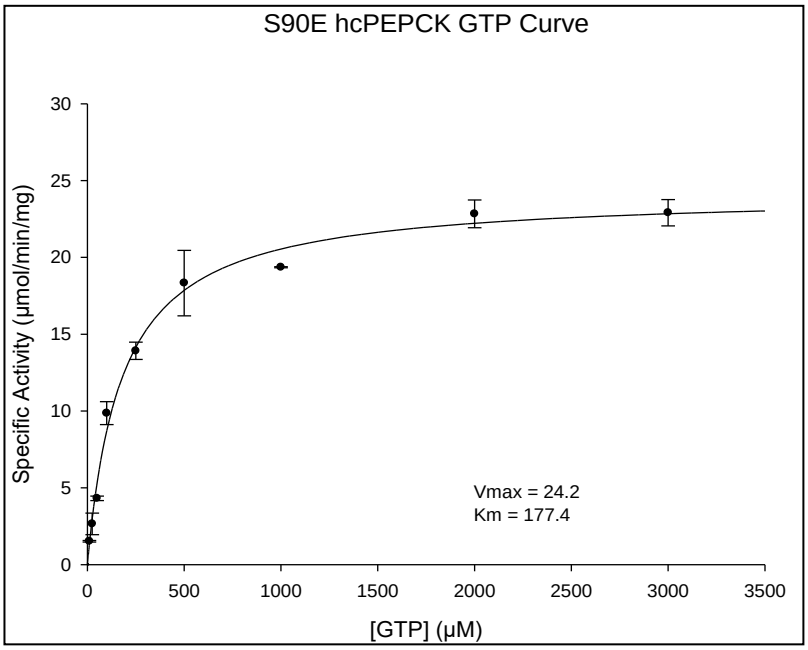


Figure B2: Michaelis-Menten curves for the characterization of S90E hcPEPCK activity in the OAA to PEP direction.

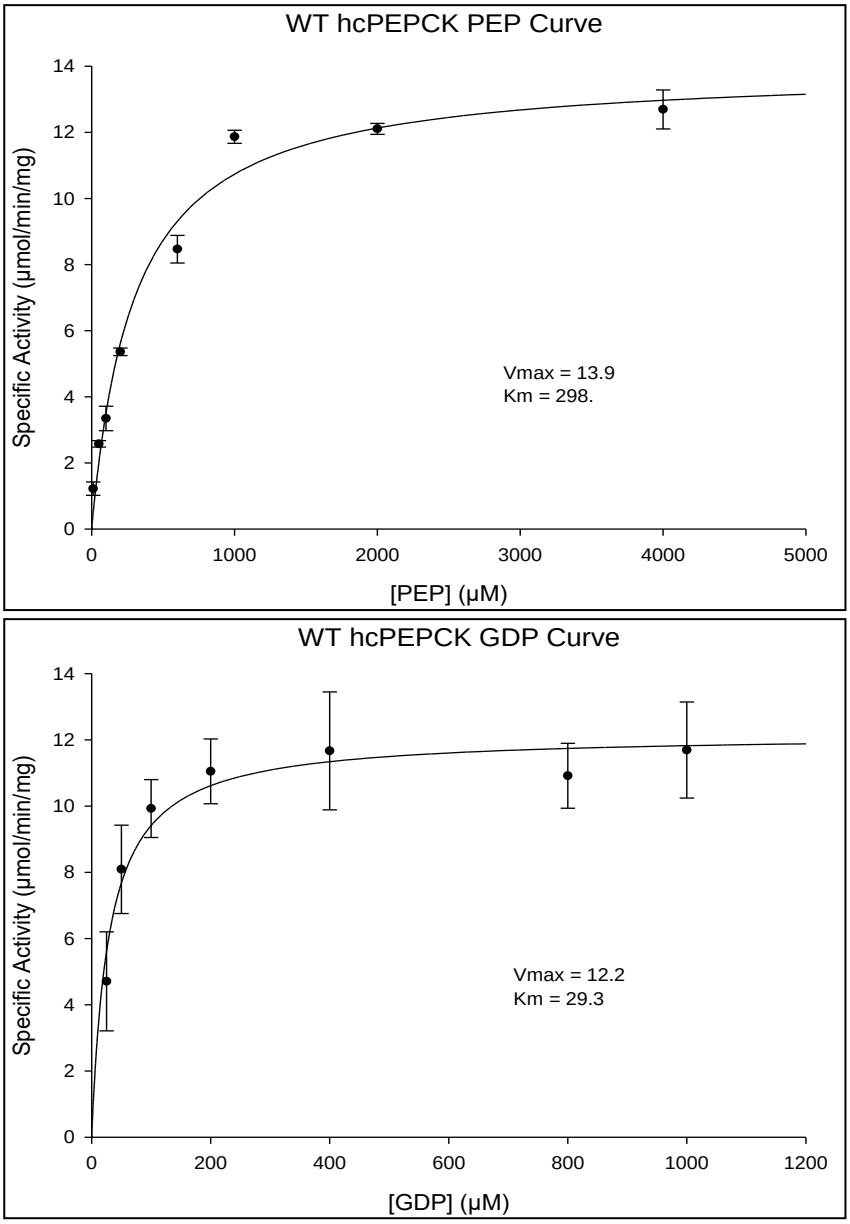


Figure B3: Michaelis-Menten curves for the characterization of WT hcPEPCK activity in the PEP to OAA direction.

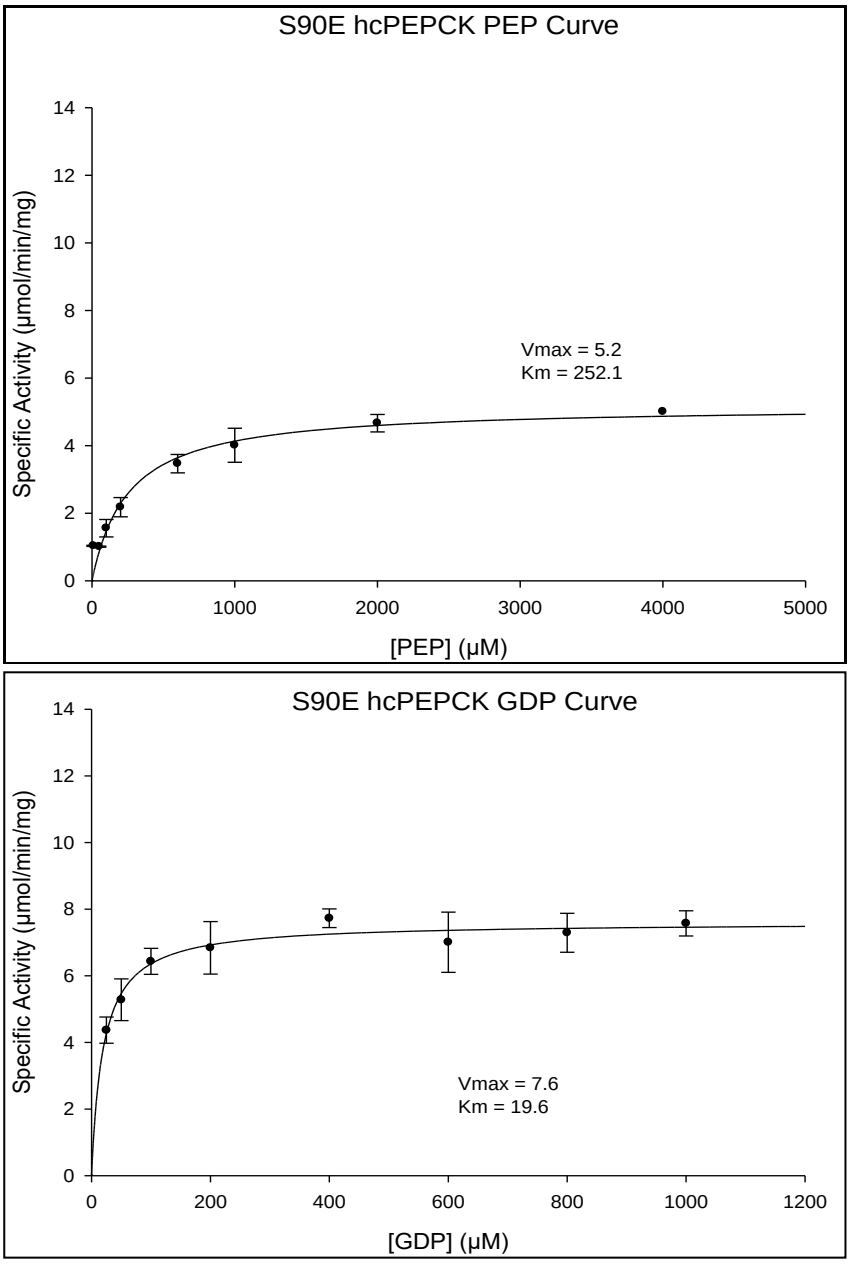


Figure B4: Michaelis-Menten curves for the characterization of S90E hcPEPCK activity in the PEP to OAA direction.

Determining Oxalate inhibition of WT and S90E hcPEPCK against PEP

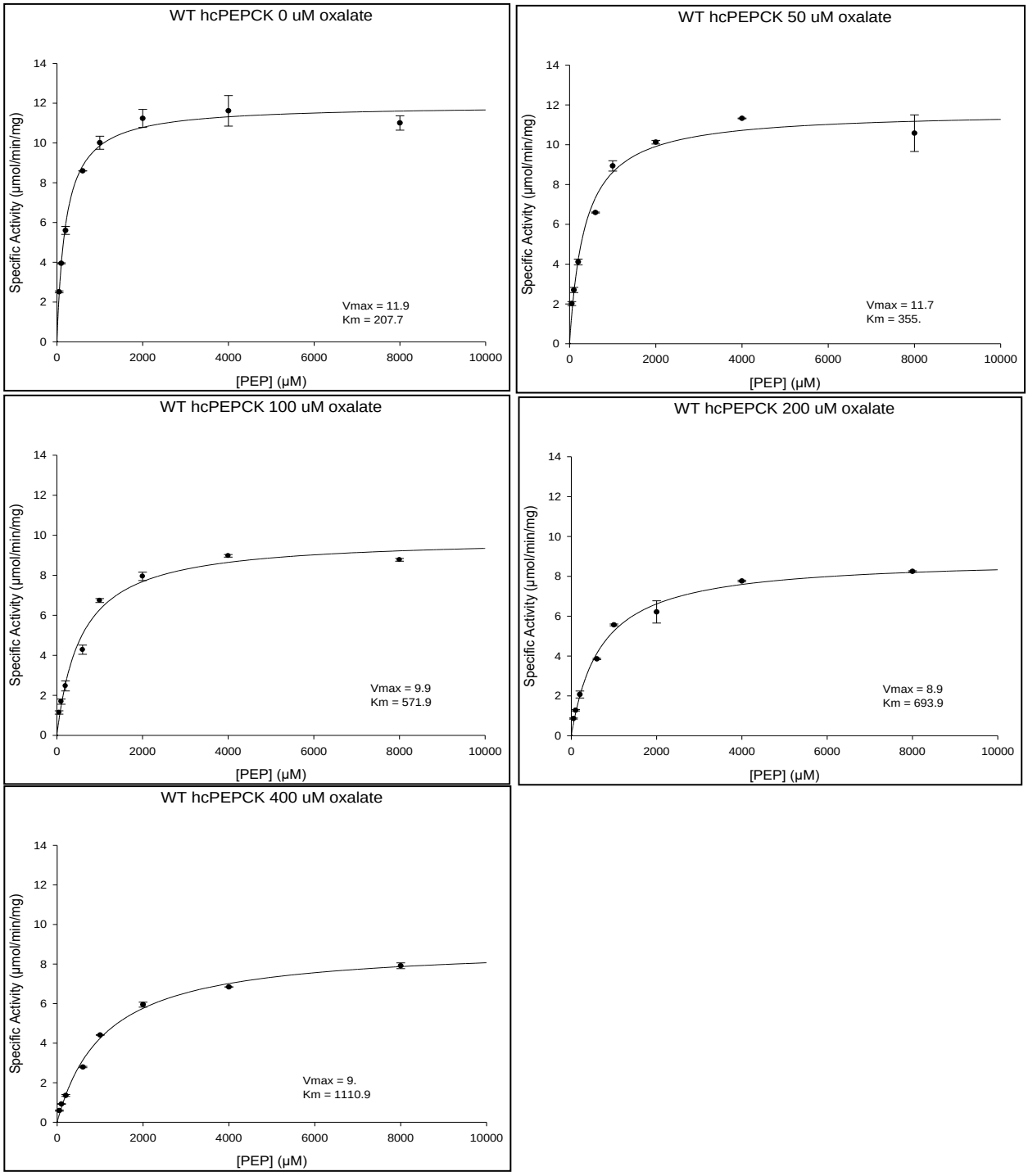
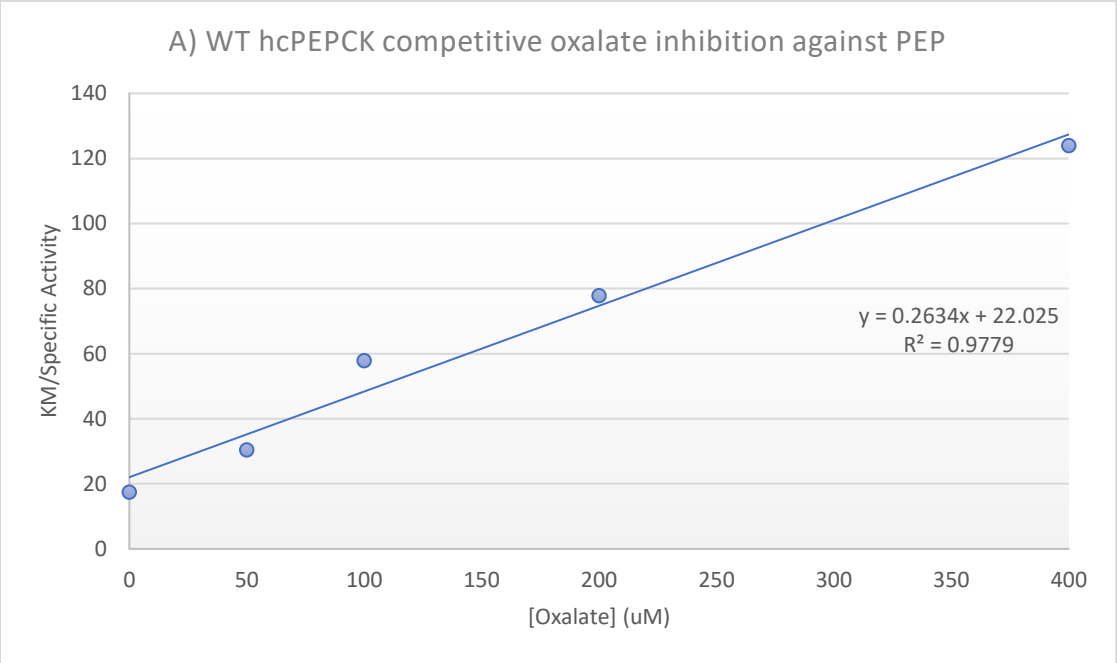


Figure B5: WT hcPEPCK PEP Michaelis-Menten curves under increasing oxalate concentrations.



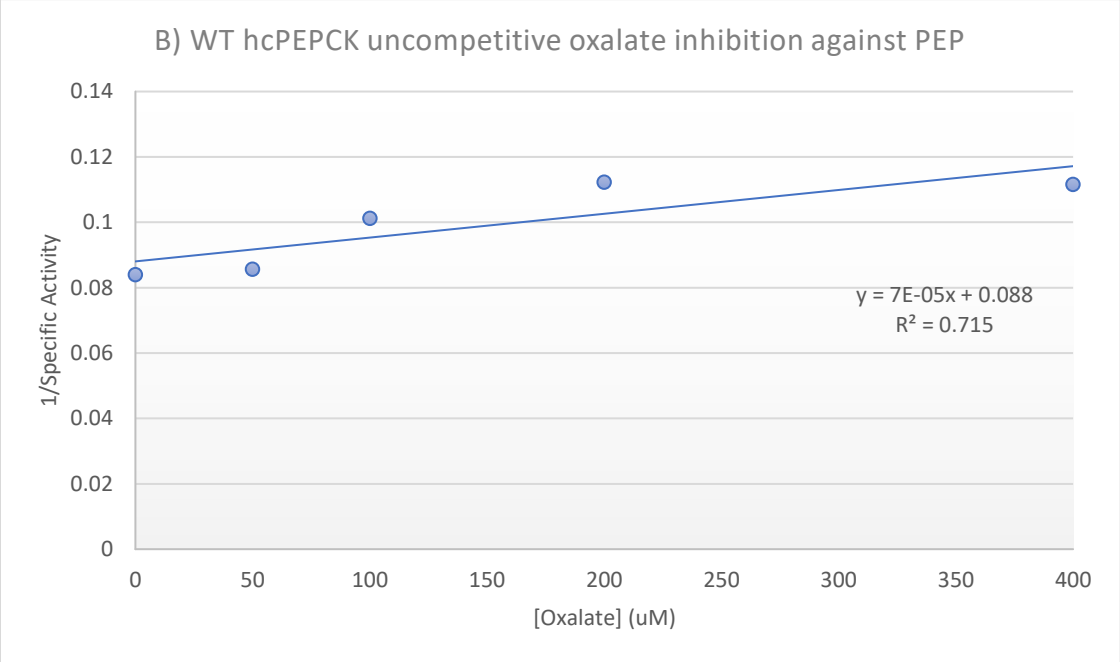


Figure B6: Kinetic re-plots of oxalate concentration vs specific activity for WT hcPEPCK in the reverse direction of catalysis. Plot A uses the ratio of KM/specific activity to determine competitive inhibition; Plot B uses the ratio of 1/specific activity to determine uncompetitive inhibition.

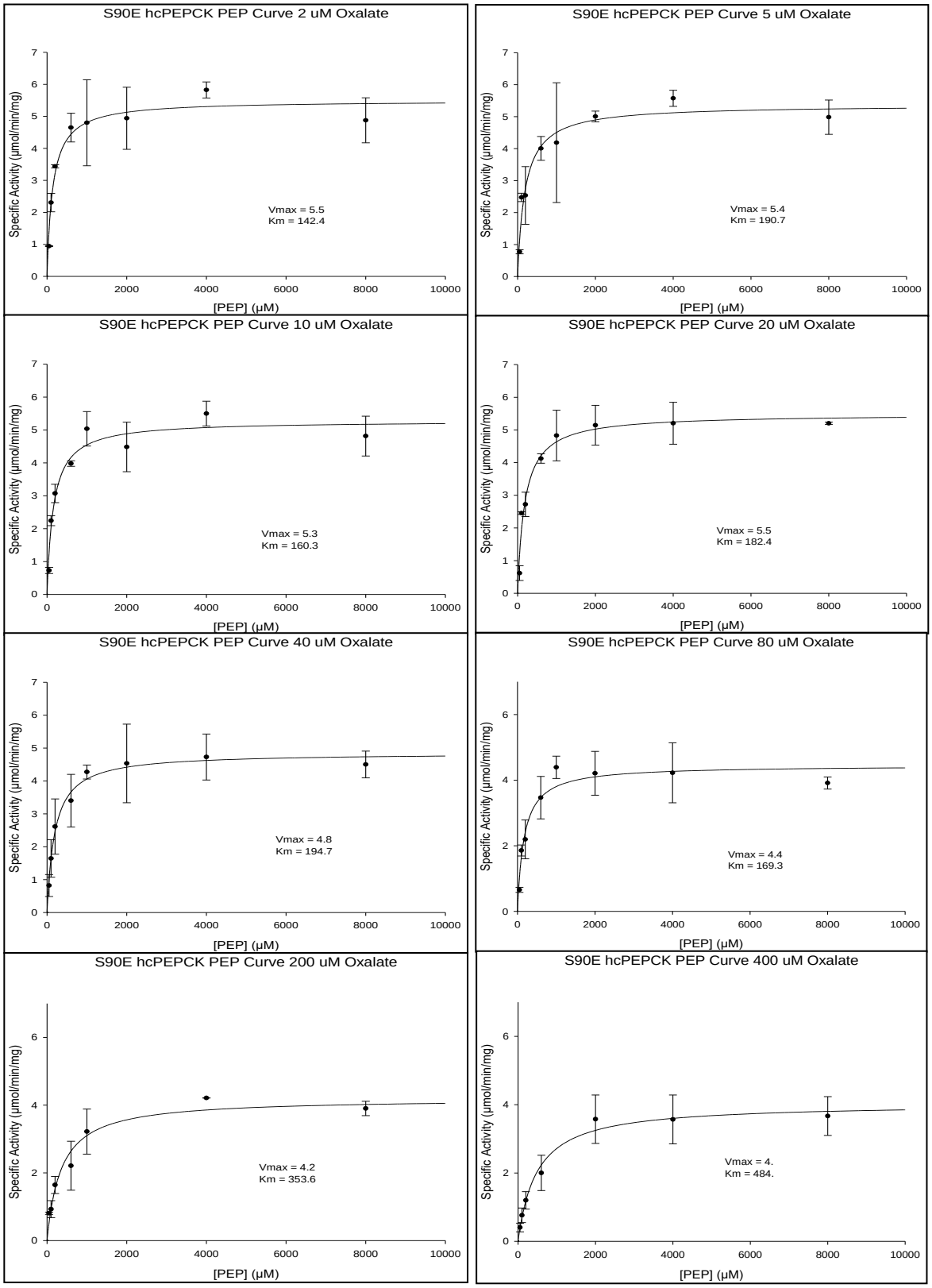
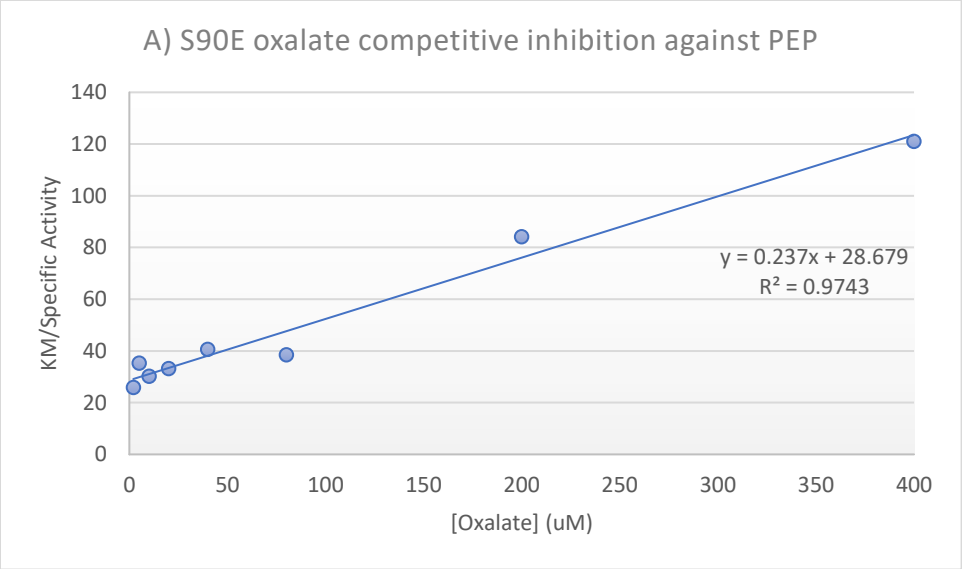


Figure B7: S90E hcPEPCK PEP Michaelis-Menten curves under increasing oxalate concentrations.



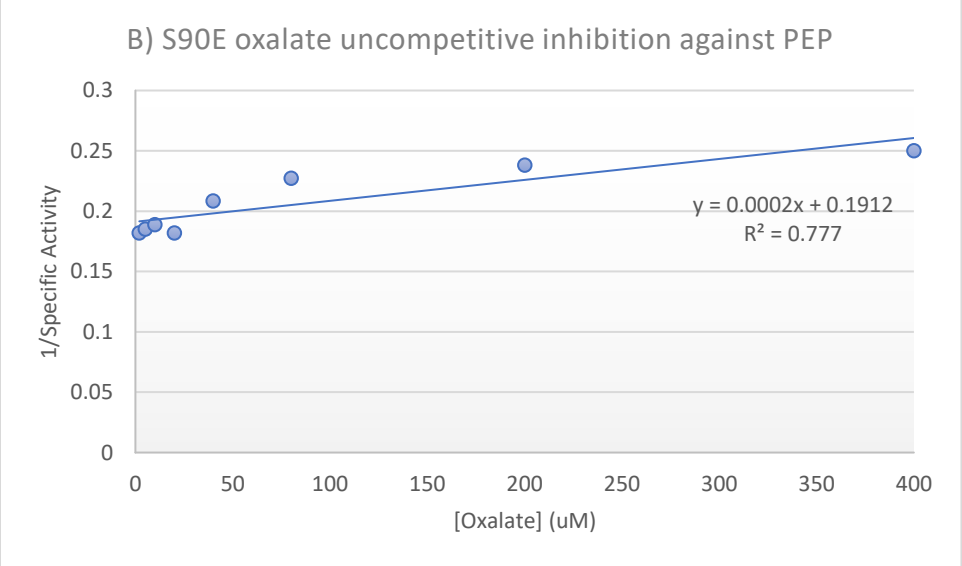
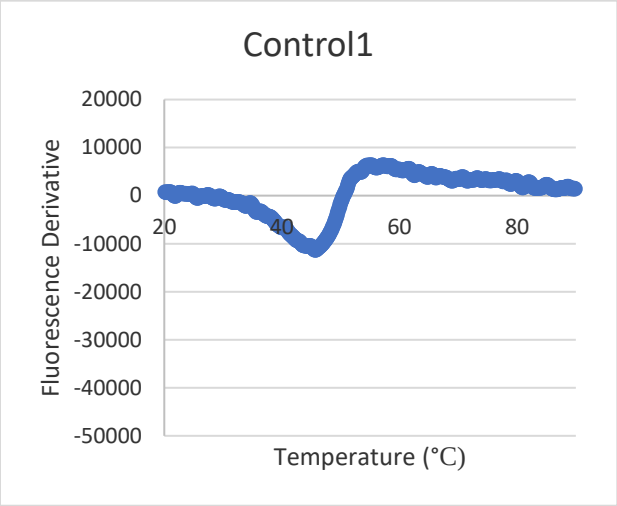
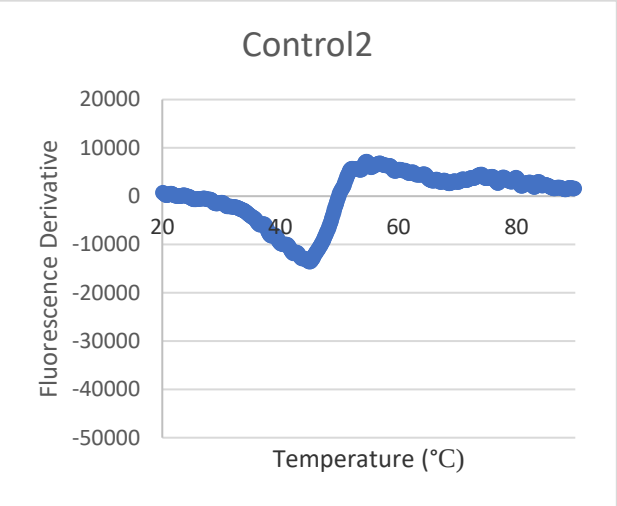
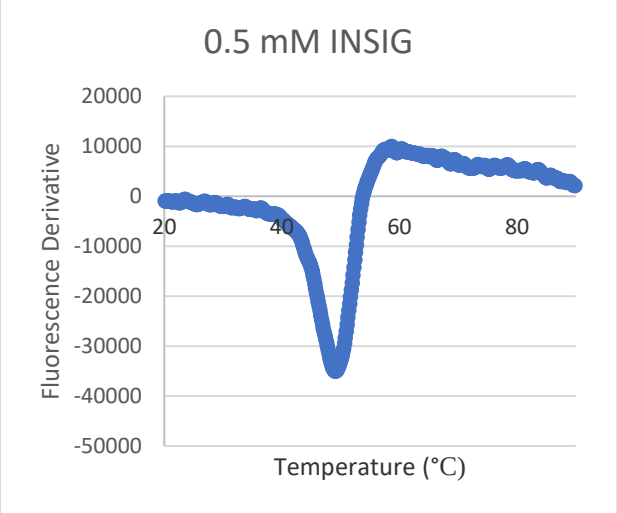


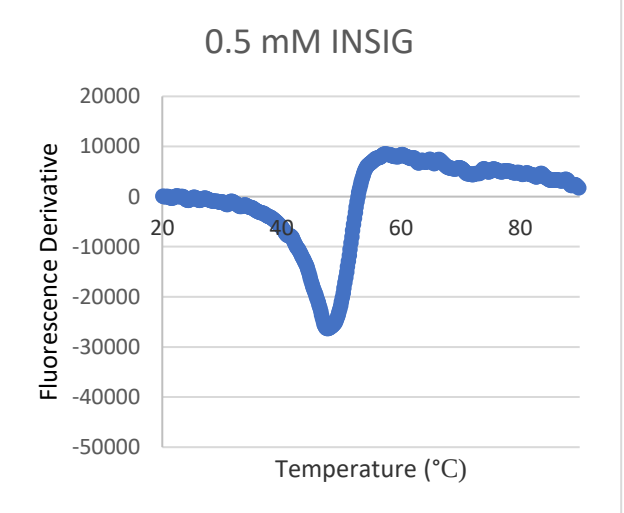
Figure B8: Kinetic re-plots of oxalate concentration vs specific activity for S90E hcPEPCK in the reverse direction of catalysis. Plot A uses the ratio of KM/specific activity to determine competitive inhibition; Plot B uses the ratio of 1/specific activity to determine uncompetitive inhibition.

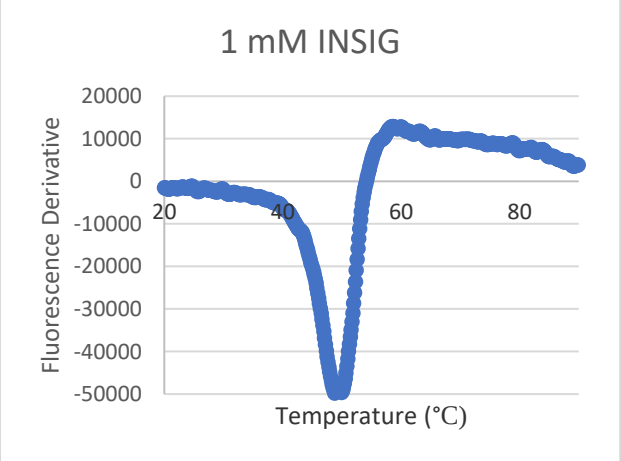
Appendix C – Raw data for DSF assays

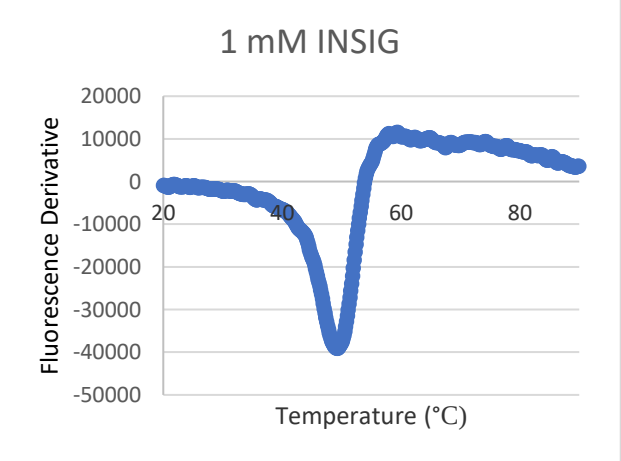


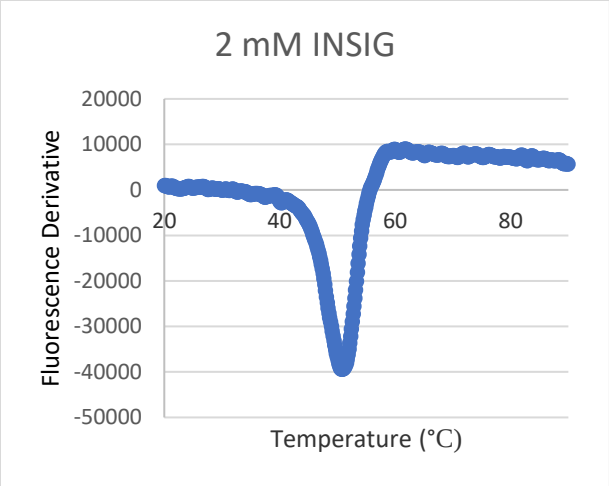


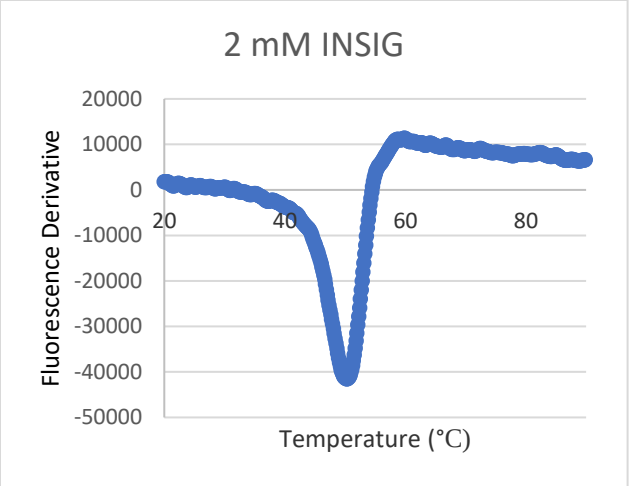


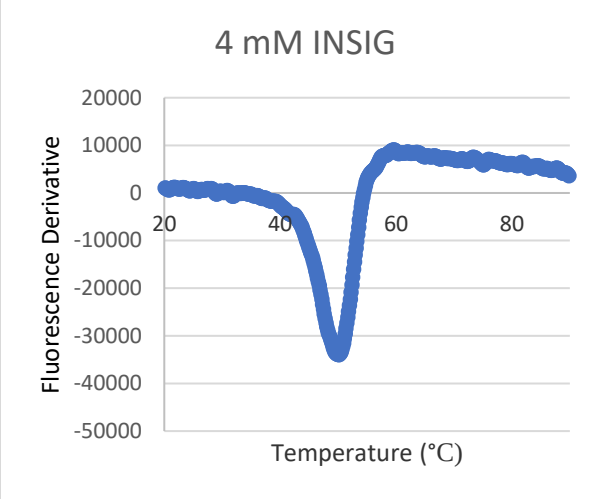


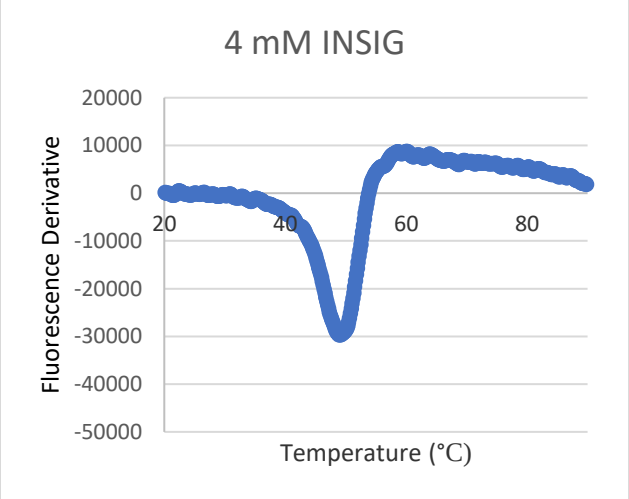


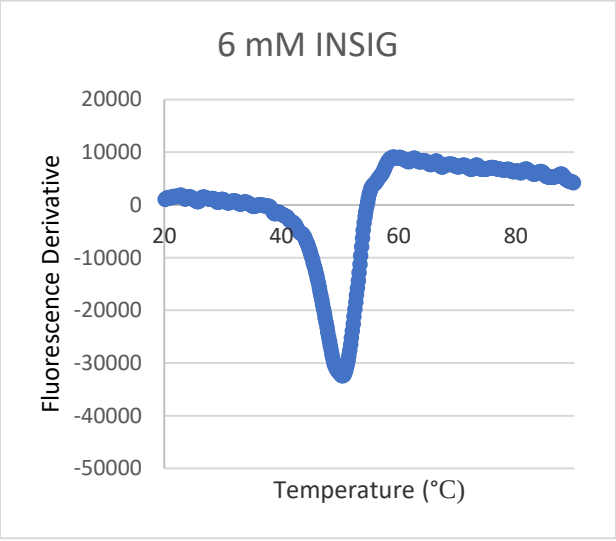


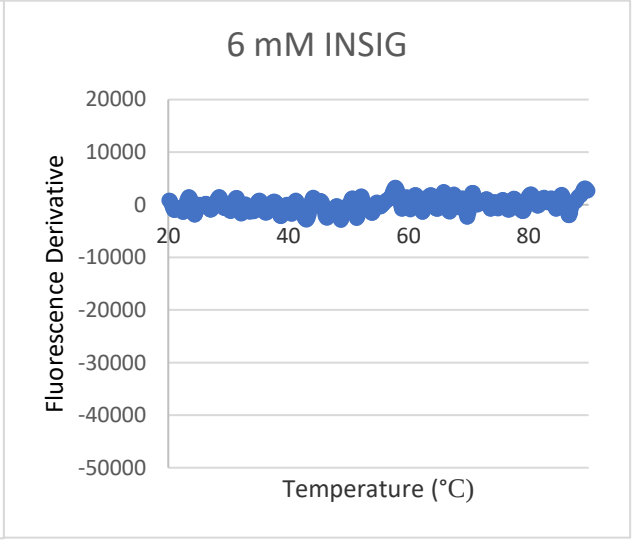

2 mM INSIG
Fluorescence Derivative
20000
10000
0
-10000
-20000
-30000
-40000
-50000
20
40
60
80
Temperature (°C)


2 mM INSIG
Fluorescence Derivative
20000
10000
0
-10000
-20000
-30000
-40000
-50000
20
40
60
80
Temperature (°C)


4 mM INSIG
Fluorescence Derivative
20000
10000
0
-10000
-20000
-30000
-40000
-50000
20
40
60
80
Temperature (°C)


4 mM INSIG
Fluorescence Derivative
20000
10000
0
-10000
-20000
-30000
-40000
-50000
20
40
60
80
Temperature (°C)


6 mM INSIG
Fluorescence Derivative
20000
10000
0
-10000
-20000
-30000
-40000
-50000
20
40
60
80
Temperature (°C)


6 mM INSIG
Fluorescence Derivative
20000
10000
0
-10000
-20000
-30000
-40000
-50000
20
40
60
80
Temperature (°C)

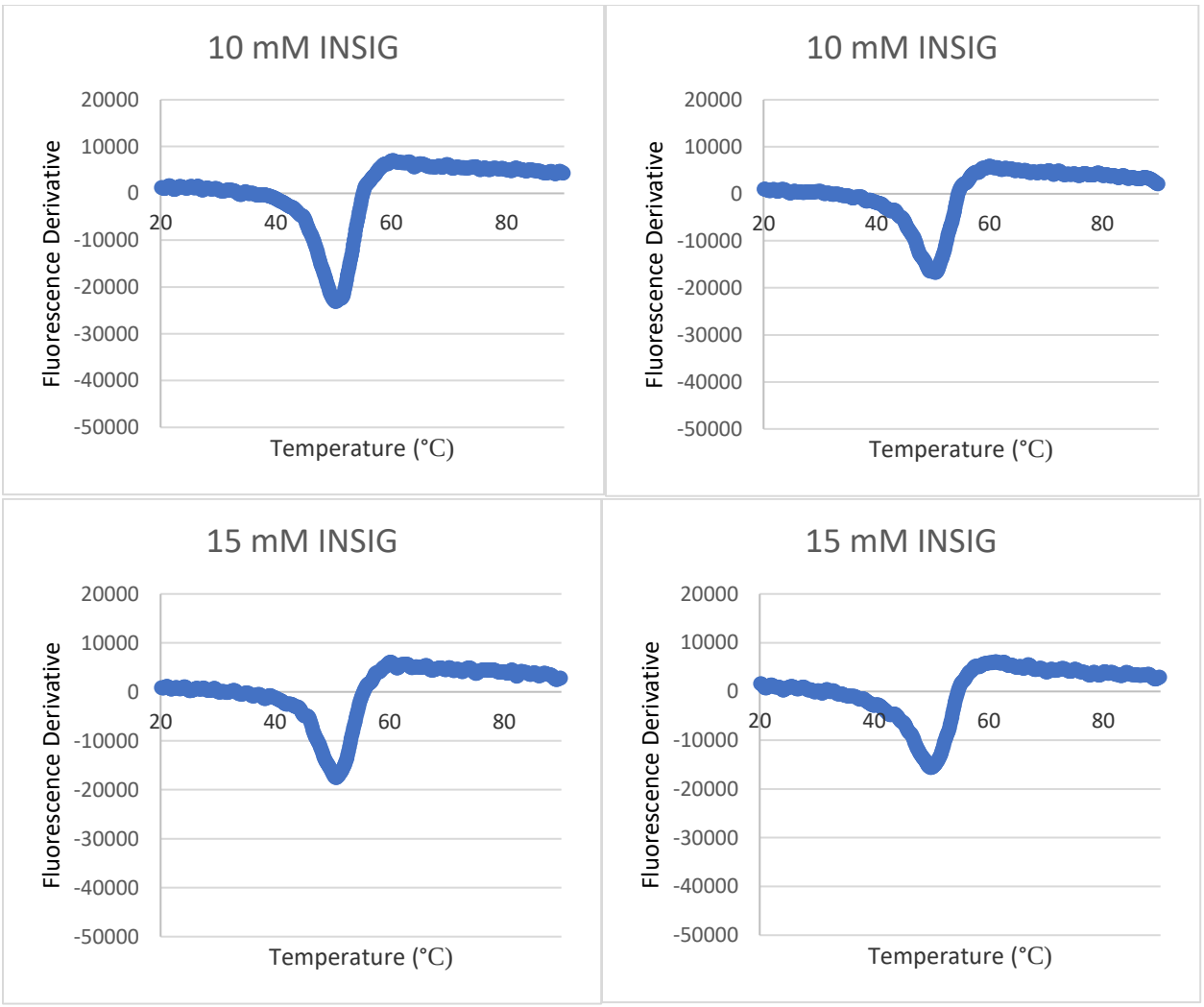
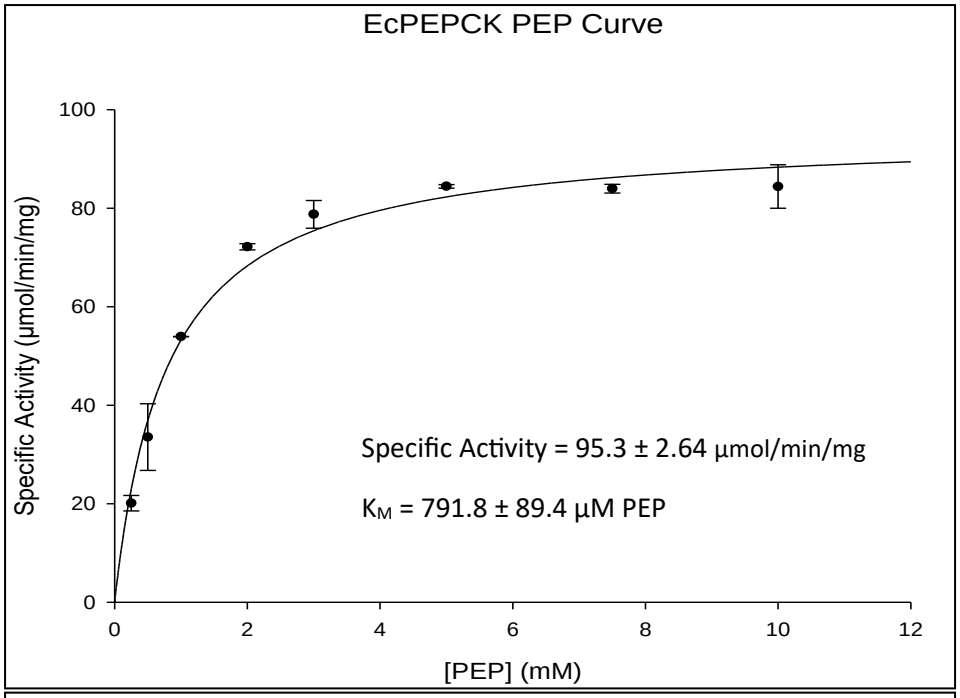


Figure C1: DSF fluorescence derivative data for S90E hcPEPCK unfolding under increasing INSIG concentrations

Appendix D – Raw kinetics data for EcPEPCK

EcPEPCK kinetic characterization of activity and substrate affinities



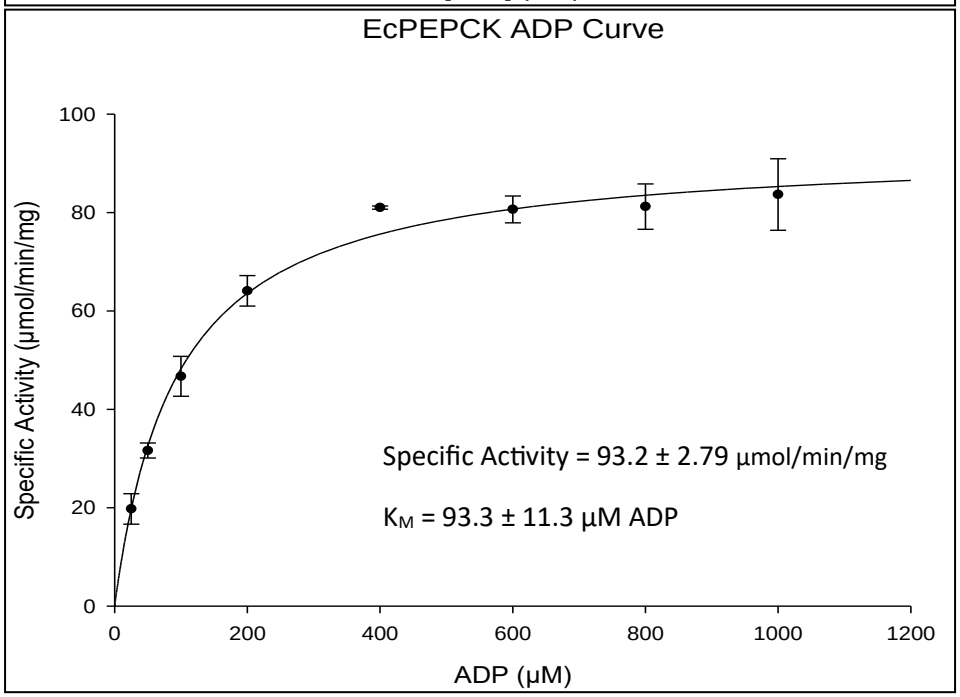


Figure D1: Michaelis-Menten curves for the characterization of EcPEPCK activity in the PEP to OAA direction.

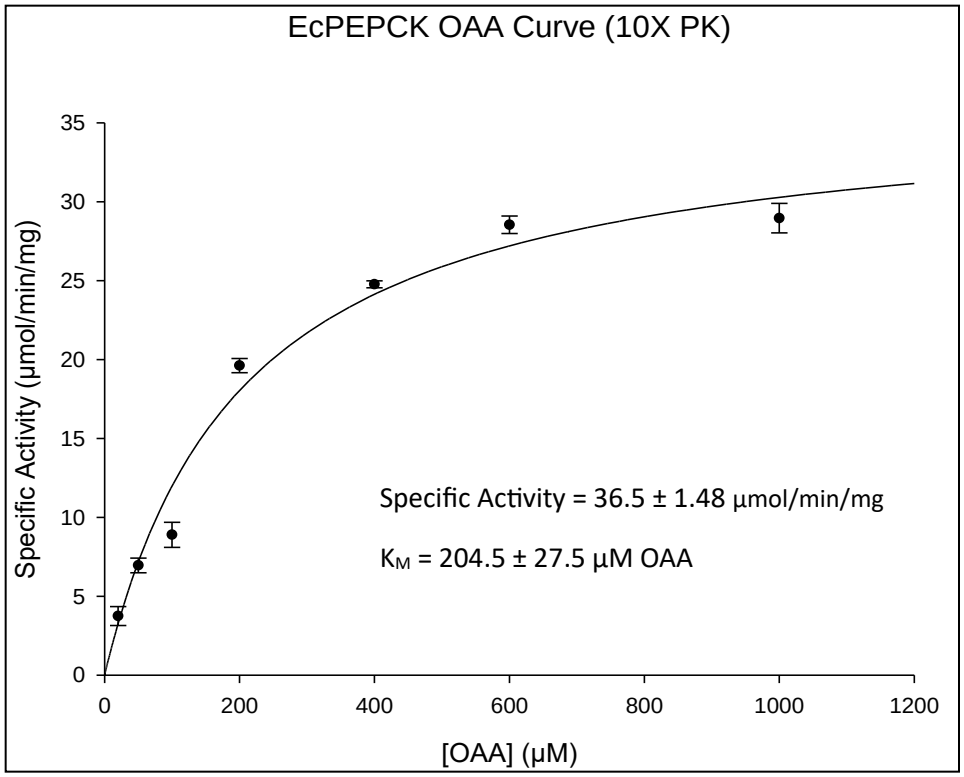


Figure D2: Michaelis-Menten curve for the characterization of EcPEPCK activity in the OAA to PEP direction.

Divalent Cation Activation of EcPEPCK

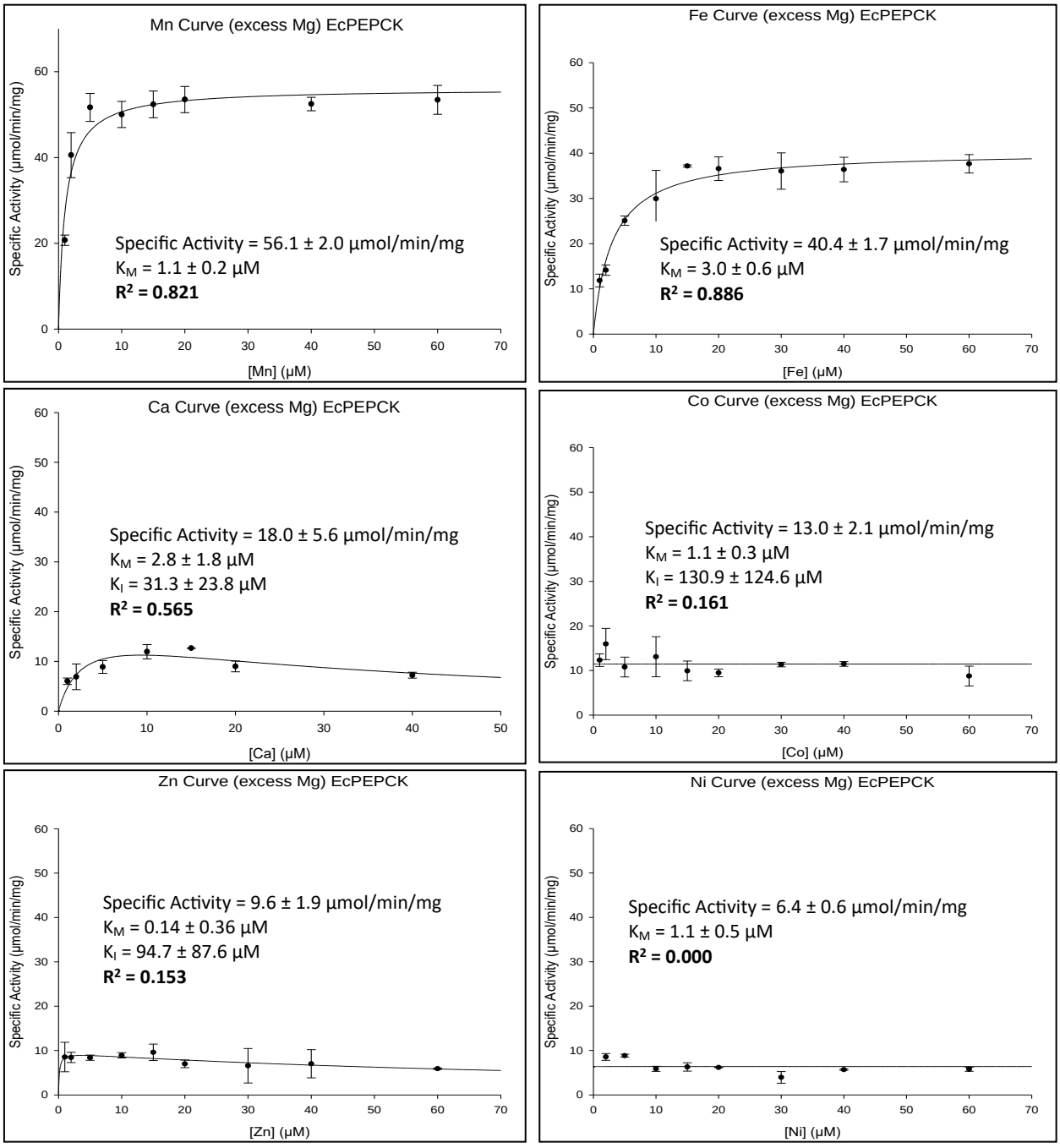


Figure D3: Divalent cation activation of EcPEPCK in the presence of saturating $MgCl_2$

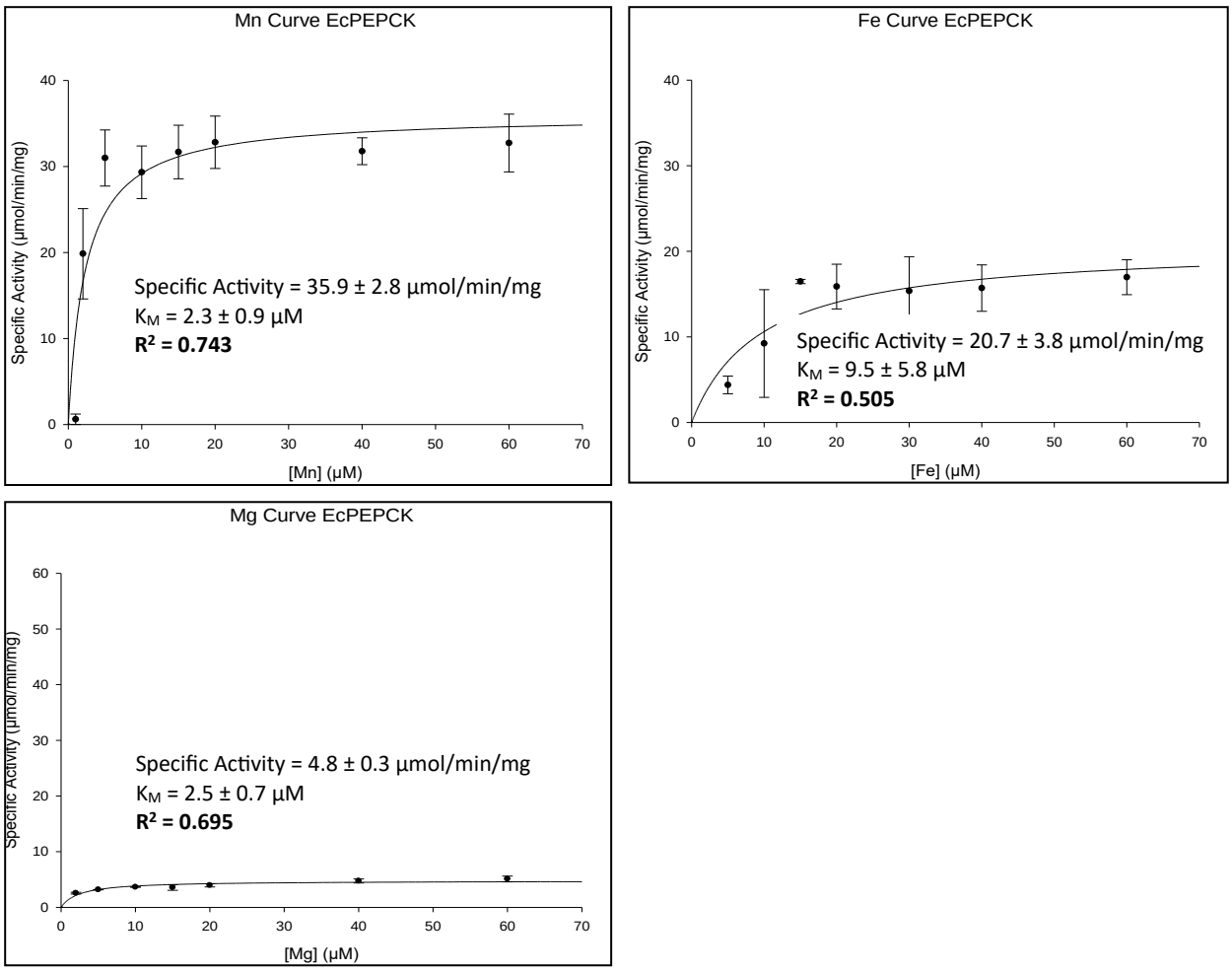


Figure D4: Divalent cation activation of EcPEPCK as the sole metal.

Determining Oxalate inhibition of EcPEPCK in both directions of catalysis

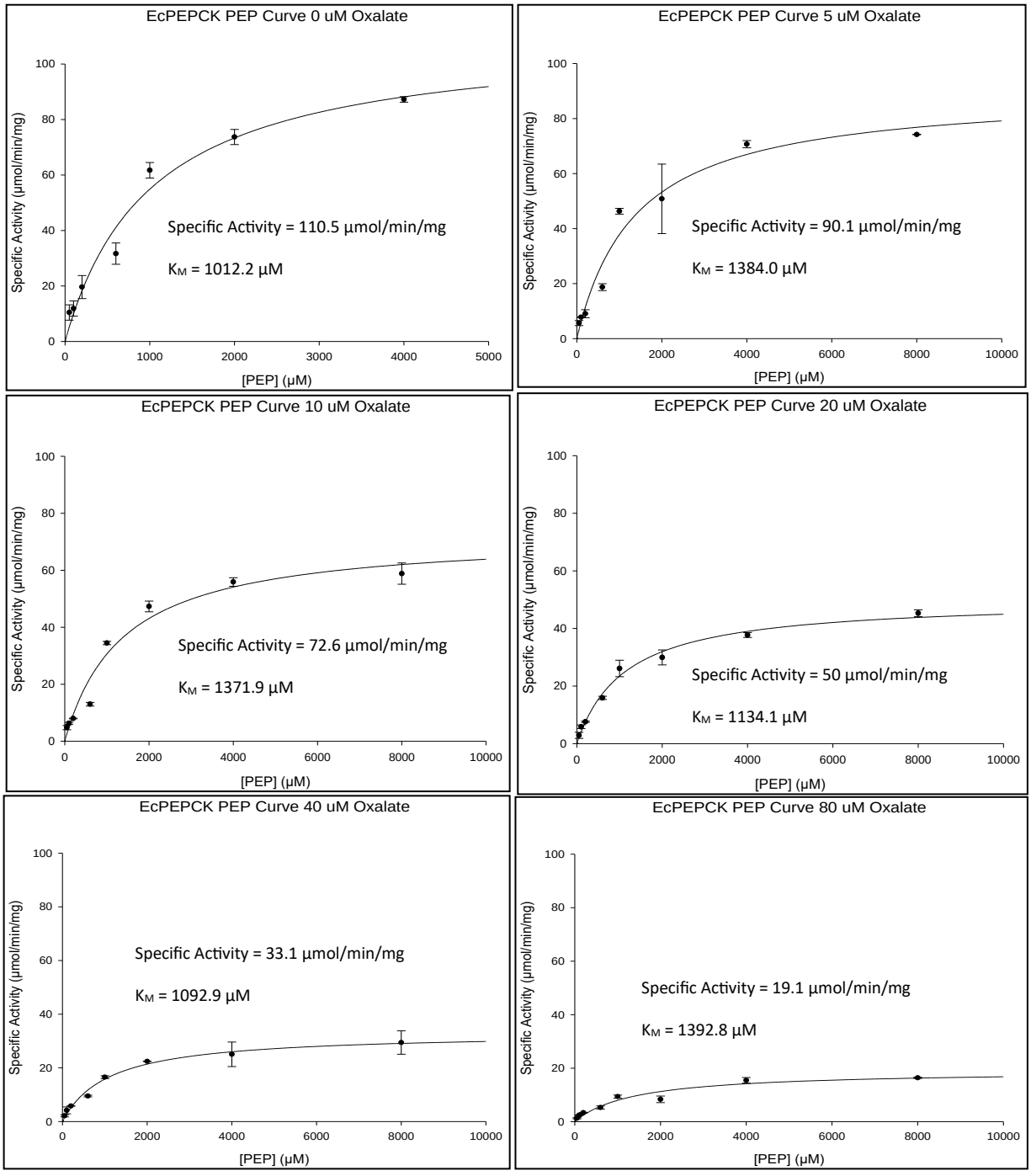
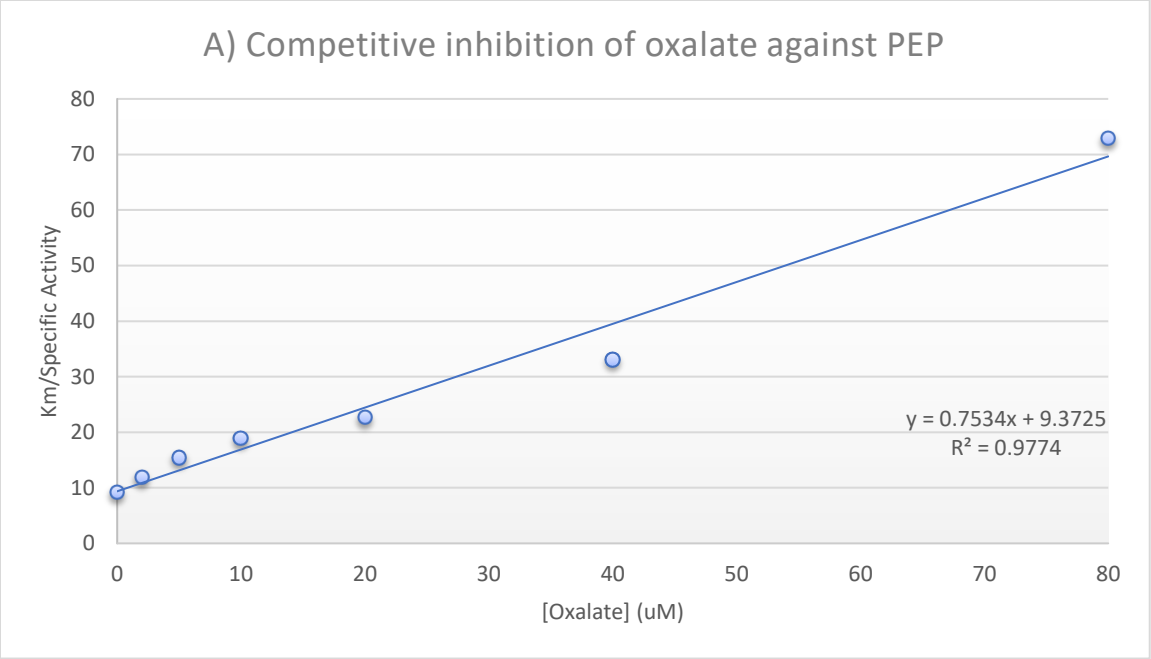


Figure D5: EcPEPCK PEP Michaelis-Menten curves under increasing oxalate concentrations.



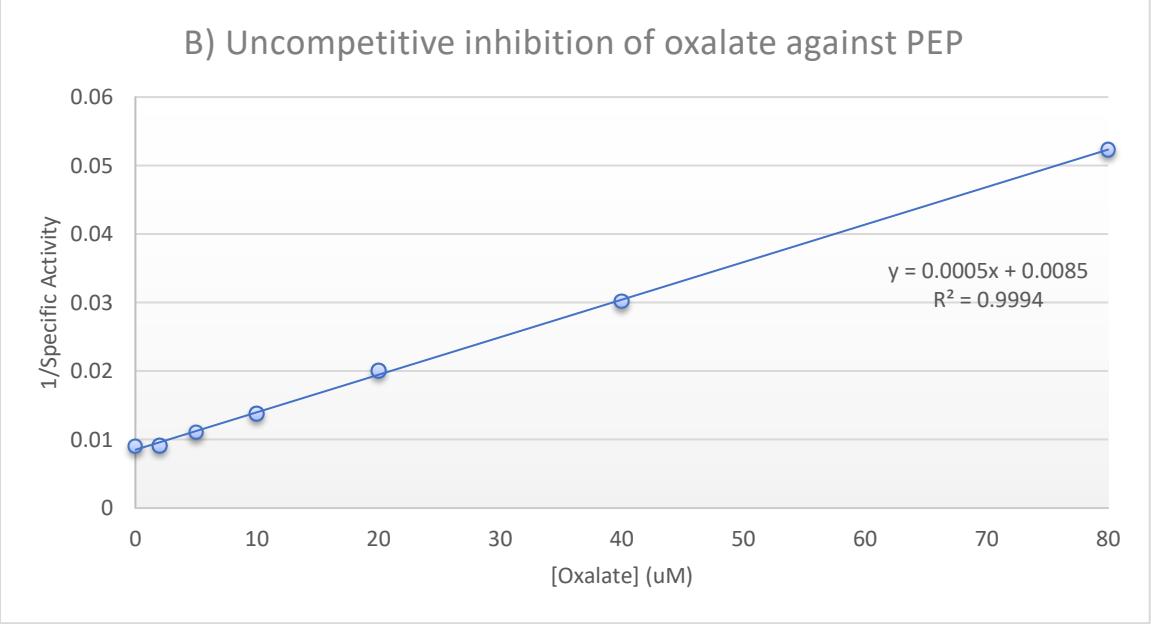


Figure D6: Kinetic re-plots of oxalate concentration vs specific activity for EcPEPCK in the reverse direction of catalysis. Plot A uses the ratio of KM/specific activity to determine competitive inhibition; Plot B uses the ratio of 1/specific activity to determine uncompetitive inhibition.

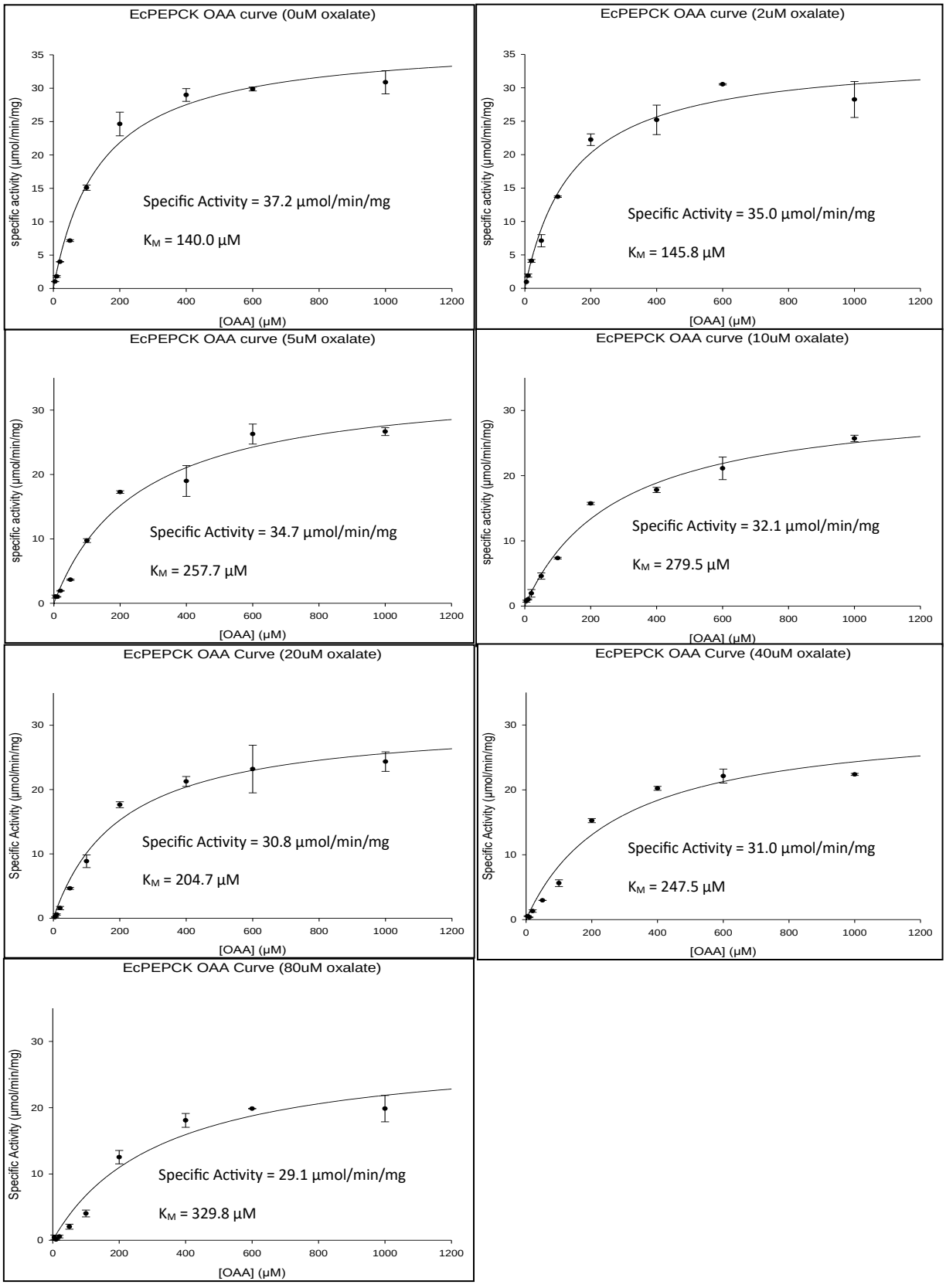
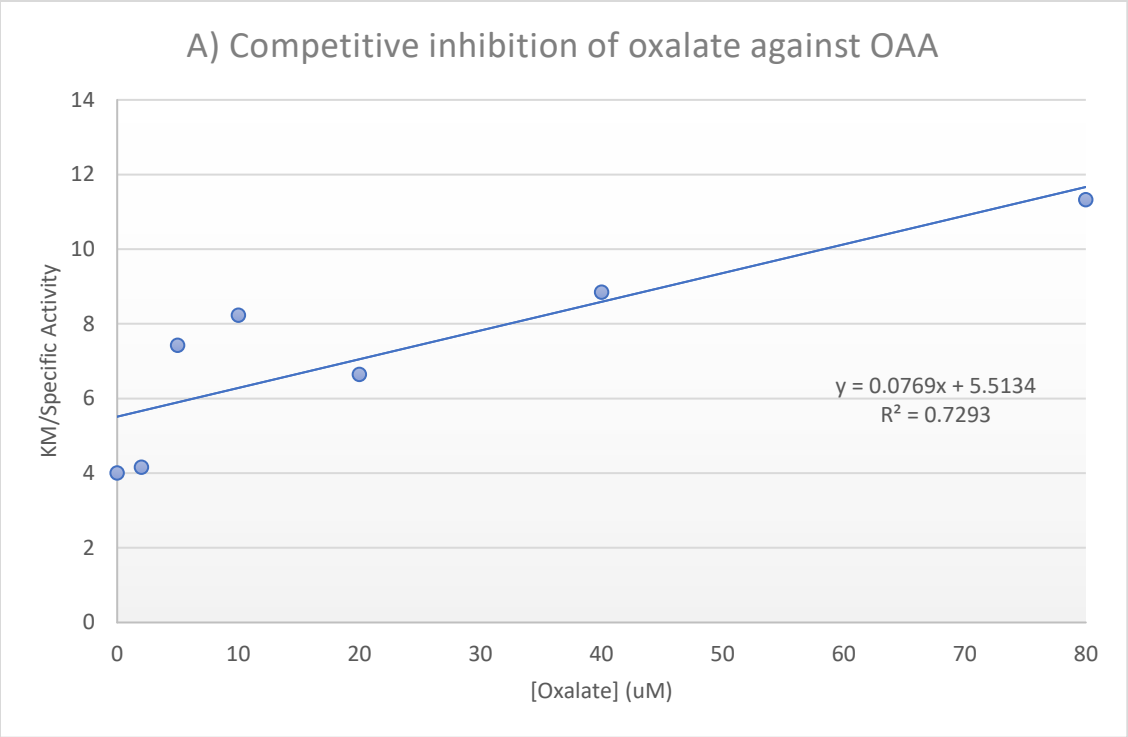


Figure D7: EcPEPCK OAA Michaelis-Menten curves under increasing oxalate concentrations.



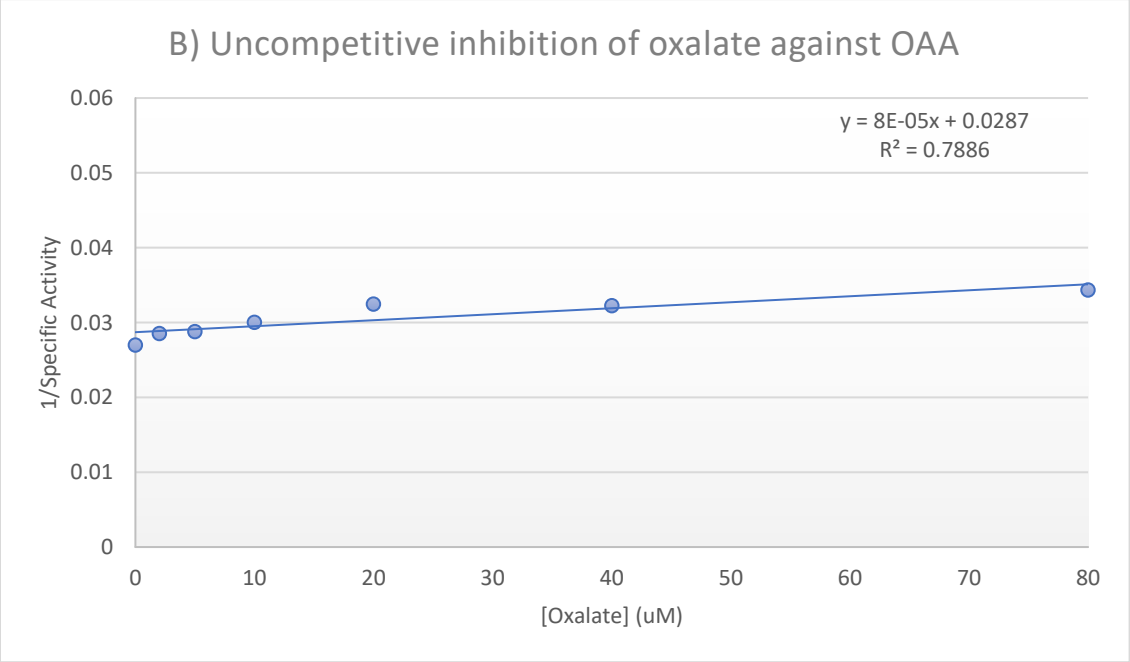


Figure D8: Kinetic re-plots of oxalate concentration vs specific activity for EcPEPCK in the forward direction of catalysis. Plot A uses the ratio of KM/specific activity to determine competitive inhibition; Plot B uses the ratio of 1/specific activity to determine uncompetitive inhibition.

Determining β-sulfopyruvate inhibition of EcPEPCK in the forward direction of catalysis

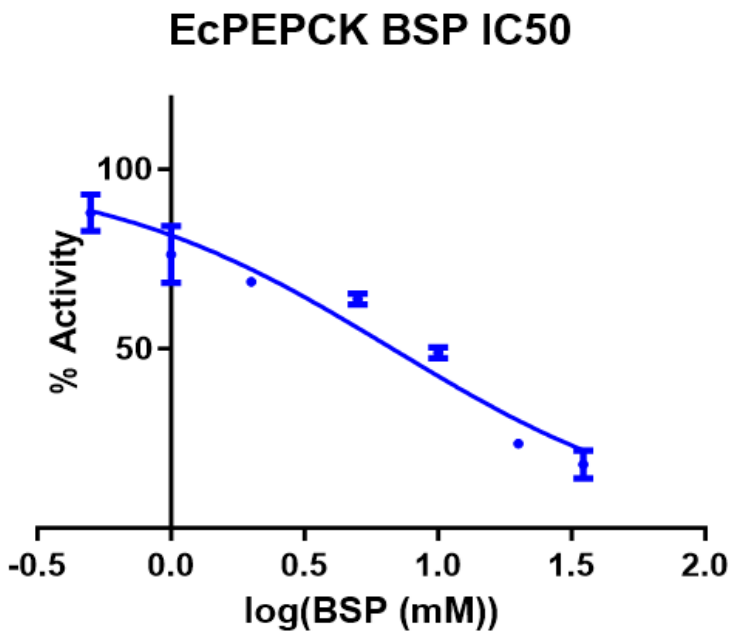


Figure D9: IC50 plot of BSP inhibition of EcPEPCK. Inhibition is plotted as percent activity against log BSP concentration using GraphPadPrism. IC50 was used to estimate the K_I for BSP for the complete set of inhibition curves.

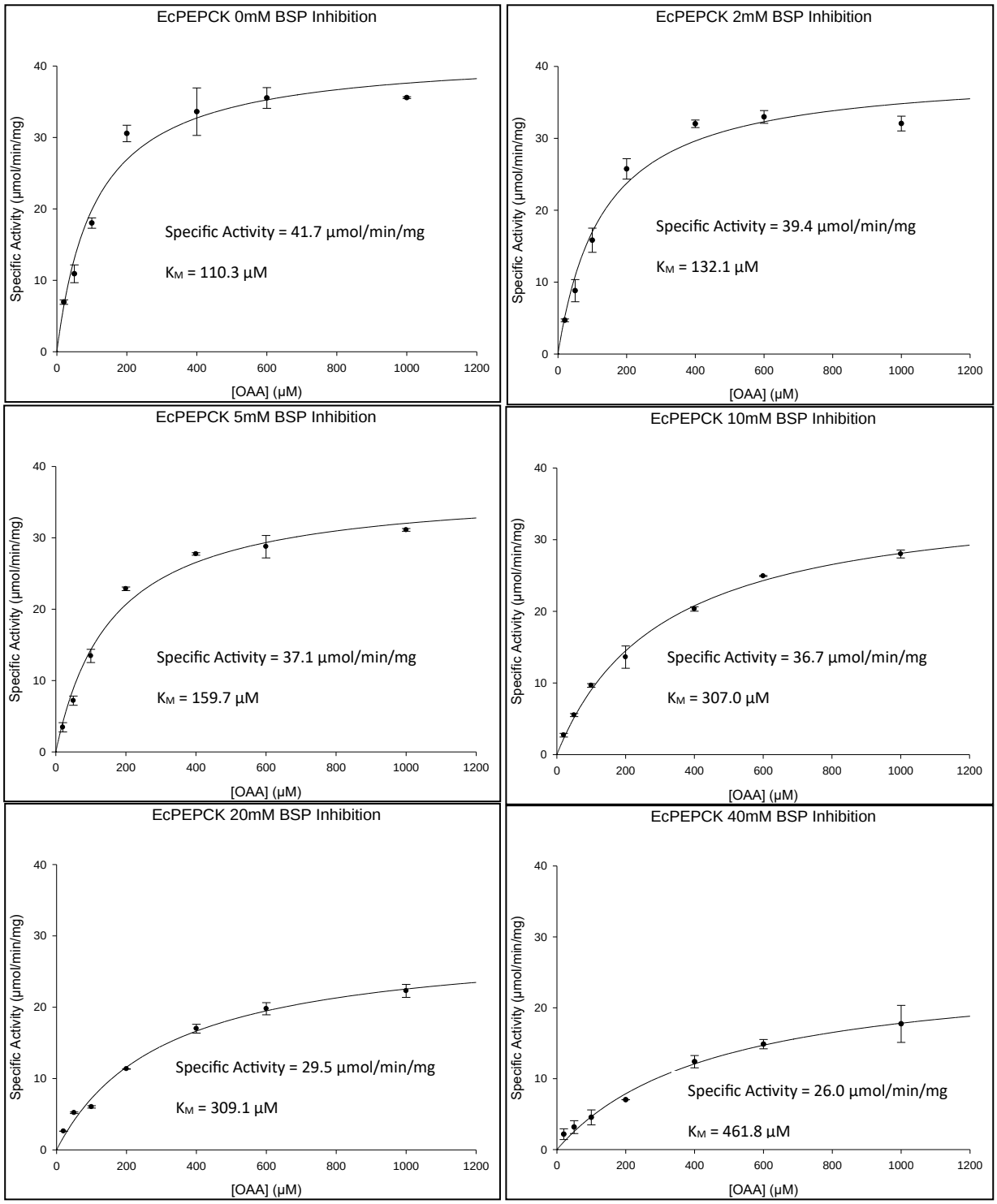
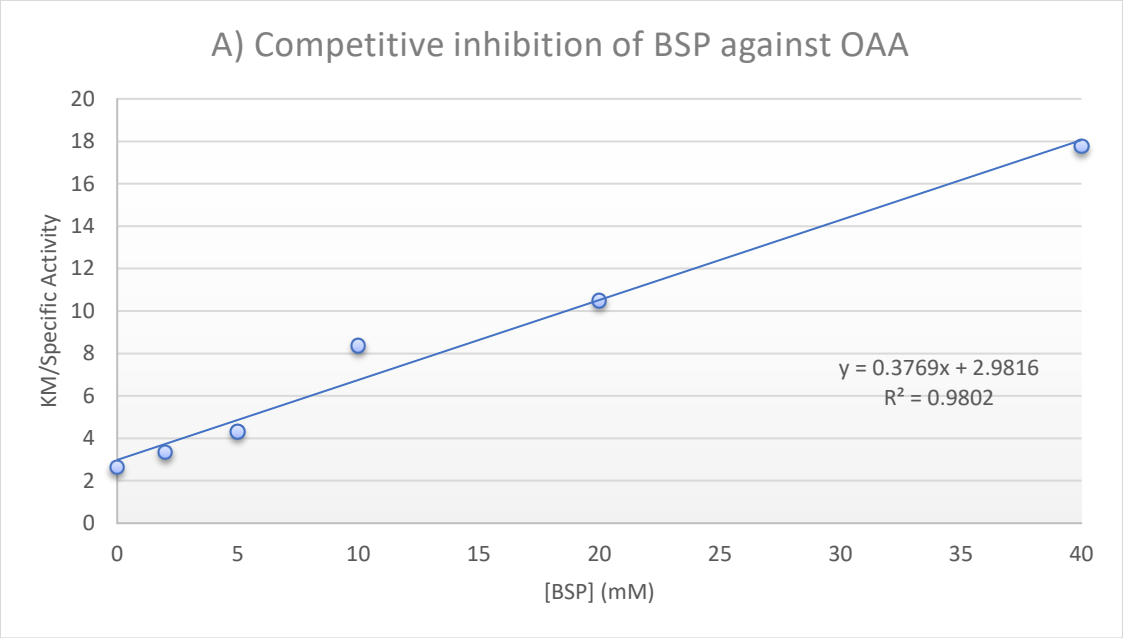


Figure D10: EcPEPCK OAA Michaelis-Menten curves under increasing βSP concentrations.



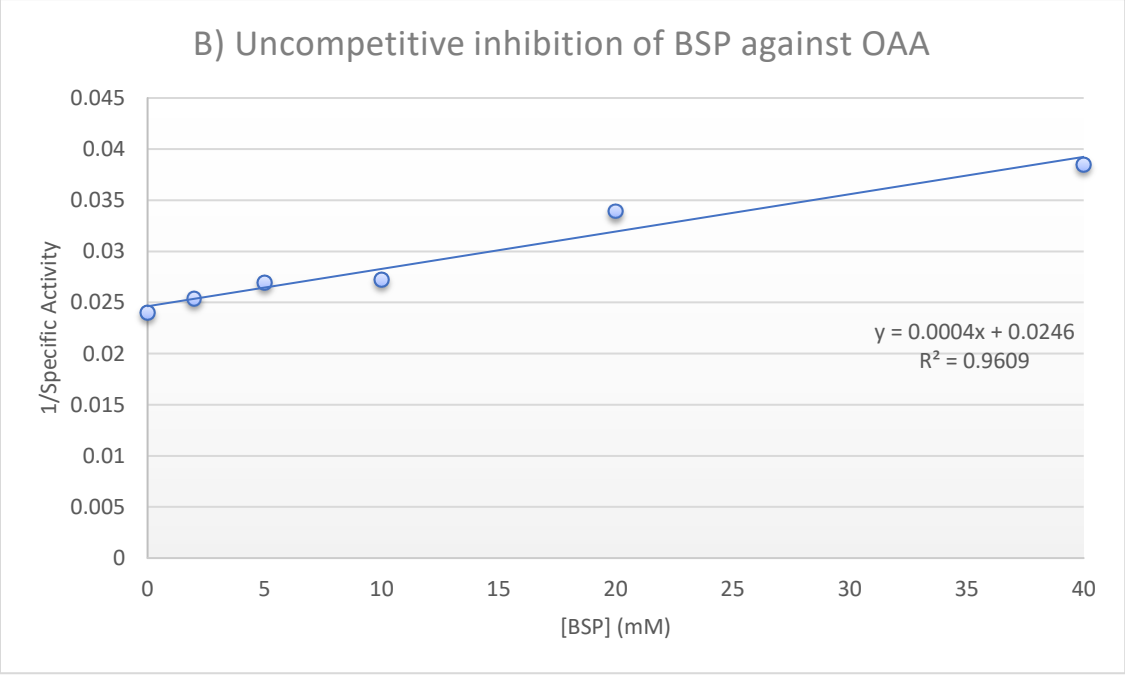


Figure D11: Kinetic re-plots of BSP concentration vs specific activity for EcPEPCK in the forward direction of catalysis. Plot A uses the ratio of KM/specific activity to determine competitive inhibition; Plot B uses the ratio of 1/specific activity to determine uncompetitive inhibition.

Determining phosphoglycolate inhibition of EcPEPCK in the reverse direction of catalysis

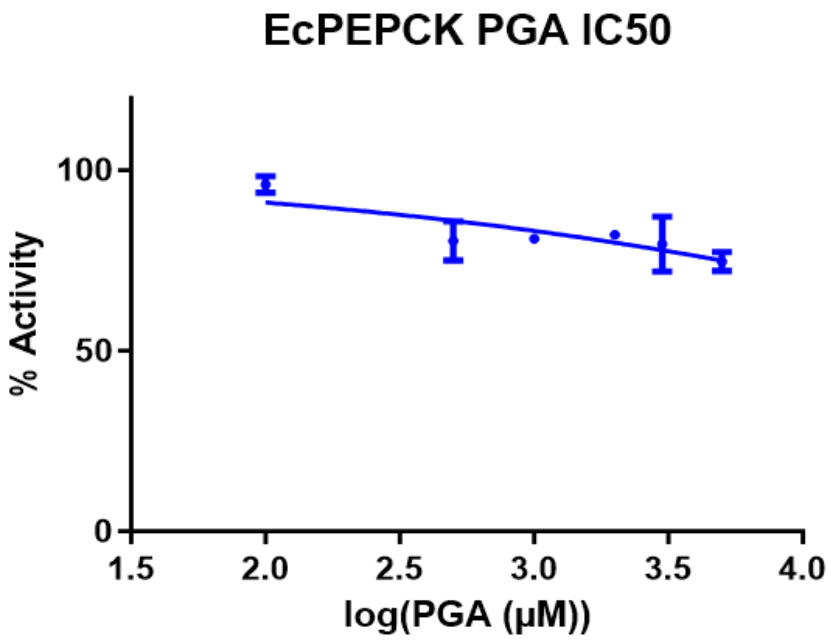


Figure D12: IC50 plot of PGA inhibition of EcPEPCK. Inhibition is plotted as percent activity against the log of the PGA concentration, and fit using non-linear regression on GraphPadPrism.

Appendix E – Crystallography Data Tables for hcPEPCK & EcPEPCK

Table E1: Crystallographic data of hcPEPCK co-crystal structures*

	WT hcPEPCK + GTP + Oxalate	S90E hcPEPCK + GTP + Oxalate	WT hcPEPCK + GTP
Wavelength (Å)	**1.541**	**1.541**	**0.9538**
Space group	P 1 21 1	P 1 21 1	P 1 21 1
Unit Cell Dimensions	a = 45.125 b = 114.507 c = 114.526 α = γ = 90.00° β = 91.61°	a = 44.803 b = 114.001 c = 114.271 α = γ = 90.00° β = 91.76°	a = 45.120 b = 112.766 c = 60.336 α = γ = 90.00 β = 107.74
Resolution Limits (Å)	35.43 – **2.10**	34.14 – **2.45**	57.47 – **1.70**
No. of unique reflections	63891	37286	59759
Completeness (%)	98.81 (96.65)	92.38 (91.69)	99.55 (96.41)
Redundancy	4.4 (4.0)	5.0 (4.8)	6.8 (5.4)
I/sigma	9.3 (1.6)	7.77 (1.45)	16.1 (0.8)
CC1/2	0.987 (0.638)	0.984 (0.537)	0.998 (0.659)
Rpim	0.071 (0.498)	0.087 (0.500)	0.030 (0.448)
Rmerge	0.125 (0.657)	0.153 (0.865)	0.072 (0.956)
No. of ASU molecules	2	2	1
Rfree	0.267 (0.332)	0.266 (0.411)	0.229 (0.291)
Rwork	0.212 (0.307)	0.219 (0.340)	0.188 (0.249)
Average B factor (Å^2)			
Protein	44.61	44.22	34.06
Water	45.18	42.17	27.66
Mn	35.66	38.45	28.47
GTP	45.56	45.60	31.09
Oxalate	30.90	37.91	-
Estimated coordinate error based on maximum likelihood (Å)	0.212	0.303	0.125
Bond length rmsd (Å)	0.003	0.009	0.010
Bond angle rmsd (deg)	1.004	1.190	1.210
Molprobity statistics (score, percentile)	1.45,98th	1.69,99th	1.35,96th
Ramachandran statistics (%) (favoured/outliers)	95.03/0.25	95.59/0.08	97.00/0.17

**Values in parentheses represent the highest resolution shell*

Table E2: Crystallographic data of EcPEPCK co-crystal structures

	ATP + Oxalate	ATP + BSP	ADP + PGA	ATP turnover
Wavelength (Å)	**1.541**	**0.9686**	**1.541**	**1.1271**
Space group	P1 21 1	C1 2 1	P1 21 1	C1 2 1
Unit Cell Dimensions	a = 54.236 b = 194.846 c = 54.162 α = γ = 90.00° β = 114.51°	a = 123.998 b = 93.219 c = 46.054 α = γ = 90.00° β = 95.89°	a = 54.527 b = 195.183 c = 54.572 α = γ = 90.00° β = 115.09°	a = 124.598 b = 93.778 c = 46.203 α = γ = 90.00° β = 95.86°
Resolution Limits (Å)	39.29-**2.00**	74.37-**1.45**	49.47-**1.73**	74.78 – **1.70**
No. of unique reflections	64968	85696	99951	53493
Completeness (%)	99.38 (86.00)	98.11 (80.74)	98.22 (93.03)	96.90 (93.70)
Redundancy	5.4 (3.2)	3.3 (1.5)	5.8 (5.4)	3.3 (3.4)
I/sigma	11.88 (1.68)	8.73 (3.18)	13.84 (1.93)	8.60 (1.00)
CC1/2	0.988 (0.758)	0.998 (0.944)	0.991 (0.750)	0.981 (0.694)
Rpim	0.060 (0.354)	0.033 (0.154)	0.044 (0.381)	0.084 (0.263)
Rmerge	0.137 (0.486)	0.134 (0.243)	0.096 (0.777)	0.124 (0.409)
No. of ASU molecules	2	1	2	1
Rfree	0.204 (0.242)	0.188 (0.275)	0.225 (0.424)	0.222 (0.260)
Rwork	0.173 (0.190)	0.169 (0.259)	0.187 (0.358)	0.195 (0.273)
Average B factor (Å^2)				
Protein	38.73	18.44	31.48	17.58
Water	35.11	38.11	32.57	36.43
Mn	31.08	11.18	33.56	10.71
Nucleotide	37.61	13.85	34.12	15.29
Ligand	23.70	16.07	33.92	-
Estimated coordinate error based on maximum likelihood (Å)	0.210	0.045	0.240	0.086
Bond length rmsd (Å)	0.006	0.011	0.008	0.008
Bond angle rmsd (deg)	0.850	1.909	1.667	1.670
Molprobity statistics (score, percentile)	1.30,99th	0.90,99th	1.43, 94th	1.09, 99th
Ramachandran statistics (%) (favoured/outliers)	96.98/0.19	98.31/0.19	96.62/0.28	96.99/0.19

**Values in parentheses represent the highest resolution shell*

Appendix F – BSP Synthesis

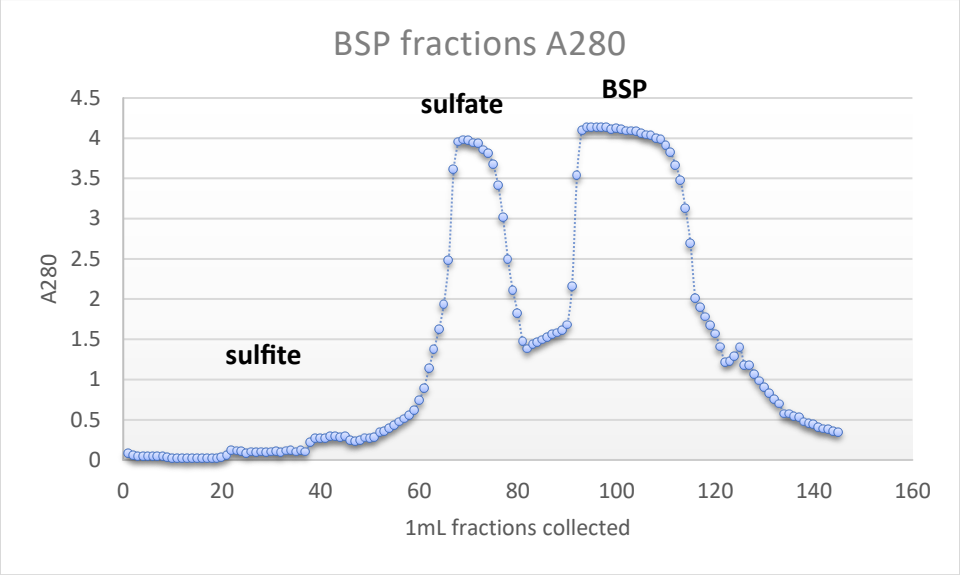


Figure F1: Chromatogram trace of synthesized BSP eluted off the Dowex1 column
Fractions eluted from the Dowex column were tested for the presence of BSP using its inherent absorbance at 280 nm. The identities of the different peaks corresponding to BSP and its contaminants, sulfate and sulfite, were approximated following the purification protocol of BSP.[87]

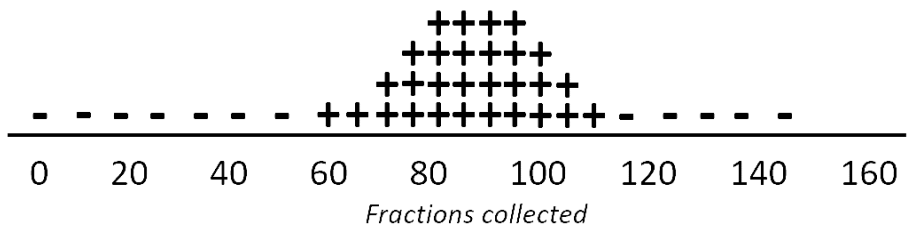


Figure F2: MDH activity for BSP at A340 nm for eluted fractions
Fractions eluted from the Dowex column were tested for BSP presence via the MDH activity assay. A positive test is indicated by + and a negative test is indicated by -.

Appendix G – ^{31}P NMR EcPEPCK Assays

^{31}P NMR Assay Controls

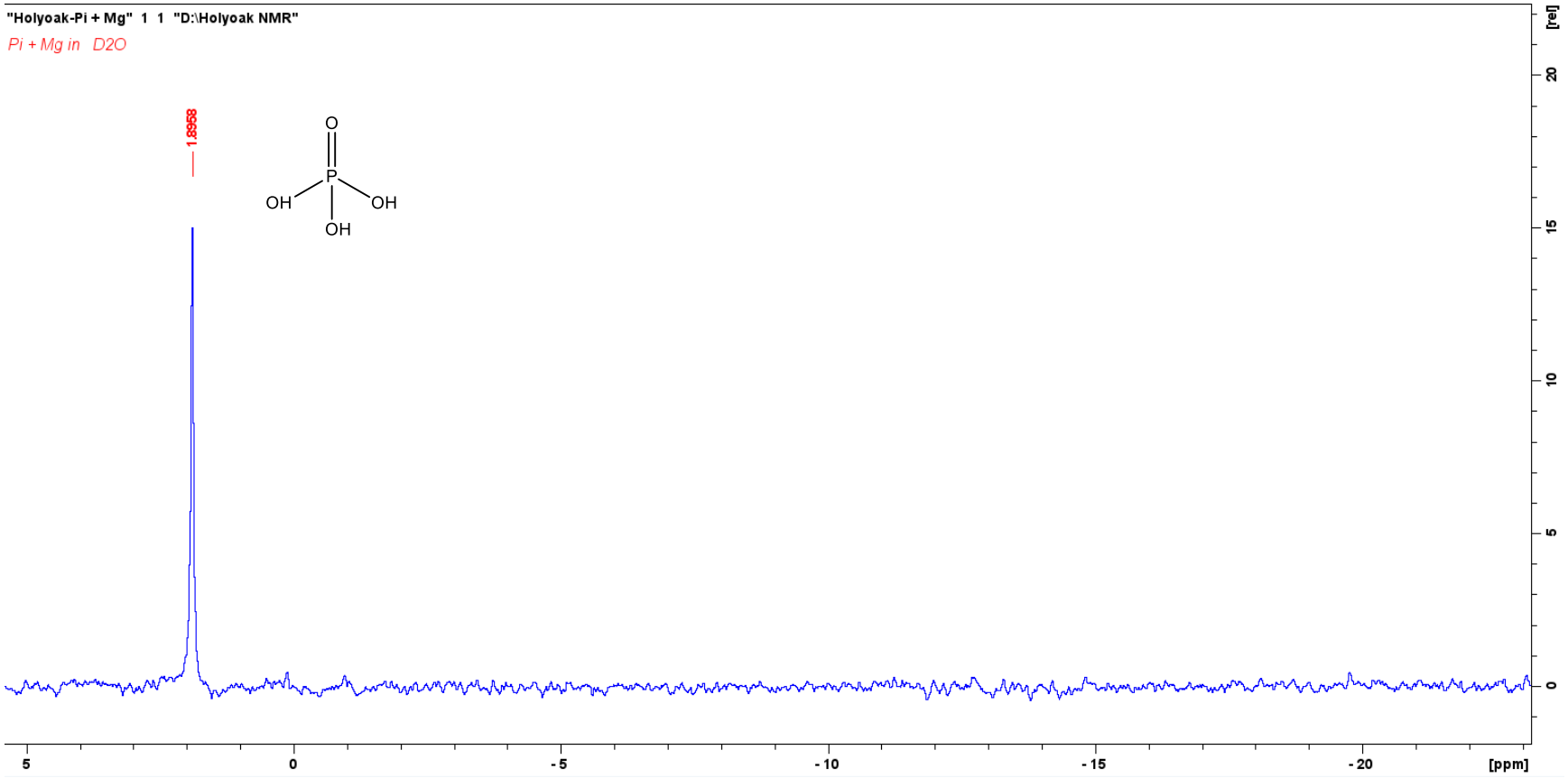


Figure G1: Inorganic phosphate (Pi) control ^{31}P NMR spectrum

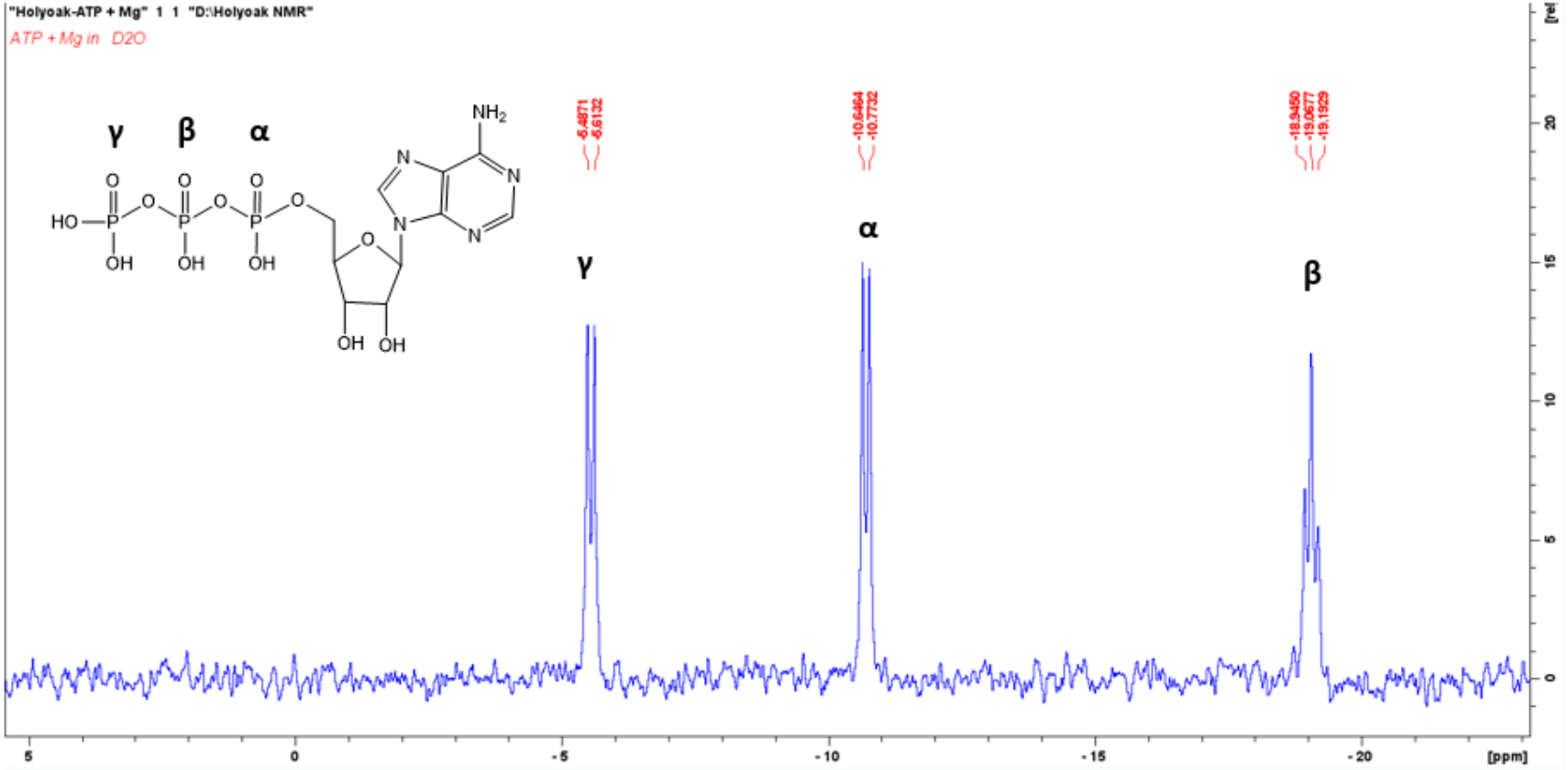


Figure G2: Adenosine 5'-triphosphate (ATP) control ^{31}P NMR spectrum

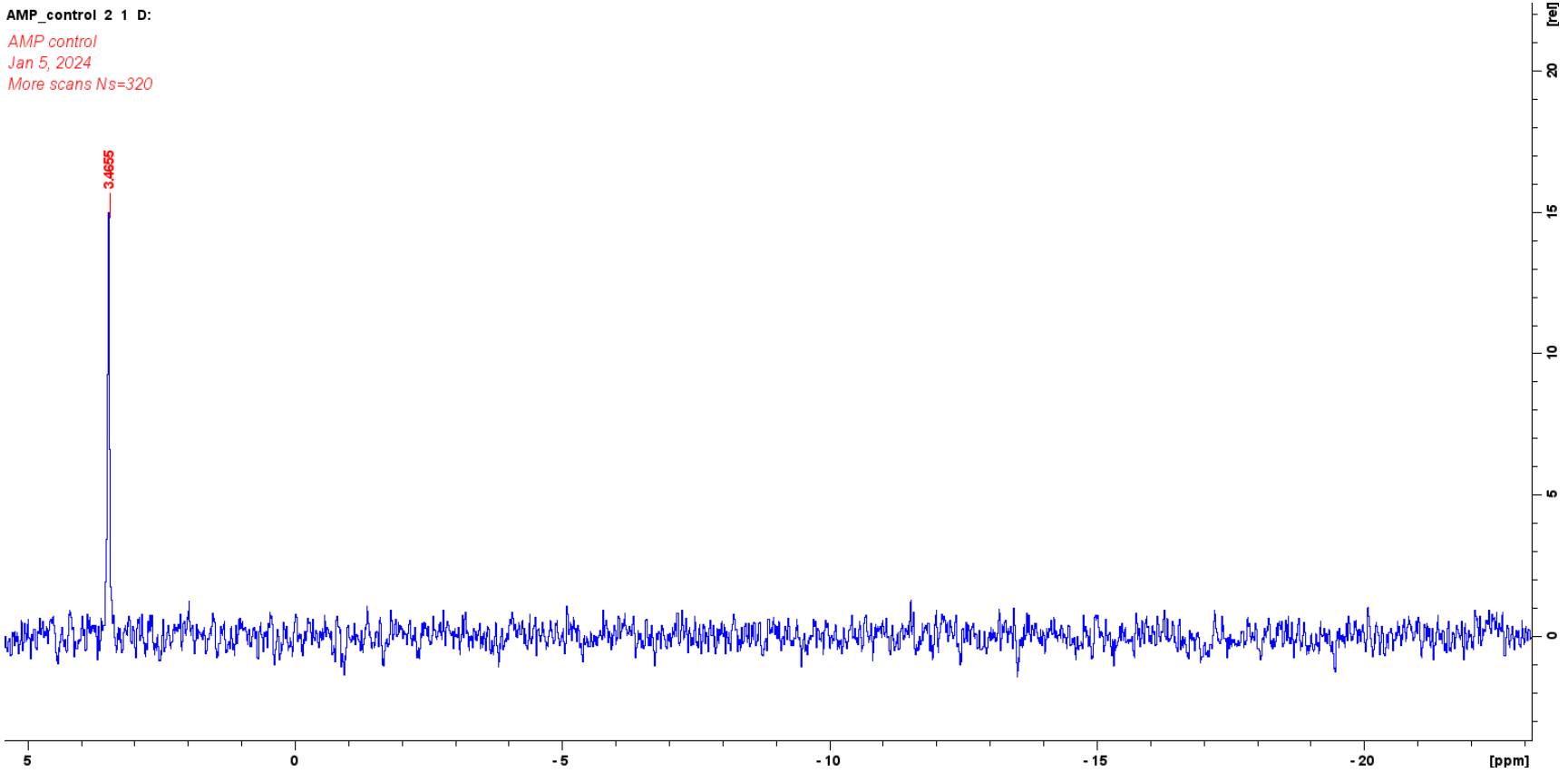


Figure G3: Adenosine 5'-monophosphate (AMP) control ^{31}P NMR spectrum

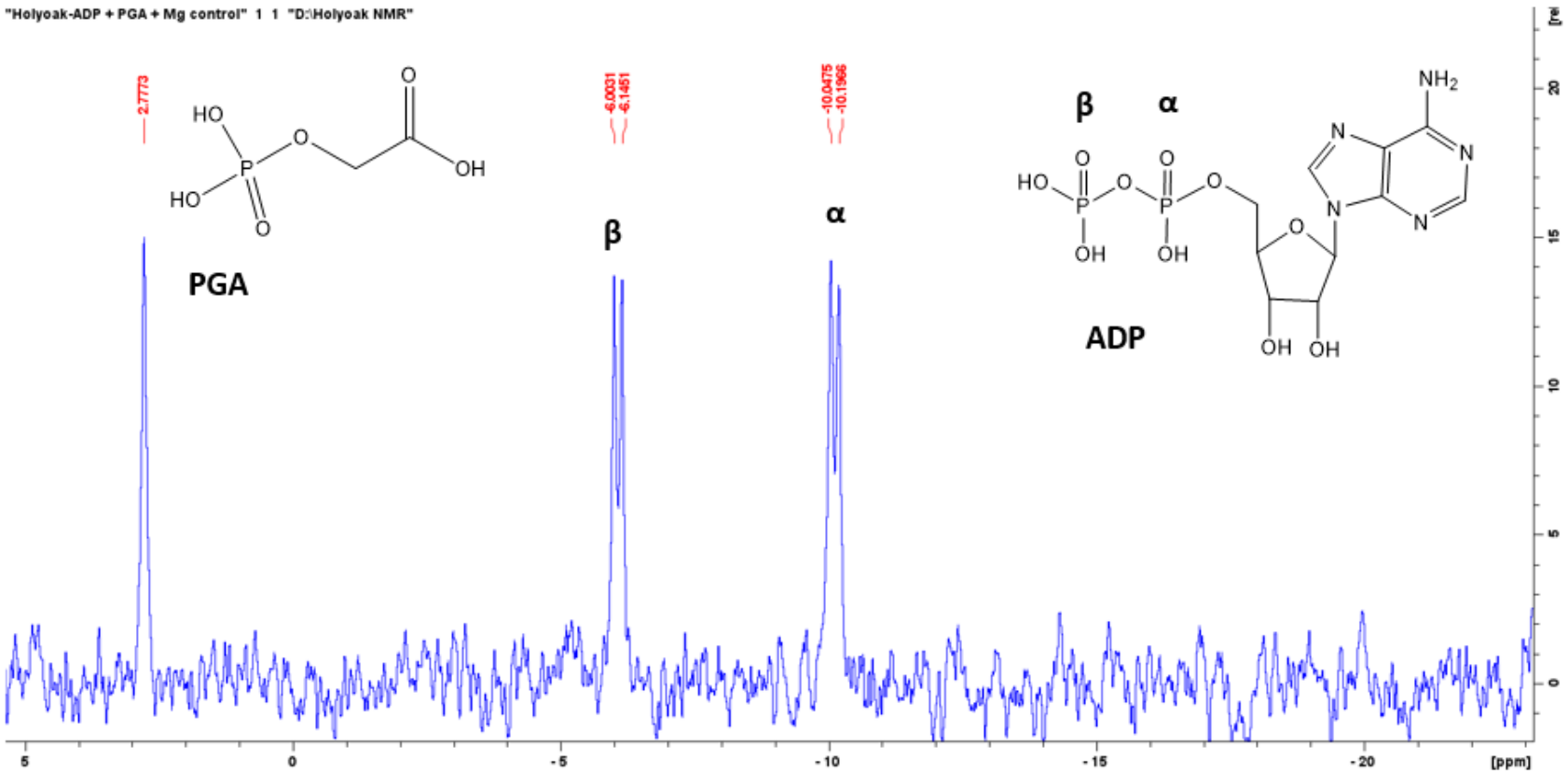


Figure G4: Adenosine 5'-diphosphate (ADP) and phosphoglycolate (PGA) ^{31}P NMR spectrum 0 hours - Control

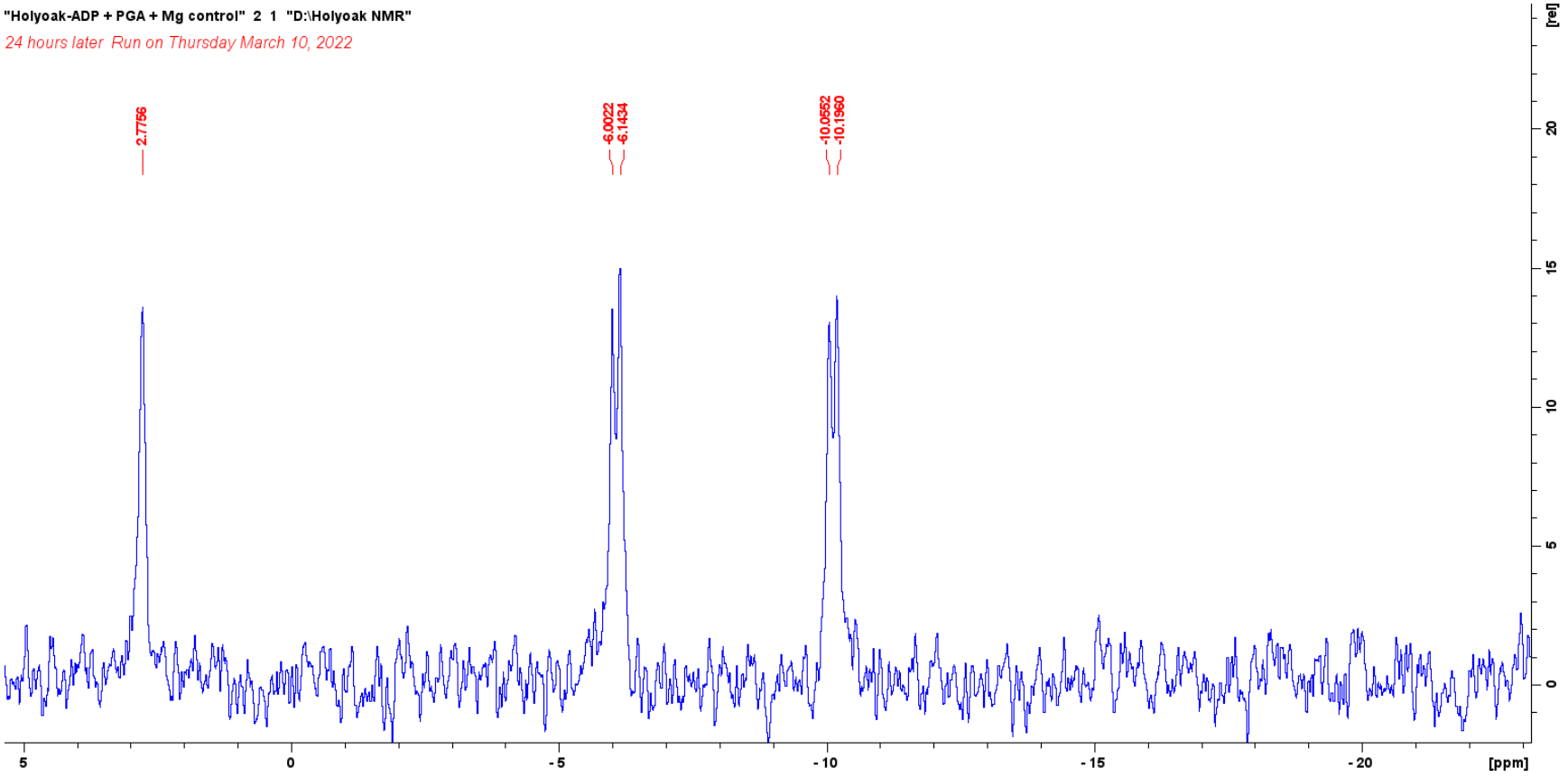


Figure G5: Adenosine 5'-diphosphate (ADP) and phosphoglycolate (PGA) ^{31}P NMR spectrum 24 hours - Control

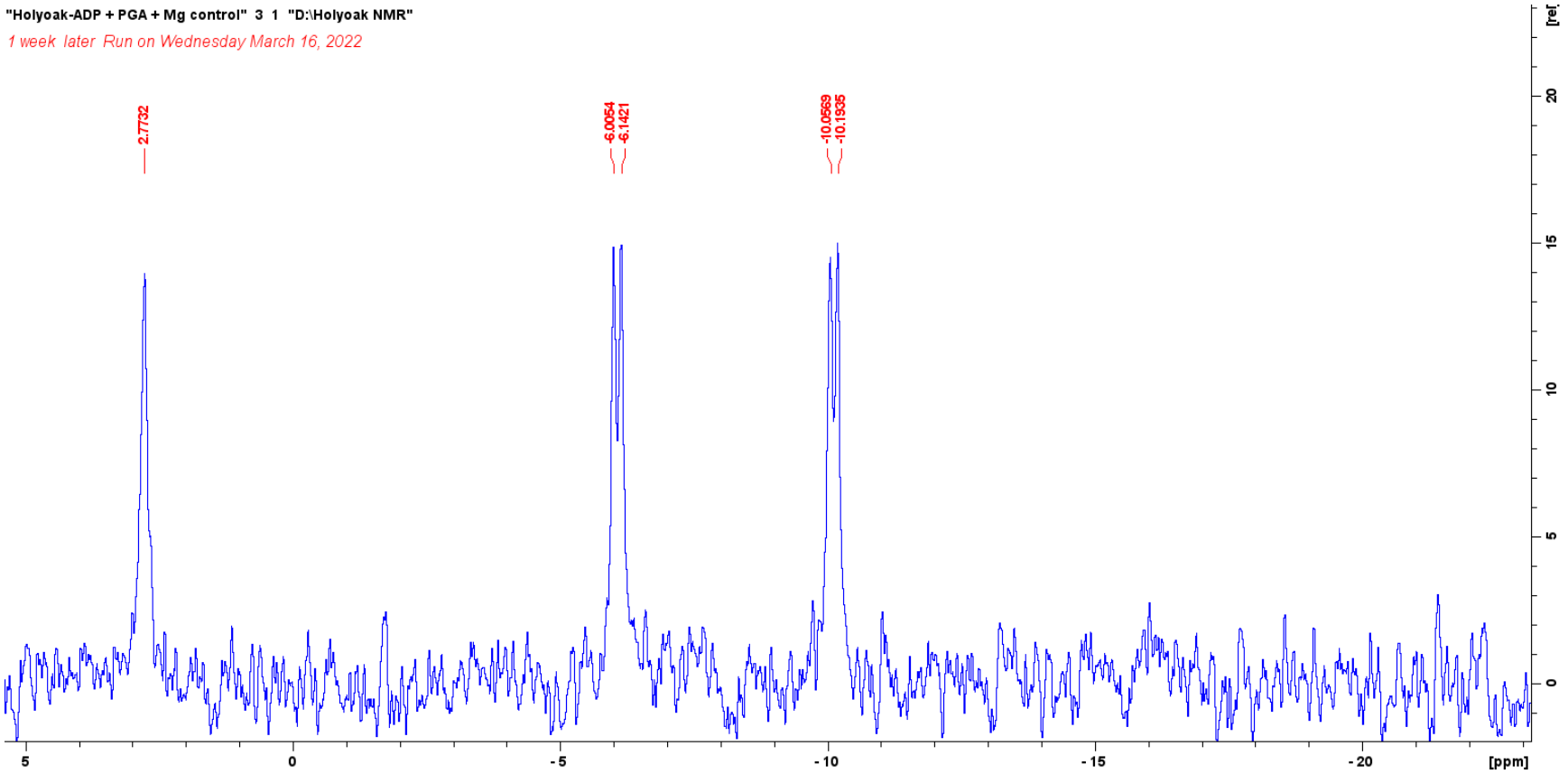


Figure G6: Adenosine 5'-diphosphate (ADP) and phosphoglycolate (PGA) ^{31}P NMR spectrum 1 week - Control

^{31}P NMR EcPEPCK Assay Spectra

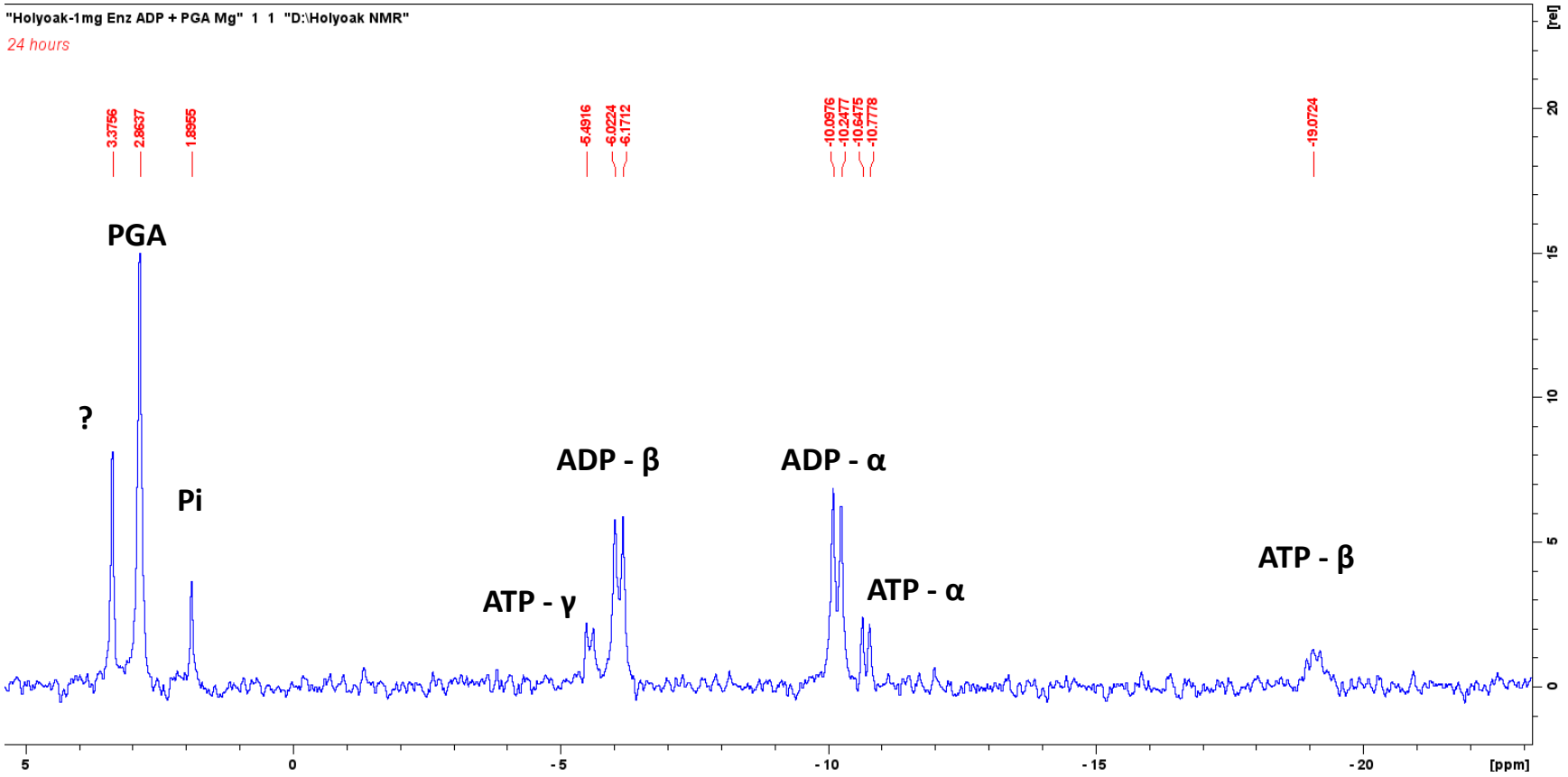


Figure G7: 1mg EcPEPCK with adenosine diphosphate (ADP) and phosphoglycolate (PGA) 24 hours

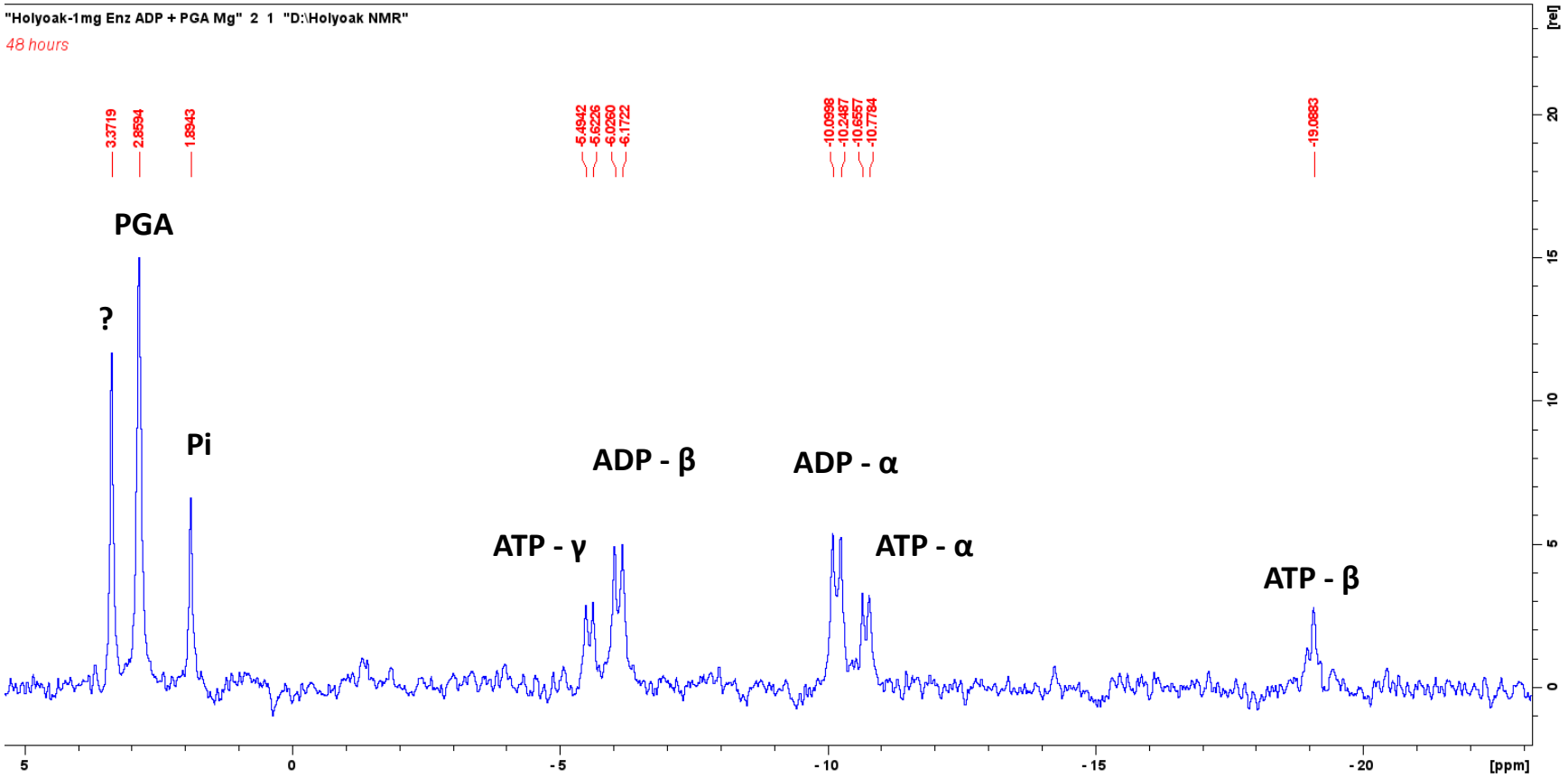


Figure G8: 1mg EcPEPCK with adenosine diphosphate (ADP) and phosphoglycolate (PGA) 48 hours

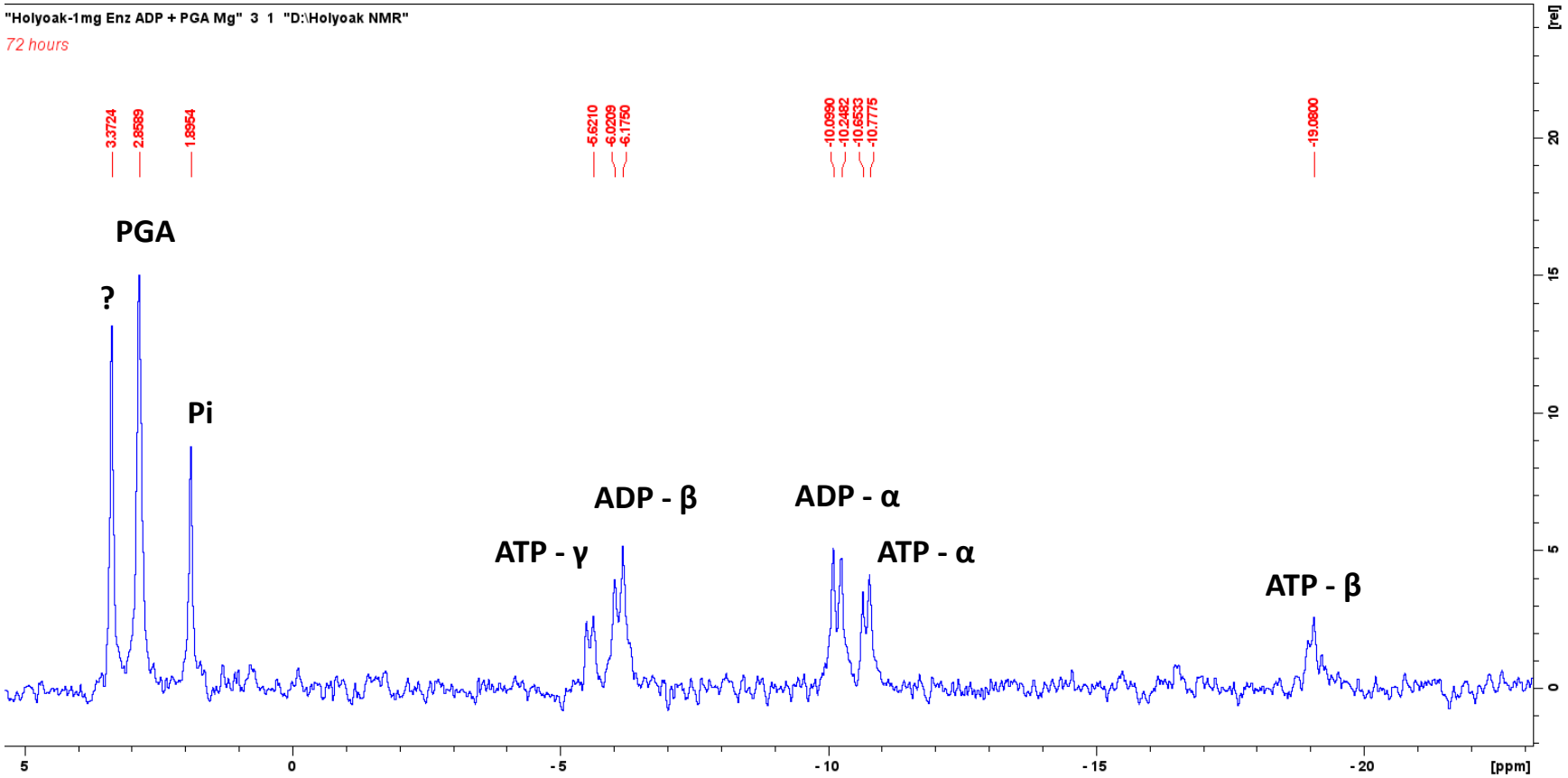


Figure G9: 1mg EcPEPCK with adenosine diphosphate (ADP) and phosphoglycolate (PGA) 72 hours

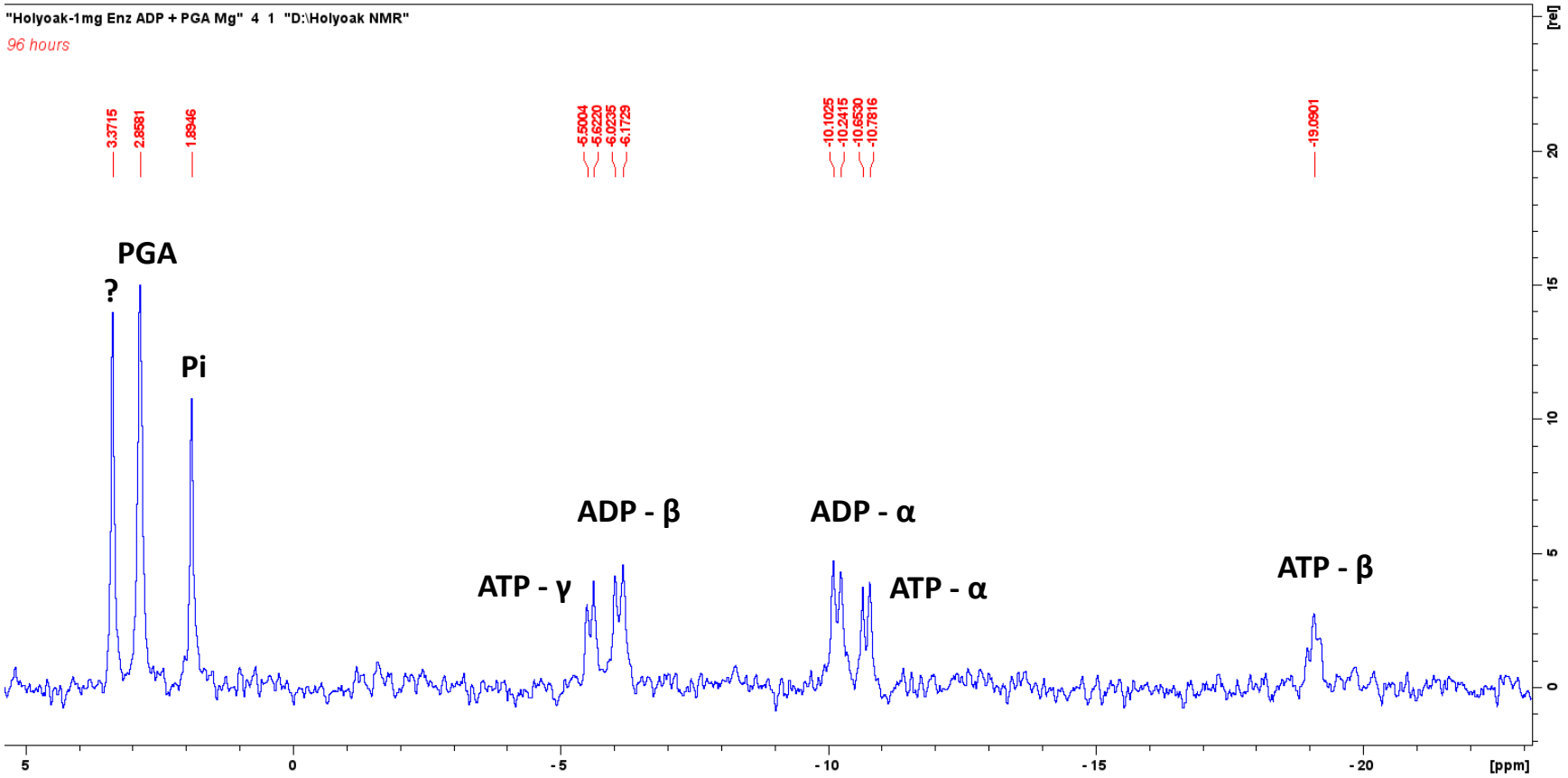


Figure G10: 1mg EcPEPCK with adenosine diphosphate (ADP) and phosphoglycolate (PGA) 96 hours

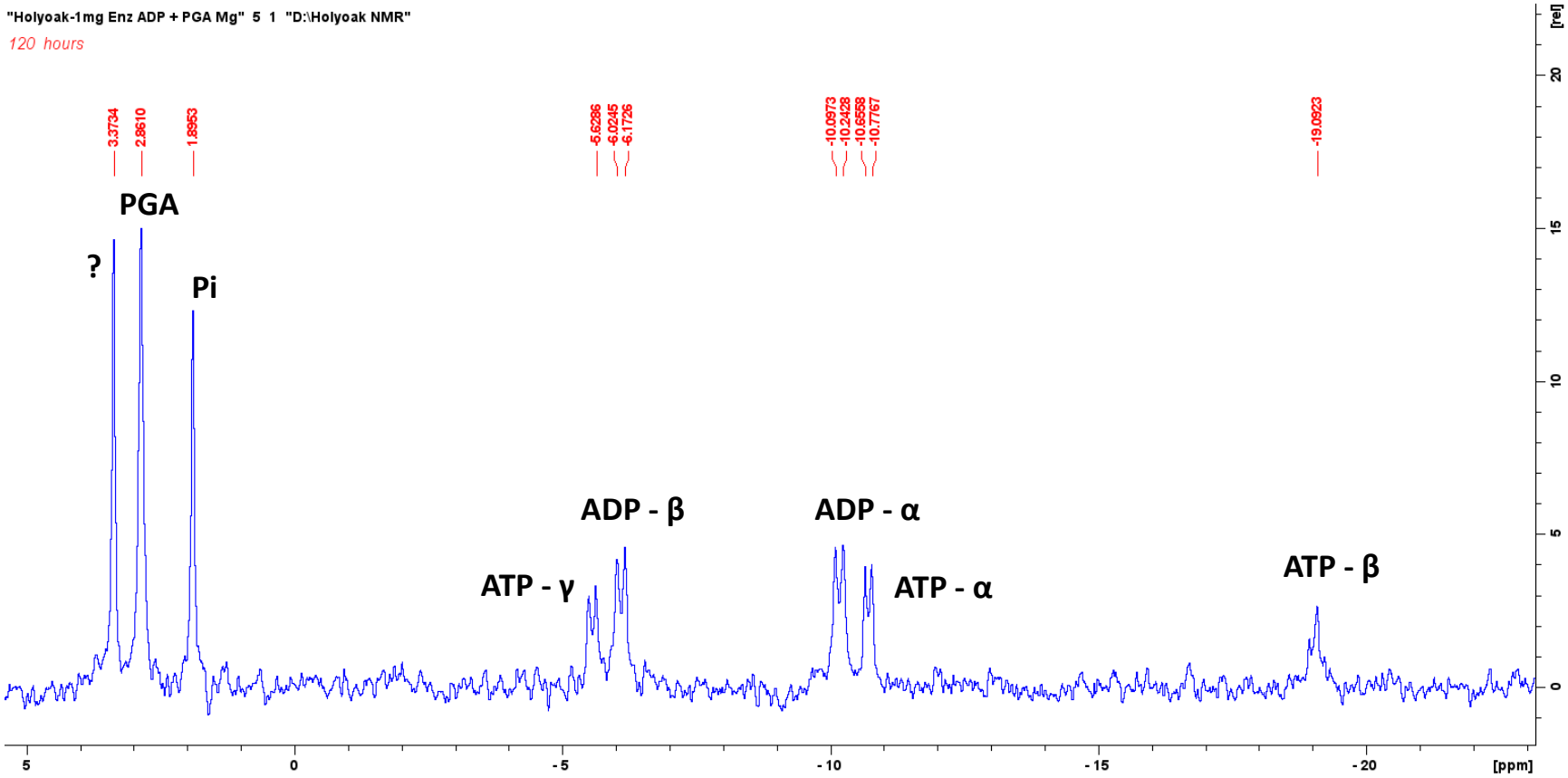


Figure G11: 1mg EcPEPCK with adenosine diphosphate (ADP) and phosphoglycolate (PGA) 120 hours

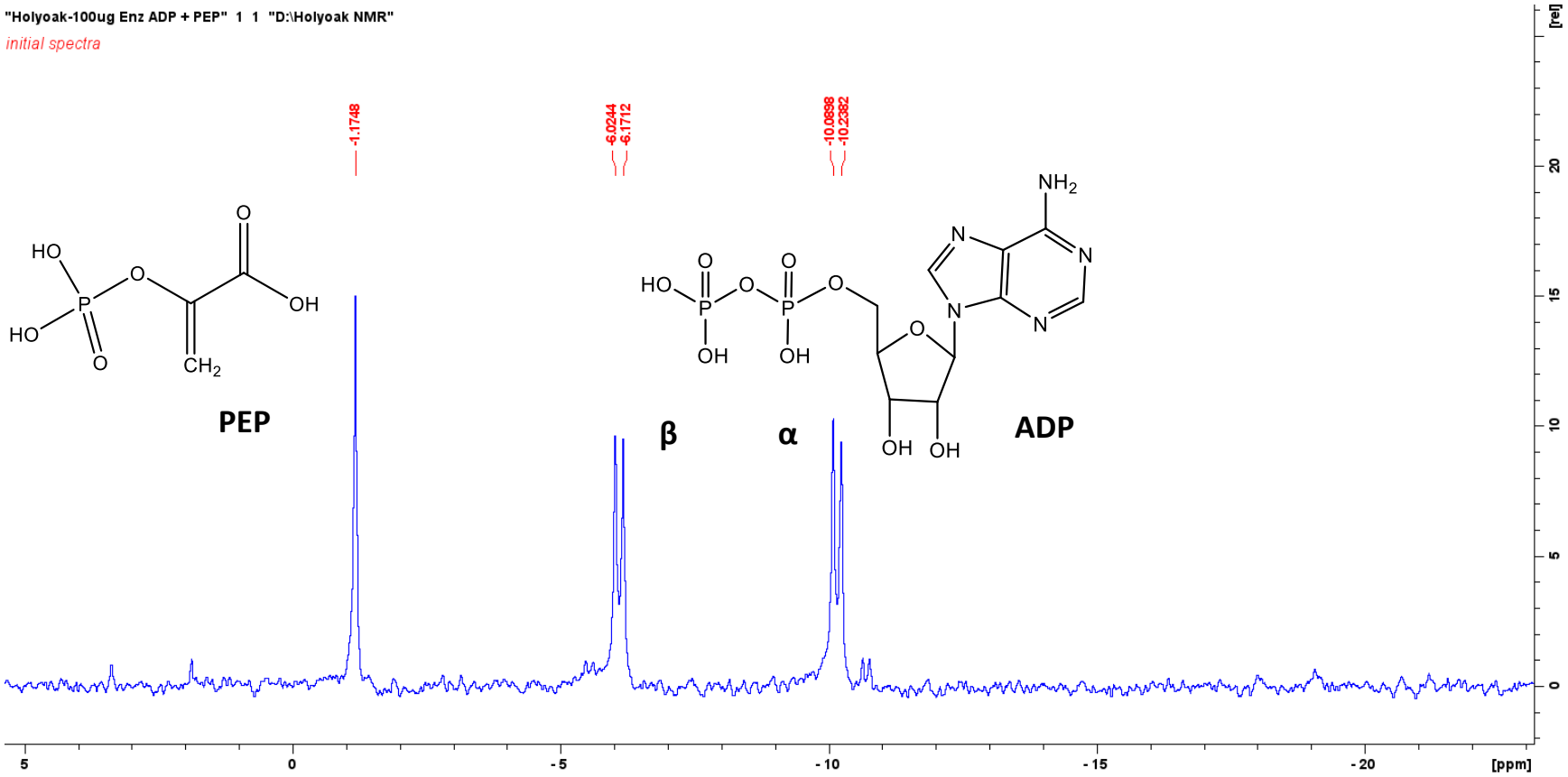


Figure G12: 100ug EcPEPCK with adenosine diphosphate (ADP) and phosphoenolpyruvate (PEP) 0 hours

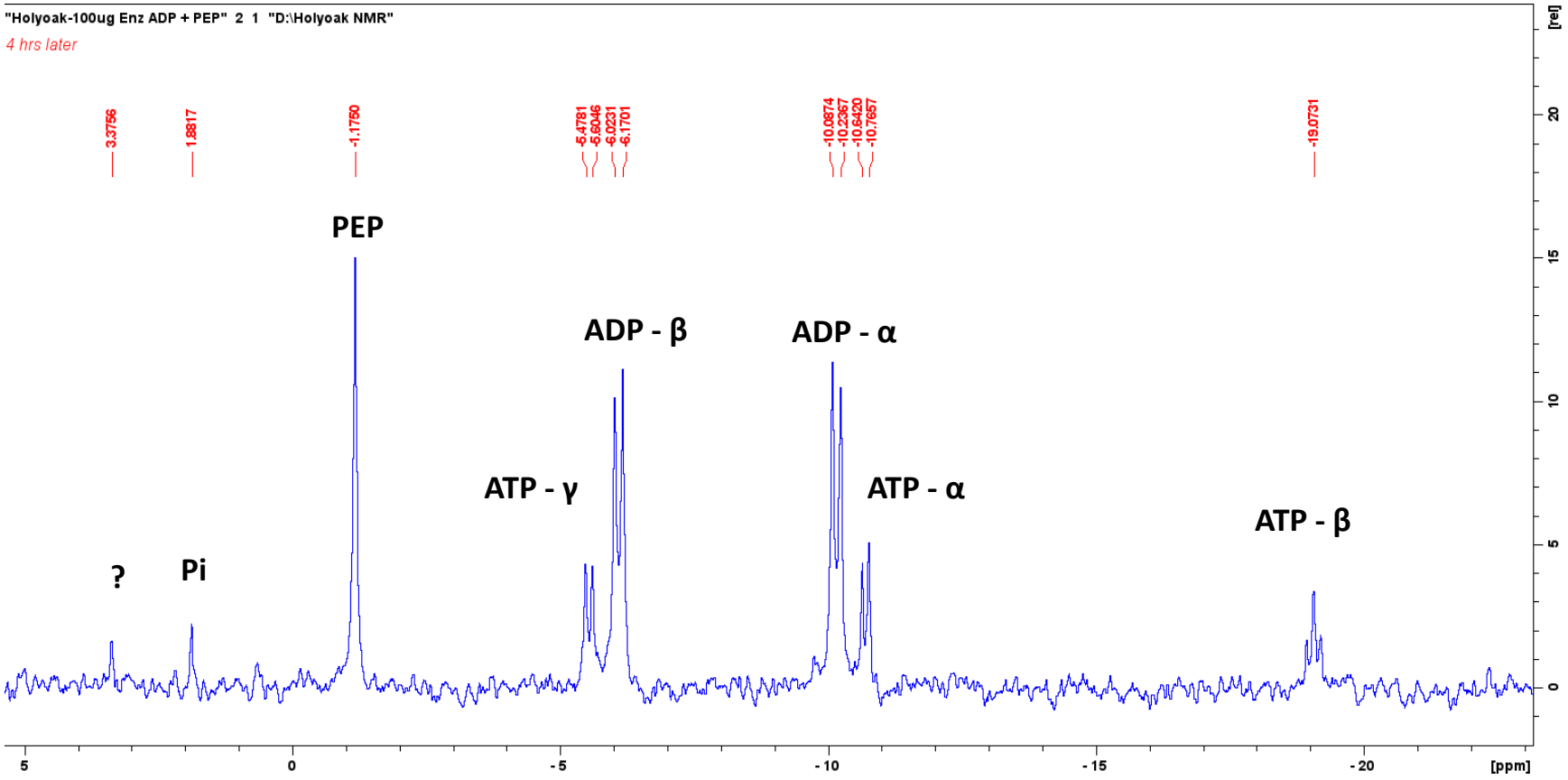


Figure G13: 100ug EcPEPCK with adenosine 5'-diphosphate (ADP) and phosphoenolpyruvate (PEP) 4 hours

Appendix H – HPLC chromatograms for ADP purity

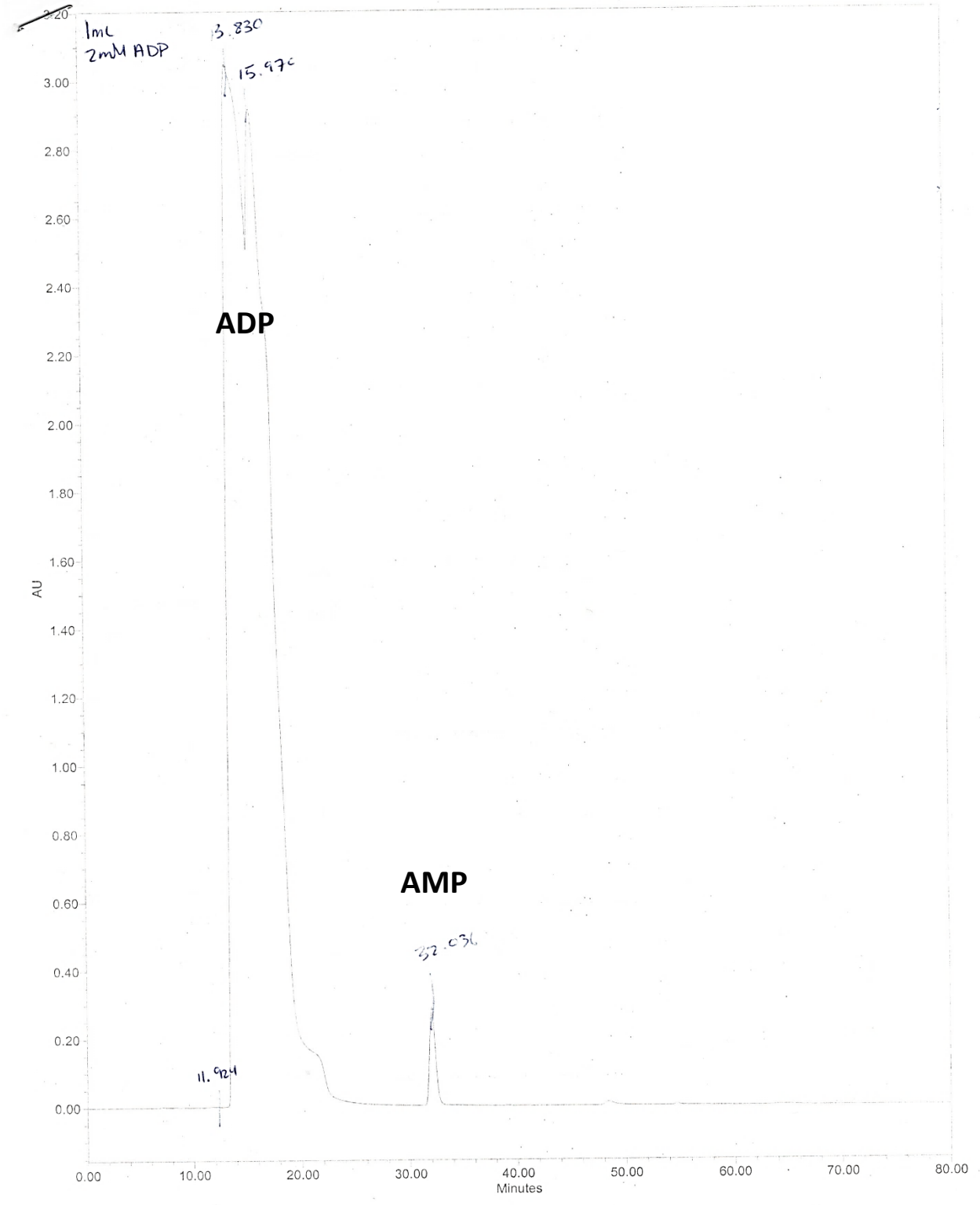


Figure H1: HPLC chromatogram for 2 mM ADP – 0 to 80 minutes, A260 detection

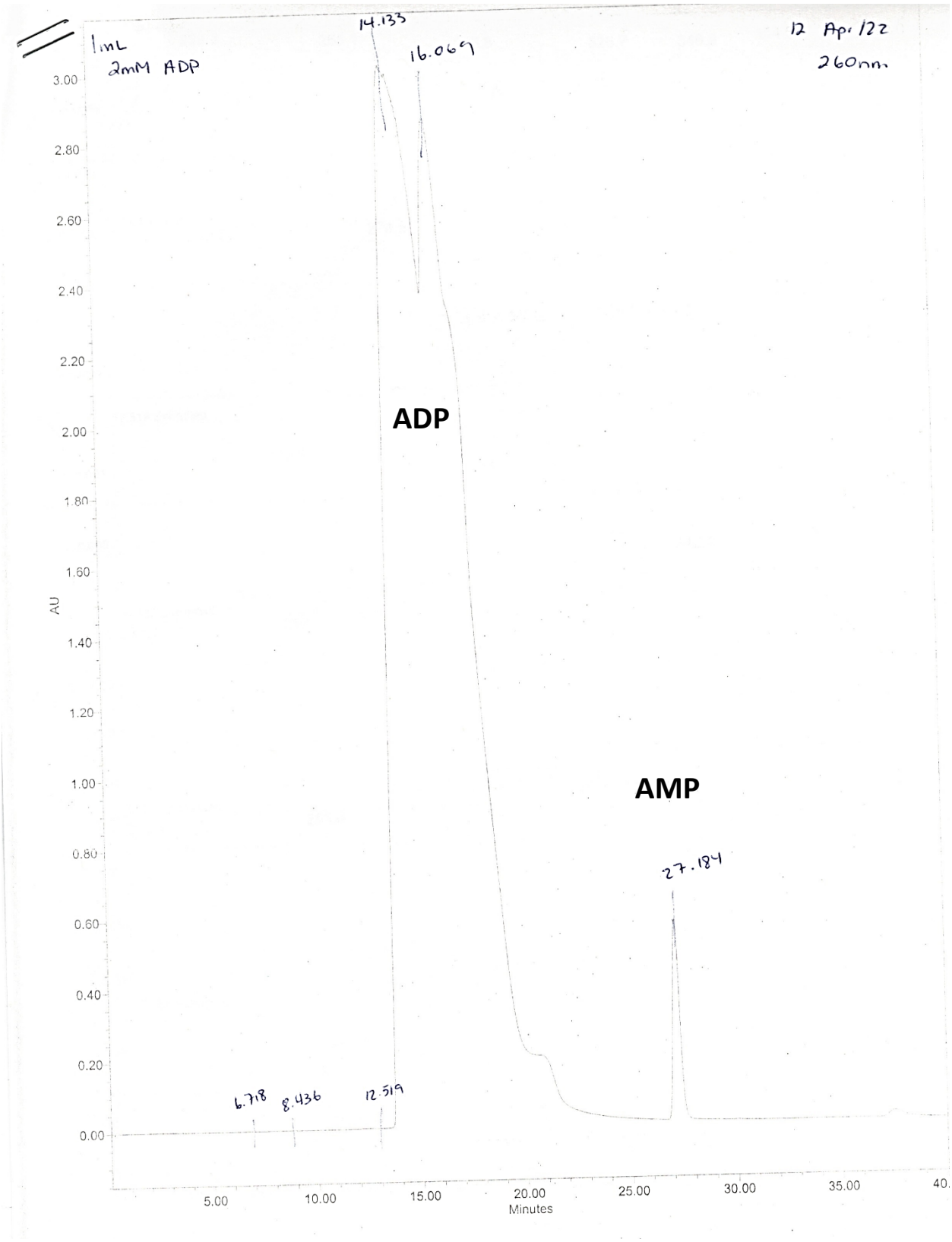


Figure H2: HPLC chromatogram for 2 mM ADP – 0 to 40 minutes, A260 detection

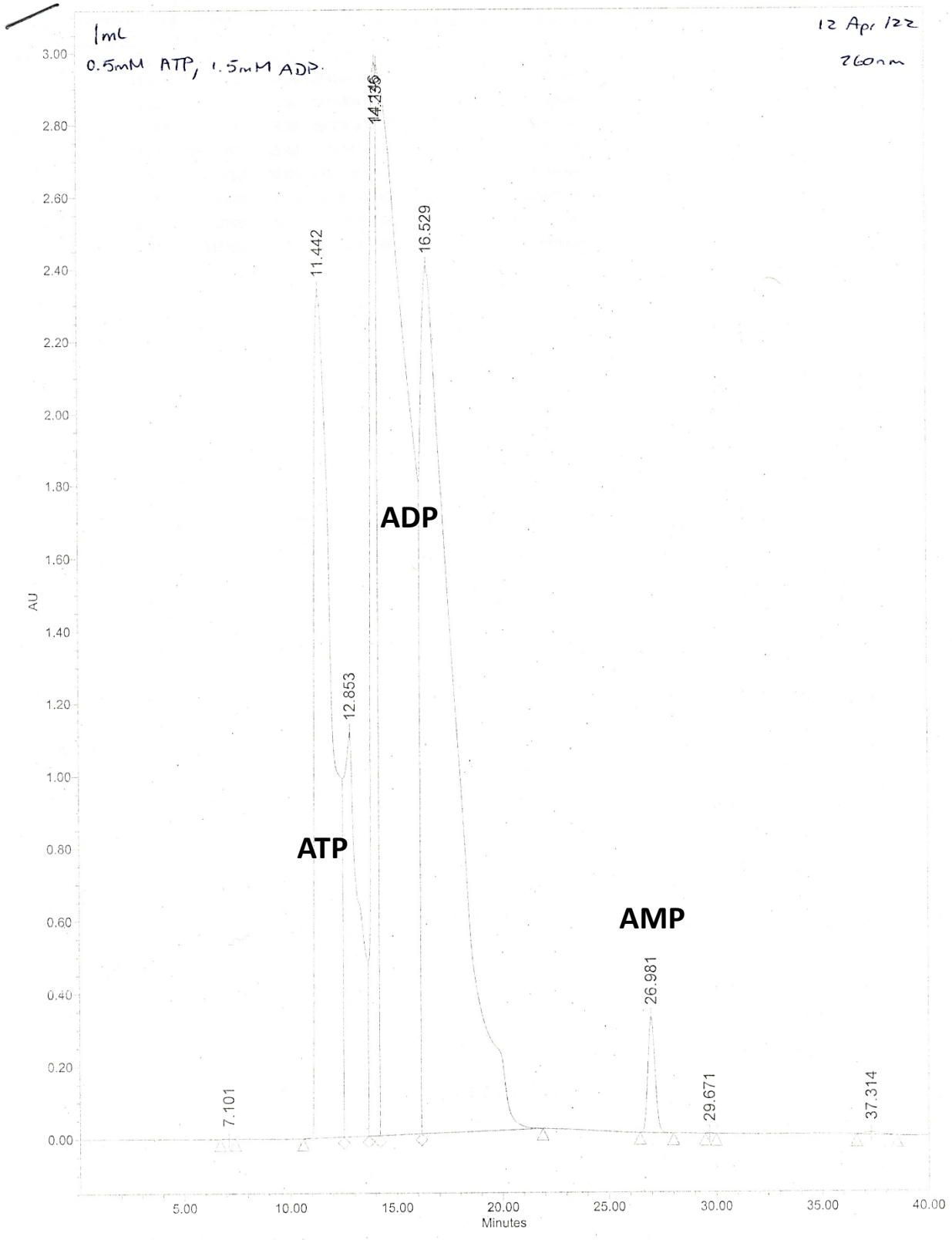


Figure H3: HPLC chromatogram 2 mM ADP + 500 µM ATP – 0 to 40 minutes, A260 detection