Functional consequences of changing structure, dynamics, and free-energy landscapes of phosphoenolpyruvate carboxykinases

by

Matthew McLeod

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Biology

Waterloo, Ontario, Canada 2020

Examining Committee Membership

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

Supervisor	Dr. Todd Holyoak Associate Professor – University of Waterloo (Biology)
Internal Member	Dr. David Rose Professor – University of Waterloo (Biology)
Internal Member (2)	Dr. Andrew Doxey Associate Professor – University of Waterloo (Biology)
Internal-External Member	Dr. Subha Kalyaanamoorthy Assistant Professor – University of Waterloo (Chemistry)
External Examiner	Dr. Gerald Audette Professor – York University

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

The complex relationship between structure, dynamics, and function will afford insights into the operational details of enzymes. To understand, regulate, and create enzymes, this complex nature must be comprehensively understood. In order to do so, perturbing a well-characterized enzyme in order to discern how changes in primary sequence, alter the structure and dynamics will begin to uncover these details. These modifications and their functional consequences will explain not only how the model system behaves, but also can be extrapolated to better understand the foundational principles of enzymes in general. These studies carried out in this thesis utilize a well-characterized enzyme, phosphoenolypruvate carboxykinase (PEPCK), as a structural and functional scaffold. Three studies investigate increasing degrees of change to this scaffold. The first perturbations of this scaffold start with a single point mutant that imparts a hysteretic kinetic behaviour in order to understand how an allosterically localized specific residue can control global dynamics. The second study examines three isozymes of PEPCK from different organisms that have adapted to different thermal niches. This study attempts to understand how full sequence adaptations tune enzyme structure and dynamics to different thermal availability, balancing dynamics with turnover. The final study is a comprehensive evaluation of the third class of PEPCK, a pyrophosphate-dependent enzyme, which is structurally and functionally distinct from the classical ATP and GTP-dependent PEPCKs. This study acts as a starting point for further investigations of this variant of PEPCK while also highlighting how adding different domains/lobes to a common central scaffold can drastically alter function. An Appendix is also included which demonstrates how a series of related inhibitors can have drastic differences in potency for a potential therapeutic target, the enzyme cytosine triphosphate synthase.

Acknowledgements

I would not be where I am today without the help of many people. First, I would like to thank my committee for reviewing my thesis and offering your expertise. I would like to specifically thank Dr. David Rose for being my committee member and mentor in all three of my degrees. You always had me in mind for various professional development opportunities and have been a valuable contribution in my career. I would like to thank both the Student Success Office and Centre for Teaching Excellence and all of those that I have worked with and for. Specifically, Erica McKellar and Dr. Svitlana Taraban-Gordan, thank you for giving me the opportunity to work on other aspects of my career while at UWaterloo. Thanks to the many graduate students that have befriended and help me along with my education - especially Dr. Marcie Chaudet as one of the first mentors in the lab, and pushing me to do more with my teaching career. I appreciate the patience for the countless crystallography questions I had for you when beginning and the teaching opportunities you presented to me which has opened many doors. I would like to thank my many friends and colleagues gained in graduate school in no particular order: Iain Wallace, Mohini Nema, Cody Shirrif, Dr. Justin Knapp, Quinn Abram, Dr. Mike Mansfield, Dr. Aaron Pual Frenette, and Michelle Lavery. Thank all of you for companionship, academic banter, and probably most importantly the nights off from work. Special shout-out to “Stormin” Norman Tran for the constant back and forth working out the fundamentals of biochemistry, I hope one day we can write “Biochemistry for Dummies”. Thanks to my many friends from the Kilworth and Sunview who have been important in expanding my horizons and point of view, while offering me a reprieve from the stresses of life. Without my family Karen, Art, Mike, Jess and Em aka “Boots”, I would be nowhere near the place I am today. The lessons you have taught me, and the love and support you have shown, has given me the tools that had allowed me to be as successful as I have been. There is no substitution for a great family.

Finally, no words can describe the gratitude I have for my undergraduate, Masters, and PhD supervisor Dr. Todd Holyoak. You allowed me to work in your lab even though I came asking for a position months after the deadline for the Senior Honours project. From there, the almost 7 years of mentorship have been invaluable. I am appreciative of all of the scientific lessons while never placing limitations on me. Without your guidance I wouldn’t be the scientist I am today.

Thank you all

Table of Contents

List of Figures

List of Tables

List of Abbreviations

*Ai*PEPCK (Actinomyces israelii)

E_a (Activation Energy)

$\Delta H^{\ddagger}$ (Activation enthalpy)

$\Delta S^{\ddagger}$ (Activation entropy)

A**D/T**P (Adenosine di/triphosphate)

ASU (Asymmetric unit)

βSP (Beta-sulfopyruvate)

cAMP (Cyclic adenosine monophosphate)

CTPS (Cytosine triphosphate synthase)

DTT (Dithiothreital)

G**D/T**P (Guanosine di/triphosphate)

HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)

clostPEPCK *(Hungateiclostridium thermocellum* PEPCK)

I**D/T**P (Inosine di/triphosphate)

Koshland-Nemethy-Filmer (KNF)

LDH (Lactate dehydrogenase)

MDH (Malate dehydrogenase)

MPA (3-Mercaptopincolinic acid)

MWC (Monod-Wyman-Changeux)

mtb (Mycobacterium tuberculosis)

NADH (Nicotinamide adenine dinucleotide)

NTA (Nitrilotriacetic acid)

N**M/D/T**P (Nucleotide mono/di/triphosphate)

OAA (Oxaloacetic acid)

PEP (Phosphoenolpyruvate)

PEPCK (Phosphoenolpyruvate carboxykinase)(PCK- gene name)

polarPEPCK *(Polaromonas naphthalenivorans* PEPCK)

PEG (Polyethylene glycol)

PTM (Post-translational modification)

*Pf*PEPCK *(Propionibacterium freudenreichii* PEPCK)

PDB (Protein Data Bank)

PP_i (Pyrophosphate)

PK (Pyruvate kinase)

rcPEPCK (Rat cytosolic PEPCK)

RSGI (RIKEN Structural Genomics/Proteomics Initiative)

SDS (Sodium dodecyl sulfate)

TCA cycle (Tricarboxylic acid cycle)

TIM (Triosephosphate isomerase)

TCEP (Tris(2-carboxyethyl)phosphine)

TRIS (Tris(hydroxymethyl)aminomethane)

WT (Wild-type)

1.0 Introduction

1.1 Preface

Each protein biochemist hope should be to study their model system with an increasing amount of detail, in order to later extrapolate these findings to a general unified model of protein function. Although this model is continuously improving it has some well-established principles. Some of these principles include (but are not exclusive to) seminal work from Koshland and Fisher first describing induced fit and lock-key models (later becoming a spectrum of induced-fit/conformational selection for substrate recognition).[1,2] Other descriptions by Frauenfelder, Karplus, and others showed proteins not as static structures but rather as a dynamic ensemble.[3–5] A great many more features are coming to the forefront to become widely accepted throughout the field. Each of these advancements help us understand proteins as a whole. In order to understand these mechanisms to form a unifying theory, a few extremely well-characterized proteins have been used as models. Although this likely works in describing the general features of proteins, more models are necessary to broaden the understanding of the unified model and its nuances.

It is well understood that proteins can be described as rugged, complicated, changing free-energy landscapes that take into account each atom of the system, that is protein, water/salts, substrates, among any other biomolecules present. This complexity means that it is impossible to describe accurately, but rather can be used as a relative representation of an enzyme. Throughout each step of an enzymes role, this energetic landscape changes according to the environment it finds itself in. For instance, as a free enzyme, the landscape will take into account all conformational modes that accompany this state. Upon ligand binding to the active site, the free-energy landscape must adapt resulting in a change in the distribution of sampled conformations. Thus, there direct relationship between the energetic landscape and dynamics. This free-energy landscape can be thought of as a series of barriers (local maxima) or valleys (local minima) of free-energy, the former describing transient high-energy states, while the latter describes longer-lasting low-energy ones. When an enzyme functions, as it would do when converting a substrate to a product, this system must traverse the free-energy landscape moving over the barriers to find new stable lows. Each barrier represents the probability of being able to transition from one state to the next. Each of these jumps over a barrier has associated with it an associated thermodynamic cost, which can change depending on how the enzyme has adapted or been manipulated within its environment. Evolution will tweak the primary sequence of a given protein depending on what environment it finds itself in, and how to best adapt the protein for its cellular function. As researchers attempt to understand and "map" this landscape, we may use single point mutations, or more numerous changes throughout the sequence, to perturb a proteins' energetic landscape in order to determine the consequences of these changes. Traditionally, mutational studies, swapping out residues that are deemed important, are usually reserved to studying only a few changes in the proteins primary sequence targeting specific interactions and how these change the functionality of the enzyme. This can lend insight into what specific structural features are responsible for particular functional steps. However, single mutation studies do not give us insight as to how the entire primary sequence leads to function. Studying isozymes, that is structurally and functionally related proteins like those from varying organisms differing by sequence, can give us insight into how the entire primary sequence dictates the properties of a protein. Finally, beyond isozyme sequence variability, evolution has seemingly used set of common structural scaffolds (domains/motifs), adding together specific features to invent enzymes with new functions. The differences from the sequences, manifested as changes in structure and function, will further lend information to add to a unified model of protein function, and hopefully lead to the pinnacle of protein biochemistry, *de novo* protein design.

Here is described a series of studies using a well-studied model enzyme, phosphoenolpyruvate carboxykinase (PEPCK), as a common functional and structural scaffold to better understand how these sequence changes, at varying degrees of modification discussed, leads to functional differences. The three areas that are investigated here are a point mutation in a dynamic region of PEPCK leading to hysteretic behaviour, isozyme differences dictated from thermal adaptation of three PEPCKs, and finally a relatively uncharacterized class of PEPCK which uses the same core protein fold, but additional domains and lobes have appended leading to both structural and functional changes. Finally, the appendix contains other work completed in my doctoral studies on the inhibition of an enzyme, cytosine triphosphate synthase (CTPS) by a series of fluorinated analogues.

1.2 Phosphoenolypyruvate carboxykinase history and overview

PEPCK has a storied history beginning in the 1954 where it was initially characterized in a series of papers initially naming the enzyme oxalacetic carboxylase.[6–8] This enzymes', purified from chicken liver, chemical mechanism was generally described as oxaloacetic acid's (OAA) decarboxylation in the presence of ITP to form phosphoenolpyruvate (PEP). From this starting point and after a name change to PEPCK, the *in vivo* cellular role of the enzyme has been elucidated showing its vital role in metabolism.[9–11] A therapeutic avenue for regulating PEPCKs activity has been highlighted through its role in diseases such as cancer[12], *Mycobacterium tuberculosis (mtb)* infections[13], and diabetes[14], but it also has ascribed roles in aerobic capacity[15] and longevity.[16] This apparent importance has led to many functional studies of PEPCK from many different organisms[17–27], and many high resolution crystal structures of different isozymes[28–36], some which are published while others are only deposited in the Protein Data Bank (PDB). Coupled with this structural information, kinetic data understanding the flux of substrates through this enzyme, as well as mechanistic details of its catalysis have been uncovered.[29,37–41] Although it has been studied extensively, the story of this enzyme continues to evolve as even recently it has been implemented in lipogenic regulation through a new purported protein-kinase activity.[42]

1.3 Metabolic role

PEPCK works adjacent to the tricarboxylic acid cycle (TCA cycle) by removing anionic molecules to maintain cycle flux. In almost all cellular contexts, PEPCK removes OAA to form PEP. This PEP can have many downstream fates such as: gluconeogenesis (PEPCKs most traditionally identified role), glyceroneogenesis, amino acid synthesis, or phosphoenolpyruvate (PEP) can be converted to pyruvate through PEP carboxylase and later feed back into the TCA cycle.(reviewed in 43) The vast array of metabolic consequences of PEPCK activity have given it the title of master regulator of the TCA cycle, and as having a crucial role in metabolism as a whole.[44]

1.4 Regulation

Due to PEPCKs’ role in whole organismal homeostasis, it is likely tightly regulated. Surprisingly, there have not been that many regulatory mechanisms of PEPCK determined. At a genomic level, the expression of the genes for PEPCK (PCK) are regulated by small molecules such as cAMP[45] and insulin[46,47] or through environmental changes such as acidosis[48–50], all resulting in lower expression. DNA[51] and histone methylation[47] of the PCK gene have also been shown to regulate PEPCK, where tissue specific increases in methylation lead to decreased expression of PEPCK. There are only a few indicated post-translational modification (PTM) that regulate PEPCK, although many residues of the enzyme have been implicated as potential PTM sites. The first PTM identified was lysine acetylation[52,53], although the details are still contentious.[54] Lysine acetylation of PEPCK has been reported to increase ubiquitin labelling by UBR5, leading to recycling of PEPCK through the proteosome. A second PTM has been recently identified, although the results are still waiting for peer review. It was shown that the mitochondrial PEPCKs activity can be allosterically regulated by sulfhydration of a non-active site cysteine.[55] Finally, as previously mentioned, phosphorylation of cytosolic PEPCK can lead to a change in activity, where it may become a GTP-dependent protein kinase involved in the regulation of lipogenesis.[42]

1.5 Phosphoryl donor specificity

PEPCKs can be classified into three categories dependent on their specificity for the phopshoryl donor; ATP- (EC: 4.1.1.49/K01610), GTP- (EC: 4.1.1.32/K01596) and most recently added PP_i-[(EC: 4.1.1.38/K20370) dependent PEPCKs. Of these three families, they are all phylogenetically distinct and are widely spread throughout the tree of life. There are a few trends amongst how these enzymes are distributed throughout life, with some interesting but rare exceptions. First, an organism will usually only have a one PEPCK. C-4 Plant PEPCKs are ATP-dependent, while all mammalian PEPCKs are GTP-dependent. Of the *chordata* phylum, there are two compartmental GTP isozymes, one found in the cytosol and one found in the mitochondria. Both ATP and GTP PEPCKs can be found distributed amongst bacteria and fungi. Finally, Archael organisms have only utilized GTP-dependent PEPCKs apart from *Aeropyrum pernix* which uses ATP. Of the odd exceptions that are noted, *Candidatus kuenenia stuttgartiensis* does not have a defined nucleotide binding site and is only 23-36% identity to both ATP and GTP PEPCKs, suggesting the potential for promiscuity.[56] Finally, PP_i PEPCKs are widely distributed in both Eukaryotes and bacteria but are absent in the Archaea. There has not been a comprehensive study to look at the distribution of PP_i PEPCK within Eukaryotic and Bacterial phyla as has been completed as with the nucleotide dependent enzymes.[56]

Using Annotree[57], some high-level details can be discerned about the representation of the PEPCKs that the updated tree of life offer as much of this data was not accessible in previous studies (Table 1).[56] As seen indicated in previous studies, PP_i PEPCK are absent from Archaea, while all three forms are widely dispersed in Bacteria.[56,58] The majority of sequenced bacterial genomes are ATP PEPCKs (10148) followed by GTP (5571), and finally PP_i (417), while the majority of archael genomes have ATP (321) PEPCKs, followed by the GTP form (140). There do appear to be a few bacterial genomes annotated as containing two types of PEPCK, and 2 genomes with all three present. There is also 1 archael genome with both ATP and GTP PEPCK present. Unfortunately, Annotree does not have have the sequenced Eukaroytic genomes available for analysis. Because of the recent addition of the third PP_i-dependent PEPCK to the PEPCK family and with recent improvements bioinformatic tools, an update to the distribution of these three enzymes should be completed.

Table 1.1: Representative distribution of each isoform of PEPCK from Annotree

PEPCK Isoform	Number of Bacterial Genomes	Number of Archael Genomes
ATP	10148	321
GTP	5571	140
PP_i	417	0
ATP + GTP	104	1
ATP + PP_i	128	0
GTP + PP_i	165	0
ATP + GTP + PP_i	2	0

1.7 Global architecture

Structurally the ATP- and GTP-dependent PEPCKs are identical, with the “core” of PP_i PEPCK having a similar fold to the nucleotide-dependent isoforms despite having a general lack of sequence conservation between the the PP_i- and nucleotide-dependent PEPCKs (E-value ≥ 1). Of the most well studied nucleotide-dependent PEPCKs, the rat cytosolic GTP PEPCK and *E. coli* ATP PEPCK, there is a 30% sequence similarity. The phosphoryl binding sites are distinct between the three isoforms, as they bind different phosphoryl donors. GTP and ATP PEPCKs have a N- and C-terminal domains (residues 1-249 and 260-622 respectively – rat cytosolic PEPCK (rcPEPCK) numbering), with the active site situated between them (Figure 1.1). Furthermore, the C-terminal domain can be separated into the nucleotide binding domain (residues 260-625, 426-622) and the PEPCK-specific domain (residues 326-425).[39]

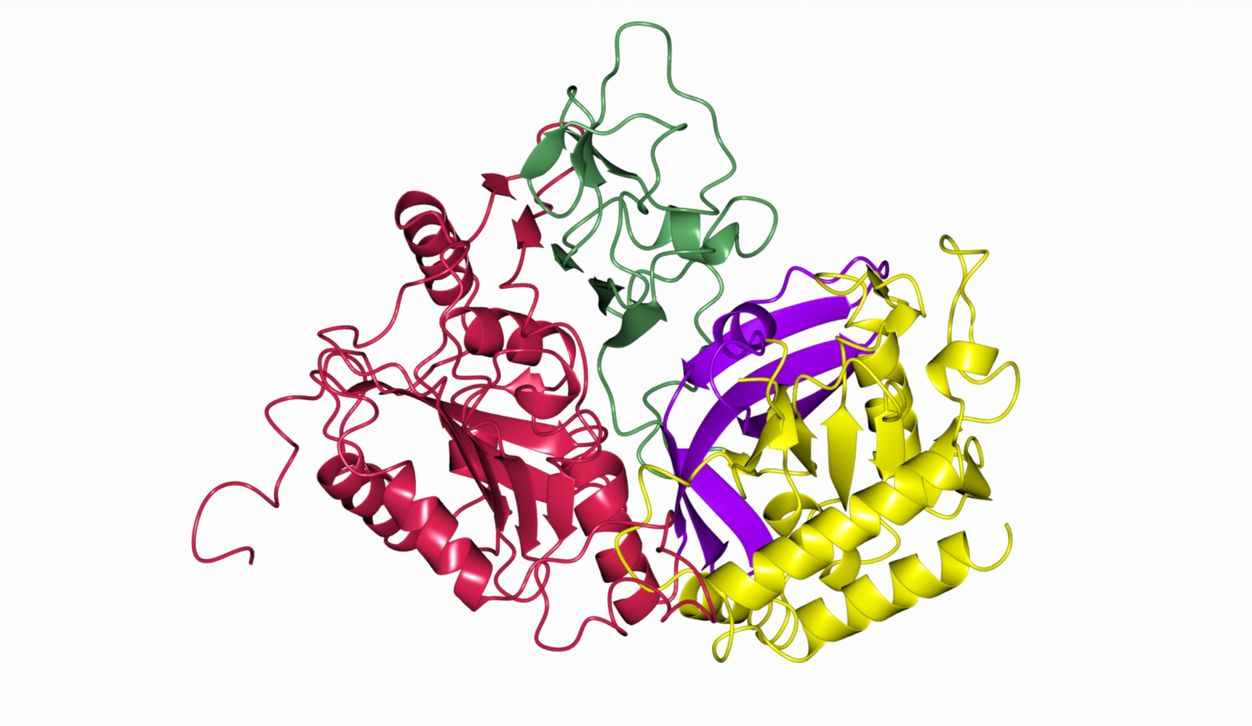


Figure 1.1: Nucleotide PEPCK domains. Rat cytosolic PEPCK (PDB ID# 2QEW) coloured by domains. The N-terminal domain is presented in red. The C-terminal domain (green) can be further broken down into sub-domains, the nucleotide binding domain (purple) and PEPCK-specific domain (yellow). ***All figures are rendered using CCP4MG with solvent removed for clarity.***

The PP_i PEPCK isoform consists of two domains, both a N- (residues 85-346, 648-721, and 758-792 *Actinomyces israelii* numbering (*Ai*PEPCK)) and C-terminal (residues 347-359, 475-496, 649-684, and 817-1047). In addition to these two domains, 4 lobes were described named lobe 1 (residues 1-84), lobe 2 (360-475, 497-649, and 793-817), lobe 3 (722-758), and lobe 4 (1048-1149) (Figure 1.2).[31] Although the additional lobes increase the mass of PP_i PEPCK by 50kDA (nucleotide PEPCK ~ 70kDA, PP_i PEPCK ~130kDa), the core of the PPi and nucleotide-dependent enzymes are structurally conserved (Cα RMSD – 2.9Å) (Figure 1.3).

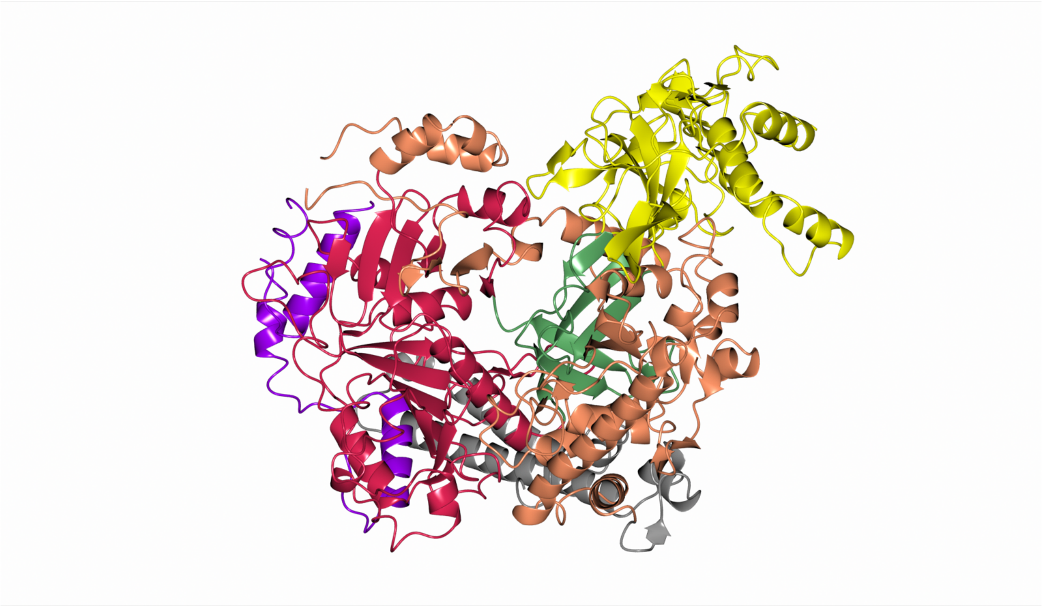


Figure 1.2: PP_i PEPCK domains and lobes. *Ai*PEPCK (PDB ID# 6K31) coloured by domains and lobes. The N- and C-terminal domains are coloured in red and green respectively. Lobe 1 is coloured in purple, lobe 2 in yellow, lobe 3 in coral, lobe 4 in grey.

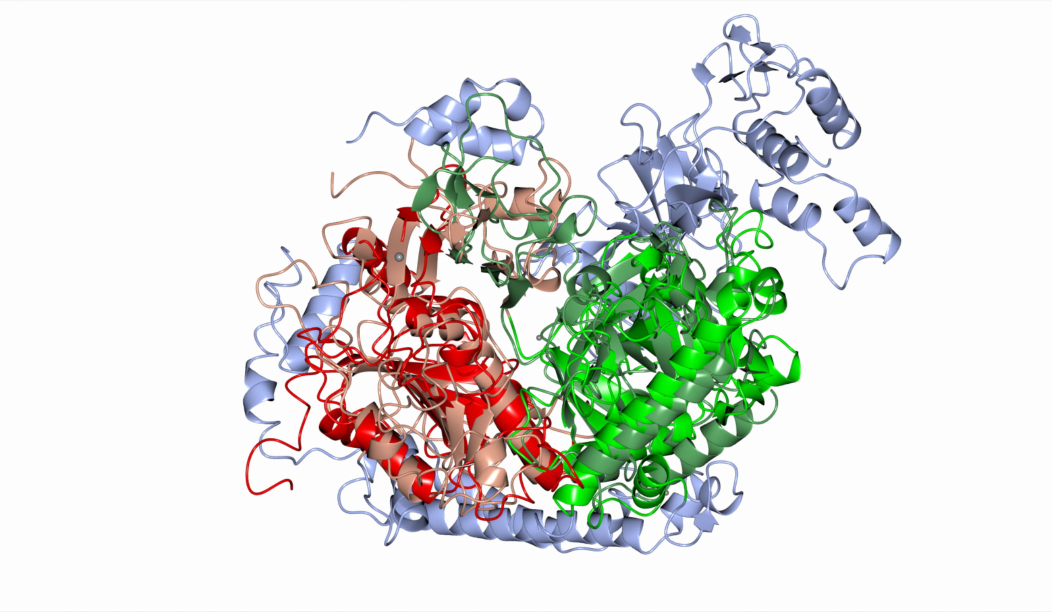


Figure 1.3: Alignment of nucleotide (GTP) and PP_i-dependent PEPCK. Alignment between rcPEPCK (PDB ID# 2QEW) and *Ai*PEPCK (PDB ID# 6K31) with domains coloured. RcPEPCK N-terminal domain is in tan, with the C-terminal domain coloured in light green. *Ai*PEPCK is shown in ice blue, with the N-terminal domain in red and the C-terminal domain in dark green.

1.8 Catalytic Loop Elements

Enzymes, in general, must undergo a series of motions in order to perform their catalytic function. Of these motions, loops are utilized to bind and orientate substrates in proper configurations to allow for the breaking and forming of chemical bonds. These active site loops, generally are mobile while unbound, and upon binding become rigid and ordered. Nucleotide PEPCKs have three active site loops that have been characterized (Figure 1.4). The generalities of these loops will be discussed below.

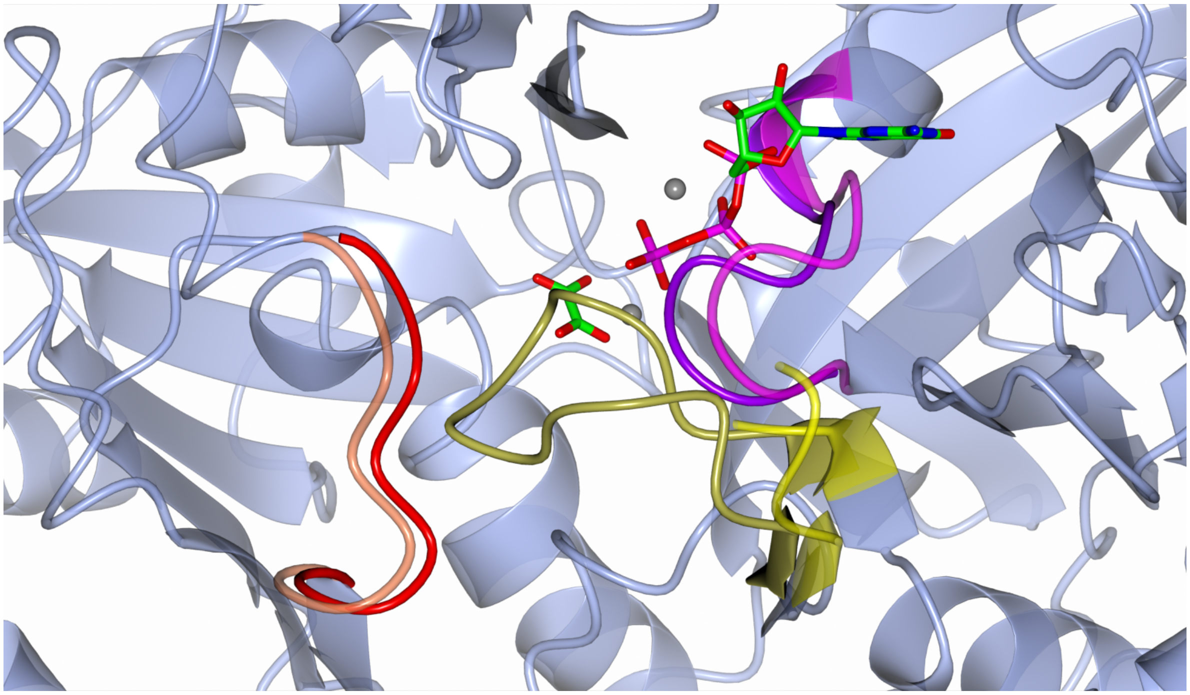


Figure 1.4: Nucleotide PEPCK active site loops. Alignment of holo-rcPEPCK (PDB ID# 2QEW, steel blue) and oxalate-GTP-rcPEPCK (PDB ID# 3DT2, hidden). The open-state R-, P-, and Ω-loops are shown in coral, magenta and yellow respectively. The closed-state R-, P-, and Ω-loops are shown in red, purple, and gold respectively. As seen, oxalate and GTP are present along with two Mn^{2+} ions. Atoms are coloured by type with carbon in green, oxygen in red, nitrogen in blue, and manganese in grey.

1.8.1 R-Loop

The R-loop (residues 85-92) is responsible for binding the substrates OAA and PEP. Upon binding, it reorders and allows for the closure of the active site lid, otherwise known as the Ω-loop, by removing steric hindrances as well as offering stabilizing bonds. A conserved R87 is responsible for substrate recognition and binding through the electrostatic interaction between the guanidine and the substrate, while E89 is responsible for stabilizing the Ω-loop in a closed conformation by forming a H-bond with H470.[39,41] Together, R87, E89, and E470 form a network which has been shown to be vital in stabilizing the closed conformation of PEPCK upon nucleotide and substrate binding.[40]

1.8.2 P-Loop

The P-loop (residues 284-292) is a kinase 1a motif which binds the nucleotide, stabilizing the negative charges of the polyphosphate tail and coordinates the nucleotide for phosphoryl transfer. This loop, like the R-loop, is important in allowing the closure of the lid. Upon entering the closed state, this loop removes steric hindrances between A287 and T465 of the Ω-loop.[39,41]

1.8.3 Ω-Loop

The Ω-loop (residues 464-474), acts as a lid to cover the active site while the chemical reaction is going on. This covering sequesters water from the reaction site, preventing unwanted chemistry from occurring. Ω-loops are a well-characterized non-regular secondary structure element in which it is described as a loop with close entry and exit points (as seen from the β-strands in Figure 1.4). Aside from PEPCKs Ω-loop, a similar structure found on triosephosphate isomerase (TIM) has also been well characterized. These two loops differ in that PEPCKs Ω-loop does not directly coordinate with the substrates[29,59], and closes through a folding mechanism as opposed to a rigid-body motion.[60] Further details of this concerted closure of the active site loops will be discussed elsewhere.

1.9 Catalytic Cycle

Extensive work has gone into understanding the reaction mechanism and conformational changes required for nucleotide PEPCKs to turnover. This overview of the catalytic cycle will be done in the context of rcPEPCK working in the cataplerotic direction converting OAA to PEP, even though this enzyme can operate bidirectionally. All other evidence suggests that these events are conserved throughout the nucleotide PEPCKs. To date there is no published data for the details of the catalytic cycle of the PP_i-dependent PEPCK.

1.9.1 Reaction Mechanism

As mentioned, nucleotide-dependent PEPCKs can freely inter-convert OAA and PEP *in vitro*, while the OAA decarboxylation reaction is predominate *in vivo*. For this reaction to occur, two metals are utilized by PEPCK. The M1 metal, is an active site co-factor that bridges the gap between the substrate and nucleotide binding pockets. In almost all cases, this metal has been shown to be Mn^{2+}. The two exceptions are Co^{2+} was found to be most activating for *Anaerobiospirillum succiniproducens*[26] (an ATP PEPCK) and all previously characterized PP_i PEPCKs.[58,61] The second M2 metal, coordinates with the nucleotides β- and γ-phosphates to stabilize the anionic charges. In truth, the M2 acts with the nucleotide as a M2-nucleotide substrate complex. The M2 metal has always shown to be Mg^{2+}

although other divalent metals can substitute for either the M1 and M2 metal, with a kinetic deficit that is typically enzyme-dependent. The reaction begins with the decarboxylation of OAA to form a reactive enolate species. This enolate is then phosphorylated by the nucleotide through an in-line transfer, creating PEP, the nucleotide diphosphate (NDP) and CO_2 from the first step. This reactive enolate species, for which oxalate is an analogue, can be protonated by bulk solvent to form pyruvate if the Ω-loop prematurely opens (Figure 1.5).[38]

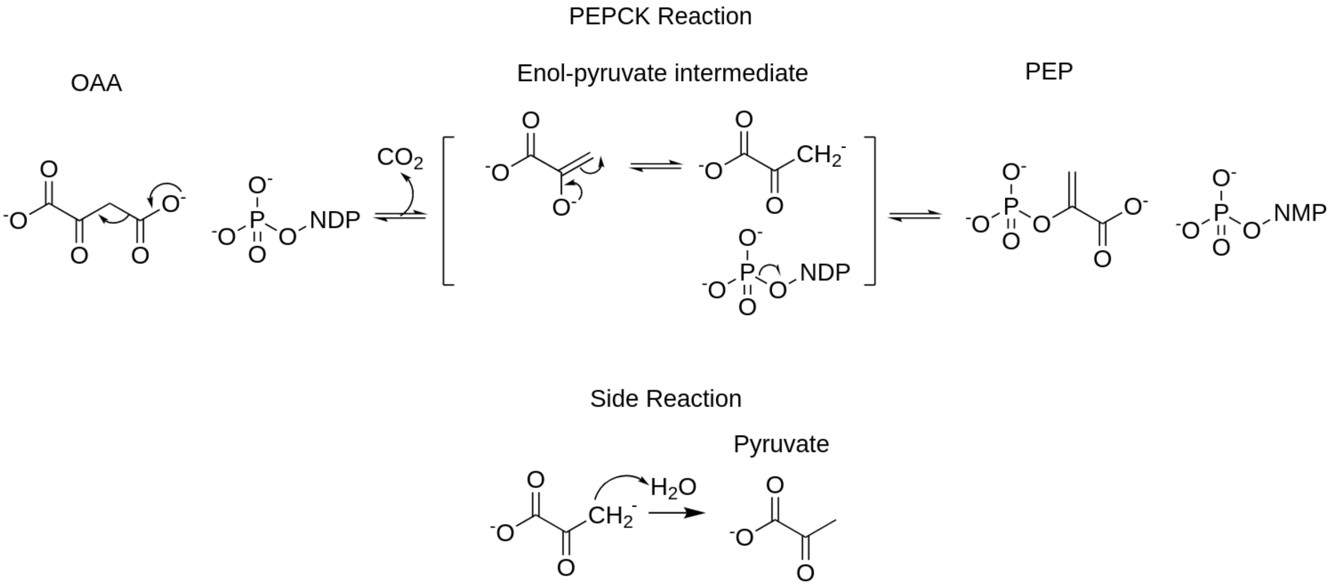


Figure 1.5: Reaction scheme for nucleotide PEPCK. ND(M)P represent nucleotide di- mono-phosphate. Reaction scheme was rendered in ChemDrawJS.

1.9.2 Binding of OAA and GTP

The first step of catalysis is the binding of the substrates OAA and GTP. Upon this binding and subsequent ordering of the active site loops, there is also a global transition of the N- and C-terminal domains rotating towards one another and the active site (Figure 1.6).[29]

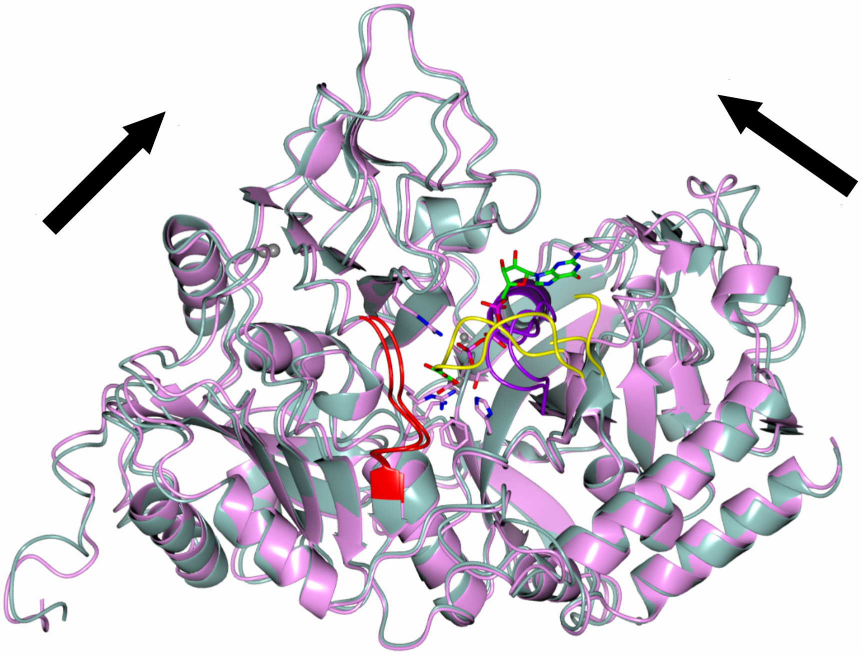


Figure 1.6: Global transitions through catalytic cycle. Superimposition of holo rcPEPCK (PDB ID# 2QEW – aqua) and oxalate-GTP bound rcPEPCK (PDB ID# 3DT2 – pink). The active site loops are coloured as R- P- Ω-loop in red, purple and yellow respectively. All atoms are coloured by type with oxygen as red, nitrogen as blue, and carbon as green.

OAA binds to the active site through coordination with the M1 metal. In addition to this, S286 of the P-loop interacts with O5 of OAA, and two arginines, R87 of the R-loop and R405, extensively coordinating with OAA.[29] After OAA (Beta-sulfopyruvate (βSP) is used in Figure 1.7, an OAA analogue) and GTP binding, the enzyme transitions from an open to closed state (Figure 1.7). In this transition, the nucleotide is shifted towards the M1 metal while also rearranging the polyphosphate tail

(Figure 1.7C). This change in nucleotide conformation is coincident with the closing motion of the N- and C-terminal domains and with the ordering of the active site loops.

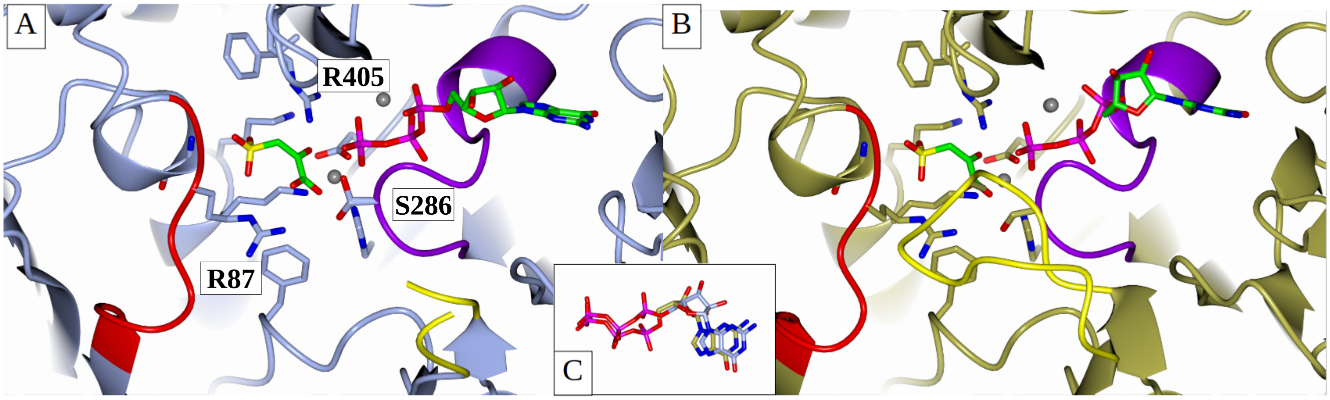


Figure 1.7: OAA and GTP binding. Panel A shows molecule B of βSP-GTP rcPEPCK complex (PDB ID# 3DT7 steel blue) in which the lid is open. Panel B shows molecule A (gold) of the same complex but the lid is now closed. Panel C shows the conformational change of GTP upon lid closure. Active site loops are coloured as R- P- Ω-loop in red, purple and yellow respectively. Atoms are coloured by type as carbon in green, sulfur as yellow, oxygen as red, nitrogen as blue, phosphate as fuchsia.

1.9.3 Decarboxylation and enolate intermediate

Prior to this phosphoryl transfer, OAA must be decarboxylated to form a reactive enol-pyruvate intermediate (represented by oxalate) (Figure 1.8). Structural evidence suggests that OAA is polarized through both R87 and N403, increasing the propensity for the carboxylate to break off. Y235 now rotates inwards to likely stabilize CO_2, although the actual CO_2 binding site has not been uncovered.[29] This reactive enolate is protected by the by the closure of the lid.[62]

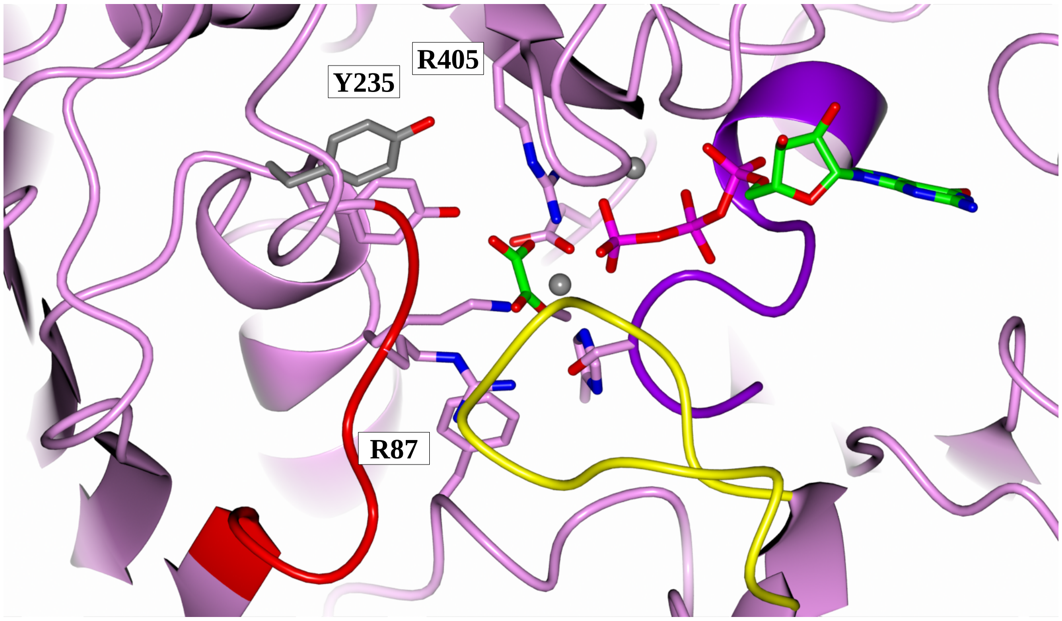


Figure 1.8: Enolate Intermediary Structure. This complex with oxalate and GTP (PDB ID# 3DT2) is shown in pink. Y235 and its transition upon decarboxylation of OAA is shown in grey (from PDB ID# 3DT7 molecule B). Active site loops are coloured as R- P- Ω-loop in red, purple and yellow respectively. Atoms are coloured by type as carbon in green, sulfur as yellow, oxygen as red, nitrogen as blue, phosphate as fuchsia.

1.9.4 Phosphorylation and release of PEP and GDP

The γ-phosphate of GTP can now be transferred to the reactive enolate to form PEP. In doing so, PEP is likely initially found in a conformation where the phosphate is occupying the γ-phosphate binding pocket as seen in the phosphglycolate (PEP analogue) – GDP structure (PDB ID# 3DTB). In order for release to occur, PEP then moves away from its direct coordination with the M1 and M2 metals (Figure 1.9A), and shifts to an outer-sphere coordination geometry (Figure 1.9B). In this position, the Ω-loop opens again to allow for PEP and GDP to be released. To accommodate PEP in the outer shell Y235 must rotate away from the active site and forms an edge-on interaction with the carboxylate of PEP (Figure 1.8 and Figure 1.9).[63]

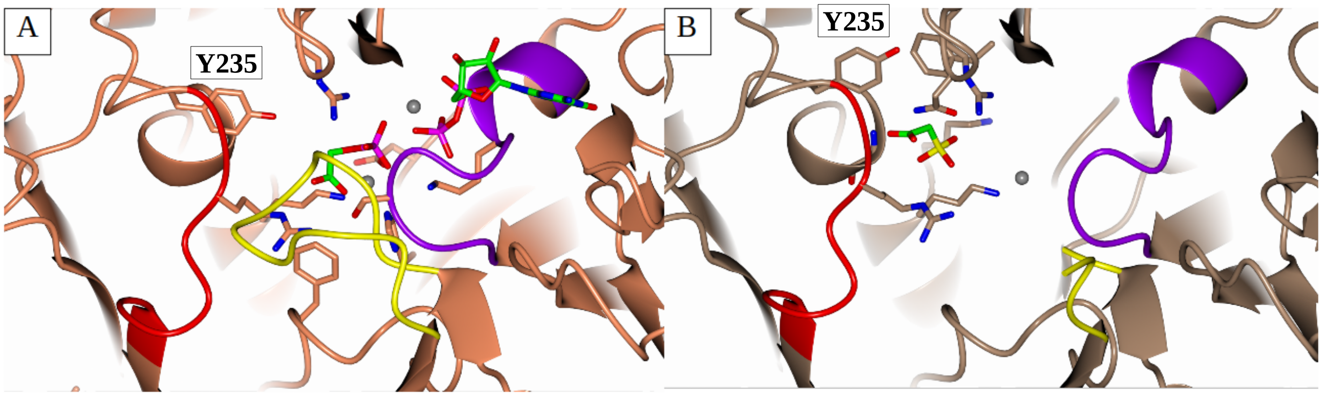


Figure 1.9: PEP and GDP binding. Panel A shows GDP and phosphoglycolate (PEP analogue) (PDB ID# 3DTB molecule A) at the suspected PEP binding conformation after phosphoryl transfer. Panel B shows PEPCK bound with sulfoacetate (PEP analogue) (PDB ID# 2RKE) at the outer-shell conformation prior to release. Active site loops are coloured as R-, P-, Ω-loop in red, purple and yellow respectively. Atoms are coloured by type as carbon in green, sulfur as yellow, oxygen as red, nitrogen as blue, phosphate as fuchsia.

With this biochemical understanding of the snapshots during turnover in nucleotide dependent PEPCKs, further studies to elucidate the coupling of these transitions and their relationship to catalysis are important to have a comprehensive understanding of this enzyme. The studies below highlight how modifications of the enzyme may not alter these end-points, but rather the ability to transition from one state to the next. These transitions and the accompanied changes in the free-energy landscape will lend new insight in the molecular motions important to the function PEPCK.

2.0 Understanding a Hysteretic PEPCK Mutant Through a Structural Lens

2.1 Introduction

Enzymes and their free-energy landscapes can be modulated and changed under differing environmental conditions. Of these conditions, the enzyme itself can be tuned through covalent and non-covalent modifications. In a simplistic sense, these modifications change the structure and dynamics, which alter the function of an enzyme by either increasing or decreasing its activity. These covalent PTM's can include but are not inclusive to; phosphorylation, oxidation, lipidation, and glycoslyation.(reviewed in 64) Non-covalent modifiers are small molecules which bind and alter the enzymes activity. Most of these are allosteric modulators have been studied with respect to modifying the quaternary state of an enzyme, although monomeric enzymes are now being recognized as being under the influence of these small molecules.[65,66]

The tune-ability described above can be a result of the change in the free-energy landscape of the modified enzyme, as well as other copies with which they are interacting within its quaternary state. This cooperativity between copies can either be positive or negative. The most well studied and recognized example of cooperativity is in hemoglobin. These studies showed the cross-talk between different copies of the tetramer where upon binding of oxygen in the first active site, the affinity for the other copies sequentially increased.(reviewed in 67) Cooperativity is also related to the time-dependent phenomenon of hysteresis. Although they are readily exchanged, the distinction between the two is that cooperativity can affect both kinetics, as well as affinity (as seen in hemoglobin), whereas hysteresis is only affecting the kinetic response.[68] The time-dependency of hysteresis is described by a slow response to the addition of substrate on the activity of the enzyme. These hysteretic lags or bursts were recognized very early, although have not been studied in depth to determine specific causes of this atypical behaviour. Seminal work by Frieden outlined common characteristics of hysteresis.[68] These common characteristics will be described in the context of a lag in activity for a monomeric enzyme, however multimeric enzymes can also undergo hysteresis through these same principles. With a non-hysteretic enzyme, the time required for a rate limiting step will be on average, the same for each molecule, and the highest free-energy barrier will dictate the maximal rate of turnover. Thus, for steady-state reactions, the rate of product formation is constant over the experimental time-frame. In the hysteretic enzyme, there is a change in the free-energy landscape with time, and therefore either the height of the barrier, or potentially which barrier is highest, changes for the rate limiting step. In general, two forms of the enzyme with variable barrier heights most simply represent this change in activity. When a reaction begins, the free enzyme will bind its substrate and form product at a given rate. If this reaction rate increases over time, the second form of the enzyme must have a different, lower, free-energy barrier, and thus is structurally and functionally distinct from the initial population of reacting enzyme. Also, if the only changing experimental parameter with time is the enzyme completing catalysis, the changes it undergoes during catalysis are producing the high-activity form. Hysteresis is best formalized with two models. This first of which is the LIST (ligand induced slow transition) model.[69,70] This model is an adaptation of MWC (Monad-Wyman-Changeux) model of allostery[71,72] and states that rather than the transfer from low to high activity states being solely induced fit (see [73] for alternative KNF (Koshland-Nemethy-Filmer) induced fit allostery model), it can be thought of as a change in equilibrium. It is likely that free enzyme can interconvert between the inactive and active forms through stochastic fluctuations, but the addition of substrate alters this equilibrium to favour that of the active form. A previous single-molecule study investigating β-glucuronidase suggested that thermal treatments only in the presence of substrates can alleviate hysteresis.[74] This suggests that an increase in the available thermal energy is not sufficient in forming the active form of the enzyme. This highlights that it is the transition through the catalytic cycle that produces the inactive state, and the presence of substrate more strongly shifts the equilibrium constant between the inactive and active form more than temperature. The second foundational model describing hysteresis is that under saturating conditions, the probability must be higher for the newly formed high-active state to turnover again to maintain this state, than being unbound long enough to be able to revert to its previous low energy form. This "memory" of the enzyme has been described as the active forms "remembers" and can relax back into the inactive form[75,76], where a slower memory implies a larger free-energy barrier between the active-inactive transition.

This transformation of a low-activity enzyme to high-activity enzyme can likely be accomplished by many mechanisms. What can be said is that under typical experimental conditions, the rates and duration of hysteresis are measured on time-scales of seconds to minutes.[68] This suggests that the transitions between low and high activity forms must be occurring also on these time scales, leaving one or many conformational changes as the most likely candidate.[68,72] Substrate binding at the active site may induce conformational change in the monomer, which then converts itself or sub-unit partner to a high activity state.[68] Alternatively, the slow response to substrate addition could also be due to an allosteric mechanism where the substrate can bind elsewhere, albeit slowly and tightly, and induce the transformation to the high-activity state.[68] Active site inhibitors, like in the allosteric mechanism, with slow and tight binding/release characteristics have been shown to manifest as burst hysteresis.[77]

Hysteretic enzymes can be great models in beginning to understand the free-energy landscape, as they can be experimentally probed throughout this process. As free-energy landscapes are impossible to precisely describe, the mechanisms of hysteresis are very challenging to parse apart. A point mutant of PEPCK has been shown to induce hysteresis, and with the mechanistic understanding of PEPCK, it may be an exceptional model to attempt to understand these causal mechanisms of hysteresis. The L153D mutation (Figure 2.1), has the aforementioned hysteretic lag in the PEP carboxylating reaction, while the OAA decarboxylating reaction is unaffected (Figure 2.2). This residue is found at the hinge region of the enzyme, in a surface loop. This hinge region is the pivot point for the N- and C-terminal domains to rotate towards one another in ordered to reduce the volume of the active site as well as position coordinating residues to do PEPCKs chemical steps. In the open state of PEPCK, this surface loop and leucine are ordered and buried towards the interior of the enzyme. As PEPCK transitions to a closed state, this loop becomes disordered and solvent exposed. The dynamic interplay between this hinge loop and global conformational changes provided the impetus behind changing this residue to an aspartic acid, becoming more hydrophilic and shifting its favourability to remain solvent exposed. Previous worked completed by Sarah Sullivan in the Holyoak Lab, showed kinetic changes for the enzyme after fully activated, but further structural characterizations are needed to understand the structural basis for this observed phenomenon. I aimed to determine various structures of this mutant in order to begin to rationalize the kinetic findings, as well as being to describe the conformational changes that may be responsible for this behaviour.

Figure 2.1: L153D mutant location. This figure is of the holo enzyme (PDB#2QEW) showing in yellow, the location of the L153D mutant situated between the N- (red) and C-terminal (green) domains at the hinge of the enzyme.

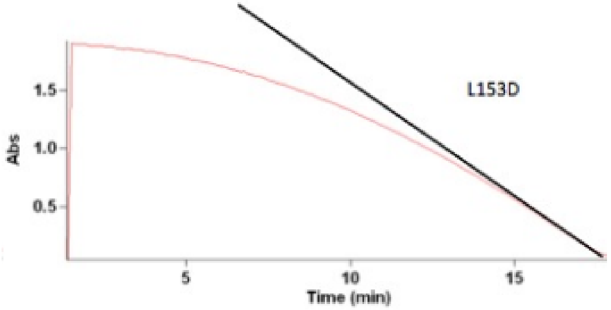


Figure 2.2: Representative kinetic trace from hysteretic mutant. This figure shows a kinetic trace readout from spectrophotometric measurements of the L153D mutant. The black line indicates the slope which was utilized to determine kinetic constants. This trace was collected by Sarah Sullivan.

2.2 Methods

2.2.1 Construct Design

Rattus norvegicus (rat cytosolic) (NP_942075/XP_342593) was cloned as described elsewhere into pESUMO-Star (Kan) vector (LifeSensors) between the BSA1 and XHO1 sites so that the final expressed protein had the SUMO 6xHIS fusion tag at the N-terminus.[39] Upon cleavage of the fused protein, the resultant protein would have no additional amino acids.

2.2.2 Expression

The expression of all GTP-dependent PEPCKs studied were completed as followed. The purification buffers are show in Table 2.1.

Table 2.1: Buffer composition for protein purification of GTP PEPCKs.

Buffer 1 (B1)	25mM HEPES (pH 7.5 at RoomTemperature) + 10mM Imidazole + 300mM NaCl + 2mM TCEP
Buffer 2 (B2)	25mM HEPES (pH 7.5 at RoomTemperature) + 300mM Imidazole + 2mM TCEP
Buffer 3 (B3)	25mM HEPES (pH 7.5 at RoomTemperature) + 2mM TCEP
Buffer 4 (B4)	25mM HEPES (pH 7.5 at RoomTemperature) + 10mM DTT

Four flasks with 50mL (supplemented with 50ug/mL of kanamycin) of LB media was inoculated and grown overnight at 37°C from glycerol stocks of PEPCK (BL21-DE3). The cells were centrifuged for 5 minutes at 5000g and the resultant pellets were resuspended into four flasks with 1L of ZYP-5052 media and 50ug/mL of kanamycin.[78] These cells were allowed to grow for ~20 hours at 20°C before they were harvested through centrifugation for 10 minutes at 5000G. The cell pellets were resuspended in prechilled B1

2.2.3 Purification

Cells were resolubilized in B1, and were passed through a prechilled French press twice at 1000PSI. All subsequent steps were completed at 4°C. The lysed cells were then centrifuged at 12000xG for 30-45 minutes. The supernatant was incubated with nickel-NTA resin for 1 hour. This resin was then rinsed with B1 until the flow-through reached an A_{280} below 0.1. B2 was than passed over the resin to elute the bound protein into 10mL fractions. These fractions were pooled and concentrated to approximately 5mL. The protein was then passed down a P6-DG column pre-equilibrated in B3. Now buffer exchanged, the protein was left overnight to stir with ~1mg of SUMO protease (purified as described in[38]) to cleave the 6xHIS SUMO fusion domain from the PEPCK. This digested protein sample was then incubated with cleaned nickel-NTA resin equilibrated in B3 for 1 hour. The resin was washed with B3 where the flow-through containing protein was collected in 10mL fractions and concentrated to less than 5mL. One final buffer exchange with P6-DG resin in B4 was completed. The final purified protein was concentrated to 40, 30 and 10mg/mL using $E_{1\%\,A280}$ extinction coefficients of 17 $mLmg^{-1}$ for rcPEPCK (L153D). The protein was frozen in 30uL drops by direct immersion in liquid nitrogen and stored at -80°C.

2.2.4 Crystallization

All L153D crystals were grown under the same conditions with variable amounts of ligands added. In brief, crystals were growth at $10mgmL^{-1}$ through vapour diffusion at room temperature. The mother liquor condition was 100mM HEPES pH 7.5 and 16-26% polyethylene glycol (PEG) 3350. Drops were set up in 2:4, 3:3, and 4:2 μL ratios of protein:mother liquor. After drops were set they were supplemented with 0.5μL of 100mM $MnCl_2$. Crystals were harvested, and soaked for 1 hour in 100mM

HEPES at pH 7.5, 30% PEG 3350, 10% PEG 400, and 5mM $MnCl_2$. Complexes of the enzyme with ligand was accomplished via a combination of cocrystallization and soaking (Table 2.2). The final crystals were cryocooled by direct immersion into liquid nitrogen.

Table 2.2: Ligand addition for the crystallization experiments. L153D GTP and ßSP complexes were collected at University of Waterloos' rotating copper anode x-ray source.

Structure	**Ligand Added to Drop**	**Ligand Added to Cryoprotectant**
L153D – GDP (GDP+PEP)	1mM GDP	1mM GDP
L153D – GDP+PEP	1mM GDP	1mM GDP + 1mM PEP
L153D – GTP+ßSP	1mM GTP	1mM GTP+ 1mM ßSP
L153D – GTP	1mM GTP	1mM GTP
L153D – ßSP	1mM ßSP	1mM ßSP
L15D – PEP	1mM PEP	1mM PEP
WT – GDP+PEP	1mM GDP	1mM GDP + 1mM PEP

Data was collected at the Canadian Light Source in Saskatoon, Saskatchewan on the 08B1-1 beamline using a Rayonix MX300HE detector or at University of Waterloo's home-source with Raxis IV++ image plate detector. Data were indexed and scaled in HKL-2000[79] and final data statistics are present in Table 2.3 and Table 2.4.

2.2.5 Structure Determination

All structures were initially solved through molecular replacement using MOLREP[80] in the CCP4 package[81] using the previously solved holo-rcPEPCK structure (PDB ID# 2QEW). These structures were analyzed to determine if they were in the open- or closed-states. Closed structures were then re-solved with the closed rcPEPCK oxalate-GTP structure (PDB ID# 3DT2). Iterative, manual real-space refinement and additions of water molecules and heteroatoms were carried out in COOT[82], followed by refinement using REFMAC5.[83] Final MolProbity statistics were collected with the online server at http://molprobity.biochem.duke.edu/index.php.[84] Final model statistics are present in Table 2.3 and Table 2.4.

Table 2.3: Crystallographic data table for the two ligand complexes. The GTP complex was collected at the University of Waterloos' rotating copper anode x-ray source.

	L153D GDP (GDP+PEP)	**L153D GDP+PEP**	**L153D GTP (GTP+Oxalate)**	**L153D GTP+BSP (GTP + BSP/Oxalate)**
Wavelength (Å)	1	1	1.54	1
Resolution range (Å)	53.2 - 2.1 (2.12 - 2.05)	48.9 - 1.7 (1.8 - 1.7)	29.4 - 2.29 (2.37 - 2.29)	52.2 - 1.7 (1.8 - 1.7)
Space group	P 1 21 1	P 1 21 1	P 1 21 1	P 1 21 1
Unit cell (Å)	60.02 118.874 86.406 90 97.521 90	60.04 118.866 86.59 90 97.396 90	45.724 119.511 60.454 90 110.952 90	60.083 118.823 86.621 90 97.273 90
Total reflections	2410808	2570756	638885	2620508
Unique reflections	73409 (7275)	129163 (12603)	25223 (1946)	132028 (13056)
Multiplicity	6.3 (5.9)	7.0 (6.8)	4.1 (3.2)	7.2 (6.7)
Completeness (%)	97.5 (97.1)	97.8 (95.8)	92.6 (71.7)	99.8 (99.1)
Mean I/sigma(I)	9.9 (1.3)	13.5 (1.5)	10.6 (1.8)	14.7 (2.3)
Wilson B-factor (Å^2)	29.4	22.5	52.3	21.8
R-merge	0.157 (1.32)	0.111 (1.36)	0.108 (0.628)	0.099 (0.815)
R-meas	0.171 (1.44)	0.120 (1.47)	0.123 (0.729)	0.106 (0.884)
R-pim	0.066 (0.585)	0.044 (0.560)	0.058(0.362)	0.039 (0.339)
CC1/2	0.909 (0.627)	0.947 (0.712)	0.937 (0.776)	0.966 (0.856)
Reflections used in refinement	73381 (7275)	129081 (12597)	25222 (1947)	131931 (13053)
Reflections used for R-free	3658 (361)	6639 (671)	1164 (85)	6783 (695)
R-work	0.228 (0.318)	0.194 (0.296)	0.210 (0.298)	0.196 (0.261)
R-free	0.268 (0.342)	0.221 (0.325)	0.271 (0.379)	0.222 (0.309)
Number of non-hydrogen atom	10092	10444	4927	10487
macromolecules	9724	9738	4815	9730
ligands	100	97	47	111
solvent	268	609	65	646
Protein residues	1238	1238	613	1238
RMS(bonds)	0.013	0.014	0.014	0.015
RMS(angles)	1.85	1.87	1.86	1.93
Ramachandran favored (%)	95.5	96.9	94	97
Ramachandran allowed (%)	4.05	3	5.56	2.84
Ramachandran outliers (%)	0.41	0.08	0.49	0.16
Rotamer outliers (%)	1.94	1.55	4.89	1.65
Clashscore	5.03	2.97	6.85	3.69
Average B-factor	32.7	25.4	58.3	23.8
macromolecules	32.8	25.2	58.5	23.5
ligands	31	22.6	56.4	24.2

Table 2.4: Crystallographic data table for the one ligand complexes. The ßSP complex was collected at the University of Waterloos' rotating copper anode x-ray source.

	WT GDP+PEP (GTP)	**L153D BSP**	**L153D PEP**
Wavelength (Å)	1	1.54	1
Resolution range (Å)	59.3 - 1.7 (1.7 - 1.7)	31.4 - 2.05 (2.12 - 2.05)	42.4 - 1.7 (1.76 - 1.7)
Space group	P 1 21 1	P 1 21 1	P 1 21 1
Unit cell (Å)	62.258 118.654 86.171 90 107.771 90	62.94 119.249 86.955 90 107.612 90	60.989 118.664 90.688 90 108.032 90
Total reflections	2140022	1740850	2129427
Unique reflections	138000 (12002)	72791 (5812)	133590 (12880)
Multiplicity	7.3 (5.4)	7.1 (6.0)	7.3 (6.0)
Completeness (%)	96.5 (84.4)	94.9 (76.2)	99.2 (96.1)
Mean I/sigma(I)	48.4 (2.3)	14.9 (2.1)	23.4 (2.2)
Wilson B-factor (Å^{2})	26.9	36.4	30.3
R-merge	0.065 (0.641)	0.114 (0.765)	0.06 (0.74)
R-meas	0.07 (0.708)	0.123 (0.831)	0.064 (0.806)
R-pim	0.026 (0.293)	0.045 (0.318)	0.023 (0.314)
CC1/2	0.987 (0.928)	0.967 (0.86)	0.985 (0.904)
Reflections used in refinement	137671 (11989)	72679 (5809)	133413 (12859)
Reflections used for R-free	6861 (573)	3671 (250)	6567 (614)
R-work	0.246 (0.534)	0.309 (0.402)	0.221 (0.391)
R-free	0.272 (0.548)	0.377 (0.439)	0.25 (0.408)
Number of non-hydrogen atom	10267	9576	9986
macromolecules	9631	9542	9560
ligands	68	32	24
solvent	568	2	402
Protein residues	1225	1214	1214
RMS(bonds)	0.015	0.015	0.014
RMS(angles)	1.92	1.96	1.88
Ramachandran favored (%)	97.1	90.4	95.9
Ramachandran allowed (%)	2.79	7.71	3.32
Ramachandran outliers (%)	0.08	1.91	0.75
Rotamer outliers (%)	2.73	4.83	1.57
Clashscore	3.43	9.63	4.84
Average B-factor	32	44	36.2
macromolecules	31.9	44.1	36.2
ligands	27.9	35.8	31.6
solvent	35.3	26	37.3

2.3 Results

2.3.1 L153D - PEP

The PEP L153D structure is as expected, with PEP binding in the outer-shell of the active site in the PEP release state (Figure 2.3). The enzyme globally is open, most similar to the holo structure (PDB ID# 2QEW). The R- and P-loop are both ordered but also in an open state, while the Ω-loop is open and disordered. Interestingly, there are differences in the two molecules of this structure in the ASU. Molecule A, is very well ordered throughout the entire structure. Molecule B on the other hand is ordered throughout the N-terminal domain but becomes quite disordered through the PEPCK-(C terminal) domain (residues 310-400 and 500-580).

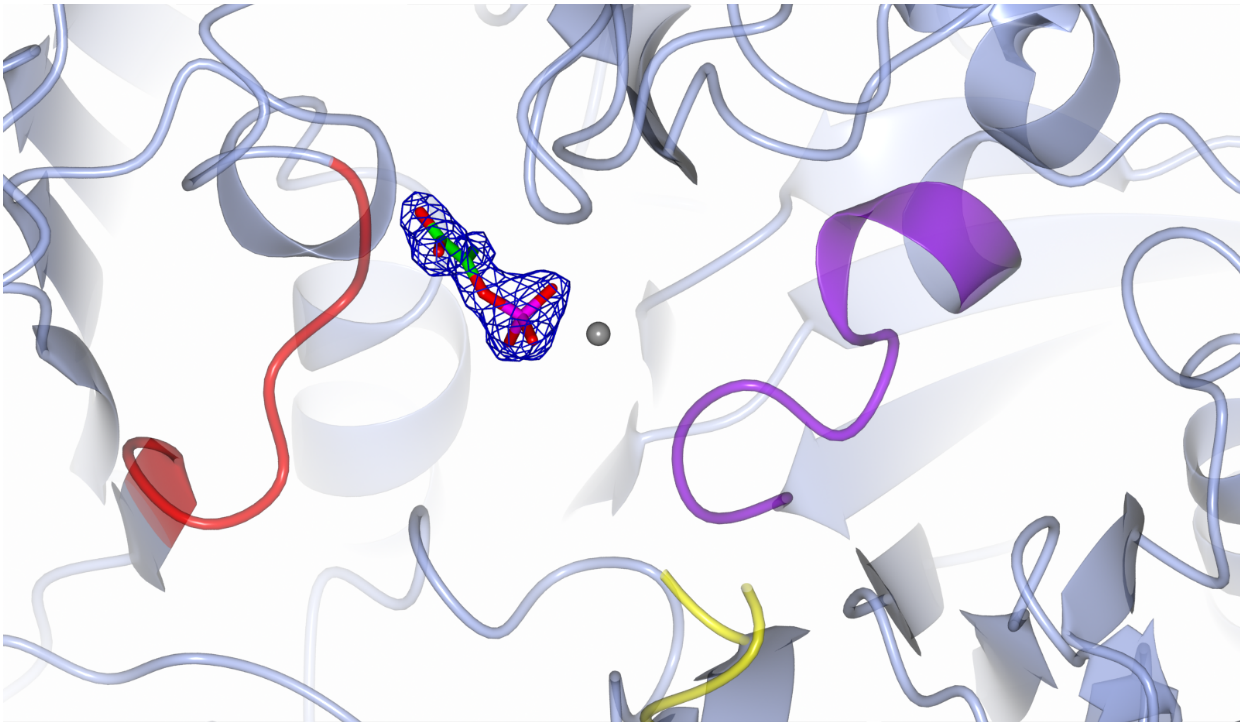


Figure 2.3: PEP bound to L153D. This figure shows PEP and M1 metal bound to the hysteretic mutant with the experimental $2F_o$-F_c electron (rendered at 1.7σ) density maps in blue mesh. The density is cropped for clarity. The R-, P-, and Ω-loop are shown in red, purple and yellow respectively. The various ligands are coloured by atom type, with carbon as green, oxygen as red, nitrogen as blue, phosphorous as purple, and Mn^{2+} as grey.

2.3.2 L153D - βSP

The location of βSP is unexpected, with βSP binding in the outer-shell of the active site in the PEP release state in both of molecules in the ASU (Figure 2.4). In molecule A, βSP is also in an the expected conformation coordinating with the M1 metal. Due to resolution constraints, the modelling of βSP in the molecule with two conformations was challenging and only the general location could be inferred. Both molecules are globally open, although the C-terminal domain is slightly rotated towards the active site. The R-loop is ordered but in an open state, the P-loop is ordered in molecule A, but disordered in molecule B, and finally the Ω-loop is open and disordered. As discussed above in section 2.3.1, molecule A is very well ordered throughout the entire structure while molecule B is ordered throughout the N-terminal domain but becomes quite disordered through the PEPCK-(C terminal) domain (residues 330-380 and 400-435).

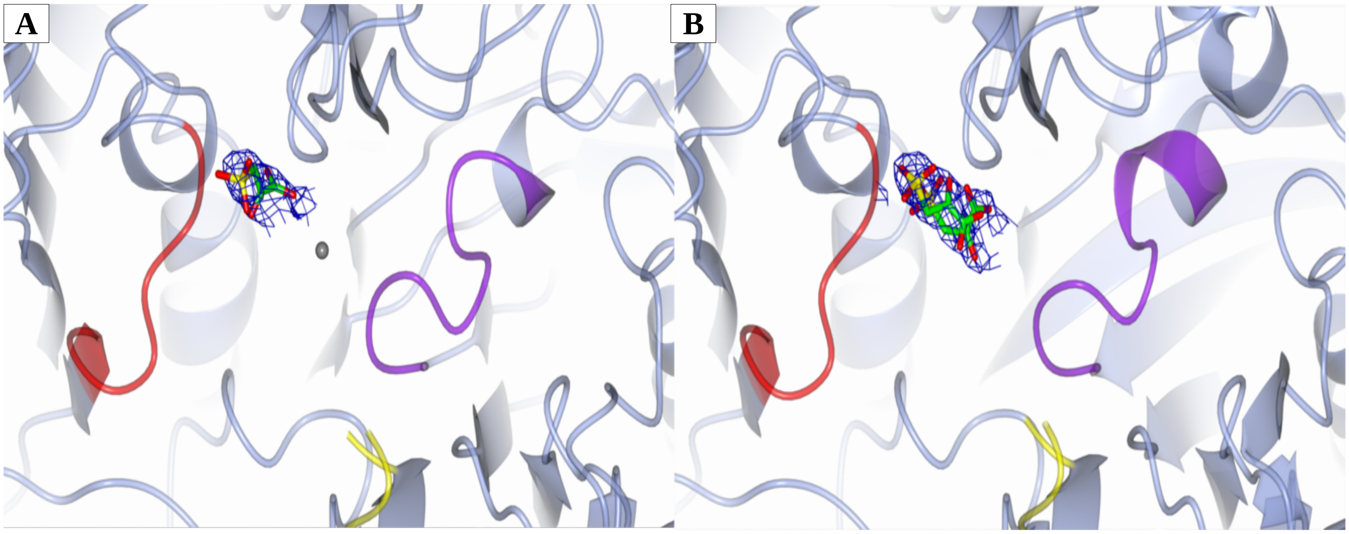


Figure 2.4: βSP bound to L153D. This figure shows βSP and M1 metal bound to the hysteretic mutant with the experimental 2F_o-F_c (rendered at 1.5σ) electron density maps in blue mesh. The density is cropped for clarity. Panel A shows only one βSP bound in the PEP site. Panel B shows βSP bound both in the PEP site, as well as in the OAA site coordinating with the M1 metal. The R-, P-, and Ω-loop are shown in red, purple and yellow respectively. The various ligands are coloured by atom type, with carbon as green, oxygen as red, nitrogen as blue, phosphorous as purple, and Mn^{2+} as grey.

2.3.3 L153D - GTP (GTP + Oxalate)

The location of GTP is as expected, binding to the P-loop (Figure 2.5). After modelling the nucleotide and M1/M2 metals, there was positive F_o-F_c electron density around the M1 metal facing the PEP/OAA site. Although the resolution is relatively low, oxalate was easily modelled in this positive density coordinating with the M1 metal as previously described. The enzyme is globally closed (PDB ID# 3DT2). The R- and P-loop are ordered and closed, while the Ω-loop is partially disordered. Throughout each molecule, both the N- and C-terminal domains they are ordered and well modelled.

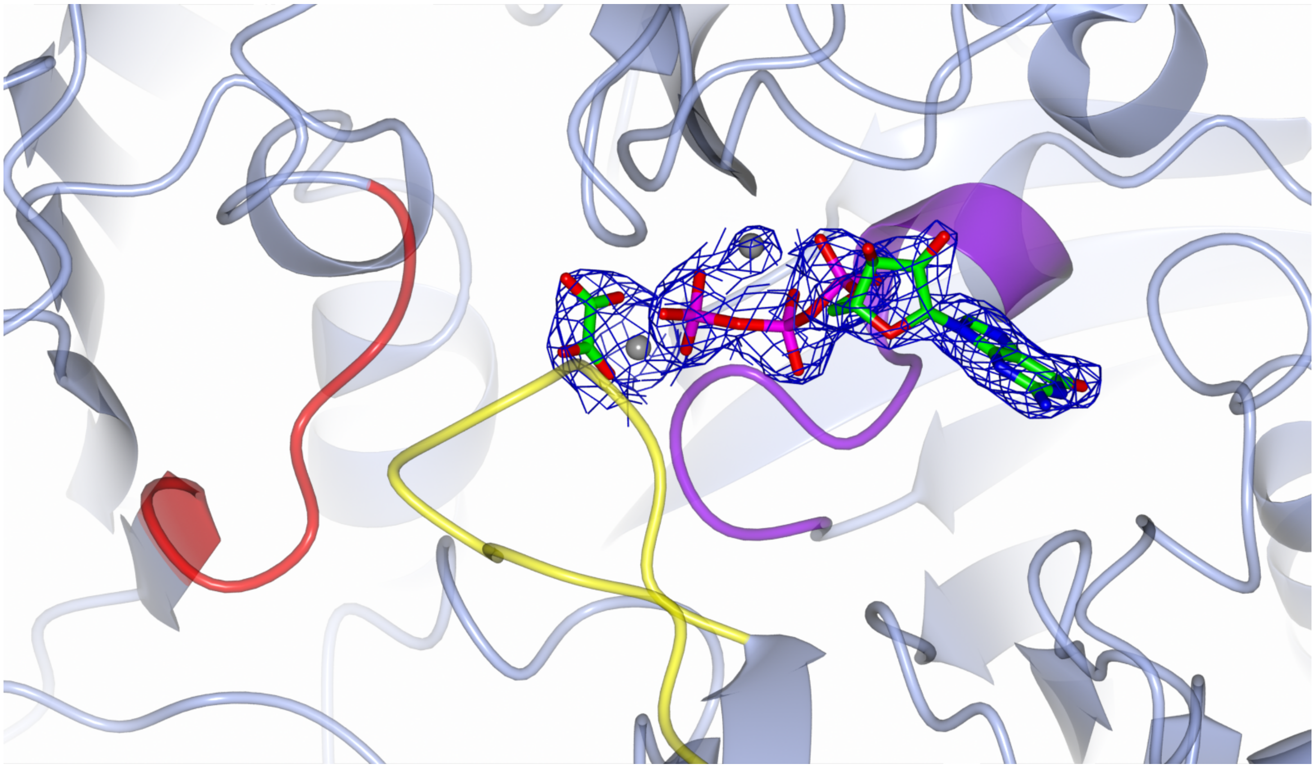


Figure 2.5: GTP structure (oxalate and GTP bound) with L153D. This figure shows oxalate, GTP and metals bound to the hysteretic mutant with the experimental $2F_o$-F_c electron density maps (rendered at 1.7σ for ligands, 3.5σ for metals) in blue mesh. The density is cropped for clarity. The R-, P-, and Ω-loop are shown in red, purple and yellow respectively. The various ligands are coloured by atom type, with carbon as green, oxygen as red, nitrogen as blue, phosphorous as purple, and Mn^{2+} as grey.

2.3.4 L153D - GTP + βSP

The location of GTP and βSP is as expected, binding to the P-loop and R loop while coordinating with the M1 metal respectively (Figure 2.6). After modelling the nucleotide, M1/M2 metals, and βSP there was negative F_o-F_c density around the sulfate. Reducing the occupancy βSP to 0.5 resolved this but created positive difference density around the carboxylate and carbonyl moieties. Oxalate modelled at the remaining 0.5 occupancy resolved well, suggesting a mixed population of βSP and oxalate. Both molecules in the ASU are globally closed (PDB ID# 3DT2). The R- and P-loop is ordered and closed, while the Ω-loop is also ordered and closed. Throughout each molecule, both the N- and C-terminal domains they are ordered and well modelled.

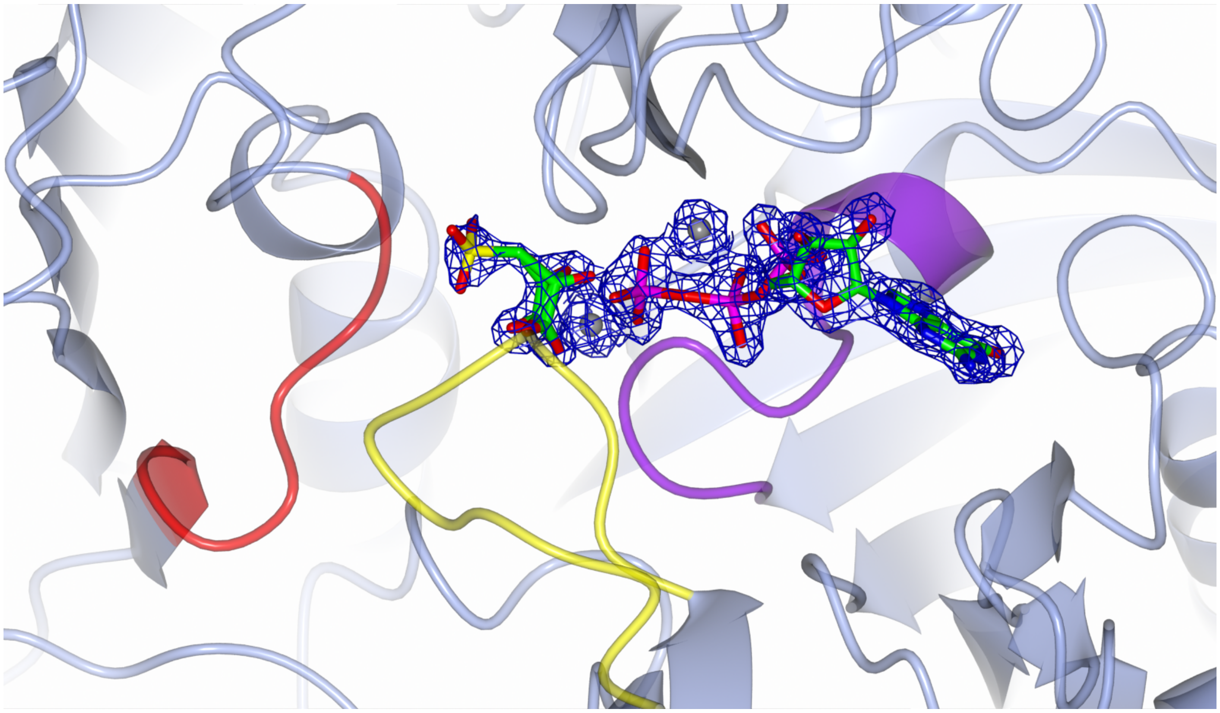


Figure 2.6: GTP and βSP structure (βSP/oxalate and GTP bound) with L153D. This figure shows oxalate, GTP and metals bound to the hysteretic mutant with the experimental $2F_o$-F_c electron density maps in blue mesh (rendered at 1.7σ for ligands, 3.5σ for metals). The density is cropped for clarity. Both βSP and oxalate are modelled with 0.5 occupancy. The R-, P-, and Ω-loop are shown in red, purple and yellow respectively. The various ligands are coloured by atom type, with carbon as green, oxygen as red, nitrogen as blue, phosphorous as purple, and Mn^{2+} as grey.

2.3.5 L153D – GDP+PEP and GDP (GDP+PEP)

The GDP is bound as expected and previously described coordinating to the P-loop (Figure 2.7). PEP however, is modelled cradled around the M1 metal, in a geometry that could be expected in a catalytically competent conformation prior to phosphoryl transfer in the OAA forming direction. Unexpectedly, this ligand complex was also found the the cocrystal structure of GDP by itself. There are two molecules in the ASU. Both molecules in the ASU are globally closed (PDB ID# 3DT2). The R-, P-, and Ω-loop is ordered and closed. Throughout each molecule, both the N- and C-terminal domains they are ordered and well modelled.

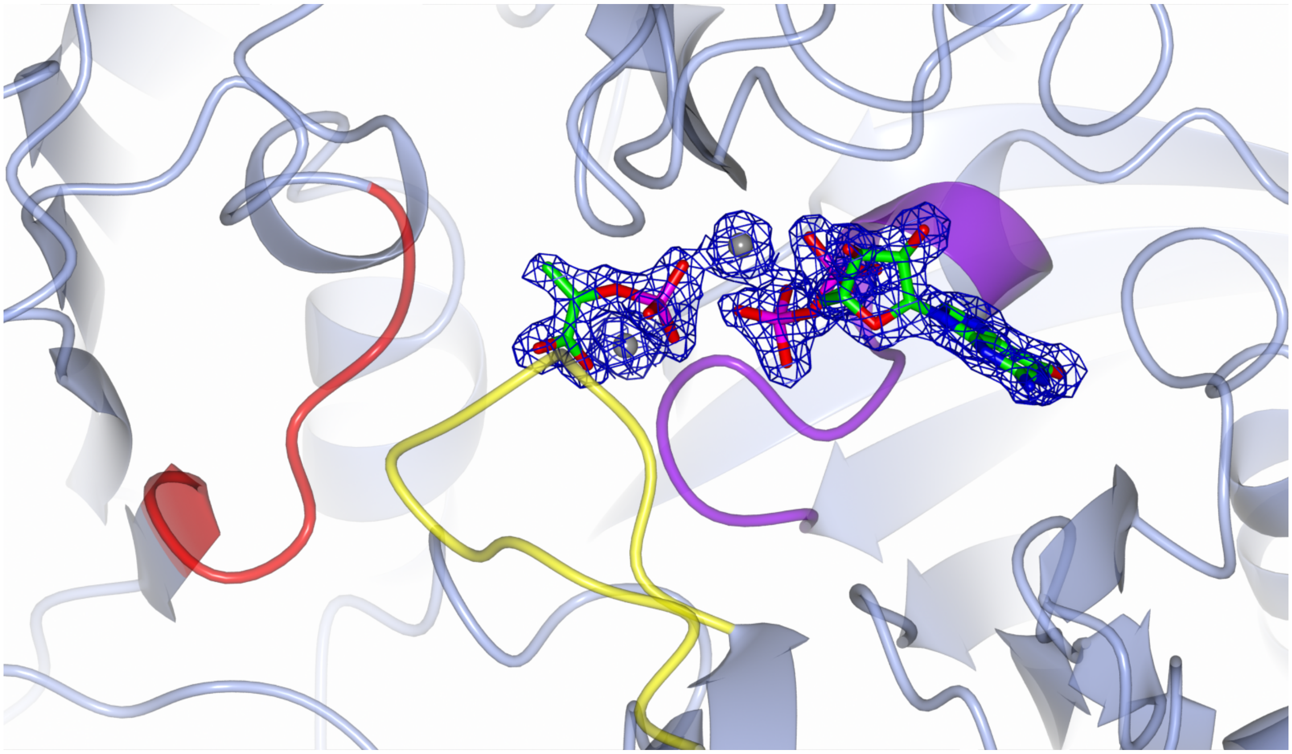


Figure 2.7: GDP and PEP bound to L153D. This figure shows the GDP and PEP structure, but also represents to findings from the GDP structure. This figure shows PEP, GDP and metals bound to the hysteretic mutant with the experimental $2F_o$-F_c electron density maps (rendered at 1.7σ for ligands, 3.5σ for metals) in blue mesh. The density is cropped for clarity. The R-, P-, and Ω-loop are shown in red, purple and yellow respectively. The various ligands are coloured by atom type, with carbon as green, oxygen as red, nitrogen as blue, phosphorous as purple, and Mn^{2+} as grey.

2.3.6 WT - GDP + PEP (GTP)

The GDP+PEP complex was found to have only GTP found in the active site coordinating as previously described to the P-loop (Figure 2.8). There is one molecule in the ASU. Both molecules in the ASU are globally open (PDB ID# 2QEW). The R-loop is open and disordered, while the P-loop is ordered and partially closed as seen in the previous WT GTP bound structure (PDB ID# 2QEY). Finally, the Ω-loop is disordered and open. Both the N- and C-terminal domains they are ordered and well modelled.

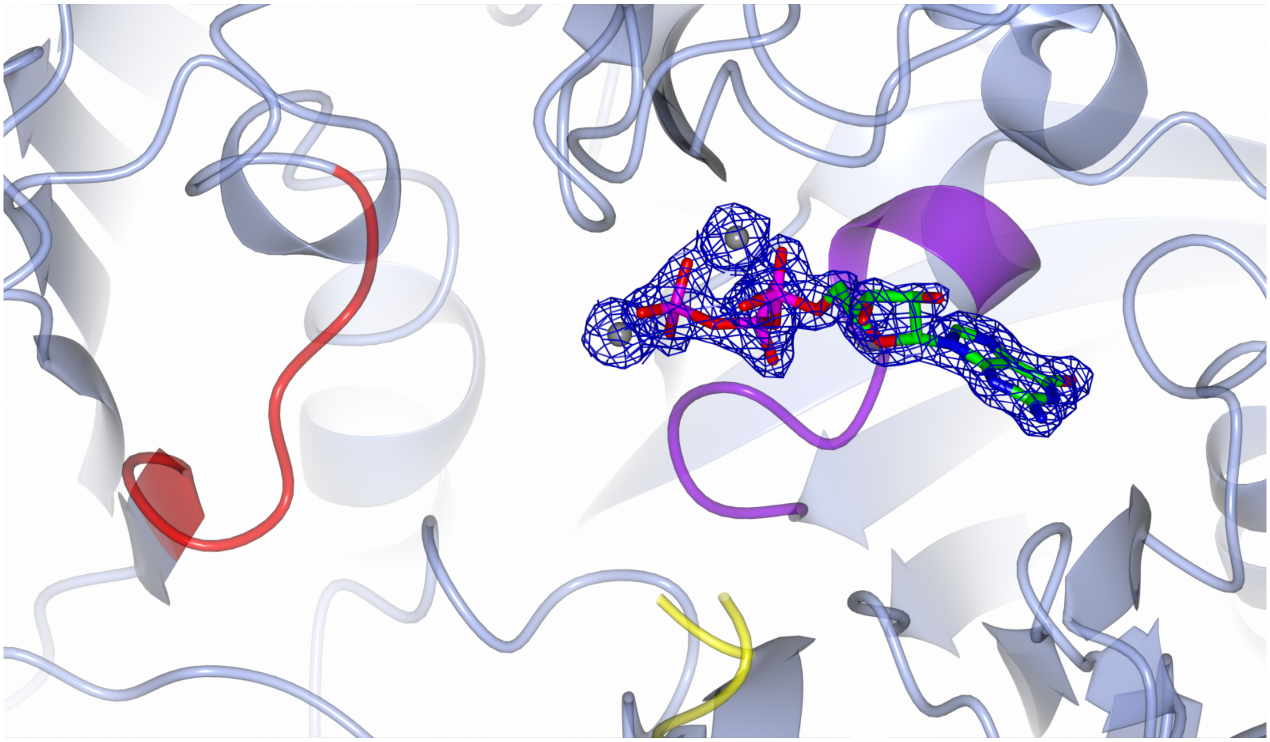


Figure 2.8: GDP and PEP structure (GTP bound) with WT rcPEPCK. This figure shows GTP and metals bound with the experimental $2F_o$-F_c electron density maps (rendered at 1.7σ for ligand, 3.5σ for metals) in blue mesh. The density is cropped for clarity. The R-, P-, and Ω-loop are shown in red, purple and yellow respectively. The various ligands are coloured by atom type, with carbon as green, oxygen as red, nitrogen as blue, phosphorous as purple, and Mn^{2+} as grey.

2.3.7 Michaelis-Menten Constants for the L153D and WT rcPEPCK

This data in Table 2.5, collected by Sarah Sullivan, is being used as a reference to understand the structural changes and these kinetic consequences. These constants were determined once the enzyme was fully activated. As seen, the mutant is turnover (k_{cat}) in the OAA decarboxylating reaction is reduced by a factor of 5, while the k_{cat}/K_m is unaffected for OAA but is an order of magnitude lowered for GTP. In the hysteretic reaction, the maximal activity is reduced by approximately one third, while the k_{cat}/K_m is increased by a factor of two, while the GDP k_{cat}/K_m is unaffected.

Table 2.5: Previously determined kinetic parameters for WT and L153D mutant.

OAA + GTP → PEP + GDP + CO_2

Enzyme	**K_m OAA (μM)**	**K_m GTP (μM)**	**k_{cat} (s^{-1})**	**k_{cat}/K_m ($M^{-1}s^{-1}$) OAA**	**k_{cat}/K_m ($M^{-1}s^{-1}$) GTP**
WT	51	55	52	1×10^6	1×10^6
L153D	10	84	9	9×10^5	1×10^5

PEP + GDP + CO_2 → OAA + GTP

Enzyme	**K_m PEP (μM)**	**K_m GDP (μM)**	**k_{cat} (s^{-1})**	**k_{cat}/K_m ($M^{-1}s^{-1}$) PEP**	**k_{cat}/K_m ($M^{-1}s^{-1}$) GDP**
WT	170	210	19	4×10^4	9×10^4
L153D	100	140	13	1×10^5	9×10^4

All rates measured in the PEP carboxylating hysteretic reaction were after ~15 minutes once the enzyme was fully activated

2.4 Discussion

Before beginning the discussion, a few assumptions should be clarified. As described above, there are many possibilities for hysteretic behaviour, many of which can be ruled out when using rcPEPCK as a model. First, this enzyme has not shown to possess higher oligomeric states and therefore can be ruled out that a gain-of-function is likely not occurring. Second, although an allosteric pocket has been shown for this enzyme behind the P-loop of the active site this is likely not mediating the hysteresis. The previous characterization of the allosteric pocket was regulated by small inhibitors[65], and more recently by anions.[85] It is very unlikely that this allosteric pocket is now being regulated by substrate addition. The structural results also support this conclusion as they do not show any occupancy in this pocket either. The only two plausible mechanisms remaining are therefore a change in the conformational ensemble and free-energy landscape leading to a more active conformer after turnover, or the production of a new allosteric pocket – likely at the mutational site. With a suitable crystal form that can form open, closed, and intermediary states, high resolution data can be obtained to attempt to discern the mechanistic origins of hysteresis in L153D PEPCK, and give further insights into hysteretic behaviour in general.

In this study, the PEP and βSP bound structure serve as the initial inactive states, while the complexes of both PEP-GDP and βSP-GTP serve as the end-points for each reaction prior to activation. The oxalate-GTP/GDP complexes would serve as the intermediary state. Although a complete ensemble of structures are likely necessary in fully comprehending the structural changes in the free-energy landscape, these initial structures can lend insight into distinguishing between the two potential mechanisms stated above.

2.4.1 Rationalizing Kinetic Data

The kinetic data (Table 2.5) described the kinetic consequences of the mutation. These measurements were completed after steady-state was achieved, some time after the hysteretic transition had completely occurred. This data shows that the mutation does not only cause the hysteretic behaviour but also changes the free-energy landscape from the hysteretic active enzyme when compared to WT. In both directions, the mutants total maximal rate of turnover as measured by k_{cat} is reduced, an effect that is clearly not bidirectional as the OAA decarboxylating reaction is affected more severely. If k_{cat} of WT is mostly controlled by rates of phosphoryl transfer[38], the new mutant k_{cat} may be at least partially driven by conformational changes associated with product release.

In the OAA decarboxylating reaction, the product release step may be impaired if the enzymes conformational equilibrium is now shifted to remain closed when in this product state (GDP-PEP-CO_2). It is evident from the structural data that the enzyme binds GDP alongside contaminating PEP and closes, suggesting a stickiness of this state. The closest WT structure available is a complex with phosphoglycolate (PEP analogue) and GDP (PDB ID# 3DTB).[65] These structures shows WT rcPEPCK open, suggesting that the mutation may be changing the free-energy landscape to form a thermodynamic well when GDP and PEP are present, leading the 5 times reduction in k_{cat} for this reaction.

In the PEP carboxylating reaction, the GTP-OAA state may be destabilized as both the GTP cocrystal, and GTP-βSP structure both seem to sequester contaminating oxalate, or some analogue like it. This would at face value suggest that L153D when bound with GTP, selects for the intermediary enolate over OAA, stabilizing the intermediate state. Inhibition data (not shown) testing oxalate against both reactions suggest that the binding constant for the enolate is unchanged from WT to the hysteretic mutant (K_i = 9μM for each). The lack of change in binding constant for oxalate suggests that rather than fully selecting for the intermediary state, L153D rather preferentially destabilizes the GTP-OAA state. Under this model, the modest reduction in k_{cat} is simply because of this the equilibrium between the intermediary state and product state is now shifted towards the enolate, potentially stalling this the carboxylation reaction, albeit less severely than the GDP-PEP stabilization in the OAA decarboxylating reaction.

The catalytic efficiency differences (k_{cat}/K_m) coincide with the structural data presented. The very modest increase in PEP binding by a factor of two could be due to the GDP form, preferentially adopting a more accepting mode for PEP binding, as evidence by the GDP structure having PEP bound. The order of magnitude decrease in binding of GTP for L153D than WT is harder to discern. The structural evidence may be found in PEP and βSP structures. Based on investigating the B-factors and electron density in the C-terminal domain, there seems to be a destabilization occurring in at least one of the molecules in the ASU of the solved structures. This destabilization is not seen in the most homologous structures of the WT enzyme (PDB ID# 2QF1 – OAA, PDB ID# 2RKA – phosphoglycolate).[29] Although these structures are not directly comparative and the crystal packing is different, it still suggest that L153D is the cause of this destabilization. Destabilization of the C-terminal domain was shown in a previously characterized mutant, E89A/D/Q[40]. This mutant was designed to remove an important interaction linking OAA/PEP binding to the closure of the lid. The E89A mutation specifically led to a closed R-loop in which it resides, but the lid was only able to partially close, as a salt bridge was removed between E89 and H470. This partial closure of the Ω-loop likely manifested as a lowly populated second conformer of the C-terminal domain. Although the C-terminal domain is disordered in the L153D and E89 mutants, it is unlikely that this is caused by the same phenomena as the disorder of the E89 mutants are when in the intermediary and Michaelis-Menten complexes, whereas in the L153D it is only when singly occupied by βSP or PEP. It does highlight that the C-terminal domain may be a more dynamic domain of the enzyme, while the N-terminal domain may be more rigid in general. Regardless, this disorder in the L153D structures suggests that the holo enzyme, and PEP/OAA bound state, are more dynamic than the WT, which may manifest as increasing difficulty to turnover, or bind other substrates. As the destabilization is in the C-terminal domain, it is not surprising that the GTP binding is preferentially affected over OAA as the GTP binding domain is a part of the C-terminal portion. This destabilization preferentially effects the triphosphate nucleotide as opposed to the diphosphate nucleotide as suggested from the previous model of GTP-OAA destabilization and GDP-PEP stabilization as manifested in the structural data. It may be that the GDP-PEP stabilization can offset the binding penalties whereas the GTP-OAA destabilization cannot.

2.4.2 Proposed Mechanism of Hysteresis

The studied loop here, situated at the hinge between the N- and C-terminal domains, shows how the coupling of this loops energetic state relates to the overall conformation of the enzyme. There must be a specific coupling between the local transitions at this hinge with the ligation state and global rearrangements necessary for activity. Altering this loop through the point mutant tested here, has changed the energetics of the loop such that, in general, the enzyme has decoupled the ligation state with the conformations necessary for catalysis. This is best highlighted through the GDP cocrystal structure preferentially binding PEP without CO_2 and closing, or the GTP structure binding some enolate-like molecule, as modelled by oxalate. This ties the dynamics of the mutagenized loop with the dynamics throughout the rest of the enzyme and explains a potential origin for kinetic deficits measured.

To understand the specific mechanism of hysteresis, that is the ability for the enzyme to relieve this kinetic deficit by turning over, is likely achieved by an alleviation of the major thermodynamic barriers caused by the mutation. As described in the rationale behind the L153D mutation, this new hydrophilic residue will prefer a disordered solvent exposed state coincident with the global closed conformation. The observed structural data presented suggests that this mutant is behaving as predicted, by preferentially closing. In this above mentioned mechanism, potentially after turnover, the decoupling of the mutagenized loop is now reset to a more “normal” coupling, where this hinge loop will now transiently order and bury towards the enzyme, shifting the global conformational to an open state and preventing the enzyme from prematurely closing without full ligand occupancy.

The unidirectionality of the hysteresis cannot be explained by the destabilization of both the PEP and βSP states. When measuring the PEP carboxylation reaction, the saturating quantities of PEP and GDP close the enzyme before CO_2 binding, causing the lag until the landscape changes to remove it. In the OAA decarboxylating reaction, the saturating ligands are OAA and GTP, and the small amounts of PEP produced are only found in the active site after the chemical steps. Once released the PEP is removed by the coupling reaction (pyruvate kinase and lactate dehydrogenase), preventing rebinding and premature closure shown in the PEP carboxylating reaction with GDP. As mentioned previously, the sticky closure upon GDP-PEP binding would still occur in the OAA decarboxylating reaction, but only once after the end of chemistry, resulting the large reduction in k_{cat}. This bottleneck likely continuously persists and cannot be removed as evidence from the lack of hysteresis in this direction.

2.4.3 Allosteric Ligand Binding and Protein-Water Interface

As mentioned earlier, one plausible mechanism for hysteresis is allosteric regulation. One interest observation that requires attention is in many, but not all, WT closed structures of PEPCK a PEG molecule is found in the hinge region of the enzyme (Figure 2.9). This suggests that this region could be a site of allosteric regulation, possibly binding small molecules to influence to conformational equilibrium. L153(D), proximity to this PEG molecule is such that it is likely directly contributing to the binding. The L153D structure with PEP does have a PEG molecule present, whereas the WT phosphoglycolate structure does not. Alternatively, the L153D structure with βSP does not have a PEG, similar to its WT counterpart with OAA. Although PEG binding is not totally consistent with closure, it is an observation that cannot be ruled out. It is by all means possible that some other molecules other than PEG could be binding at this site, influencing its dynamics. Although unlikely, during the kinetic assays when the hysteresis is monitored, some molecules present may bind to this region and influence the hinge dynamics. Further studies should be completed to determine what the consequence of ligand binding at this site are, as well as probe for specific ligands if possible.

As this hinge region and mutation are all on the surface, it is likely having a dynamic effect within the protein as well as in regards to the solvent fluctuations. It has been noted the coupling of both α-(bulk) and β-(hydration shell) fluctuations to protein dynamics through a proposed “solvent-slaving” phenomena.(reviewed in 86) This mutation from a hydrophobic leucine to hydrophilic aspartic acid may also be influencing the conformational dynamics of the enzyme through the solvent interface. This, however, is experimentally impossible to determine without the use of *in silico* techniques and even then is likely an extremely large feat.

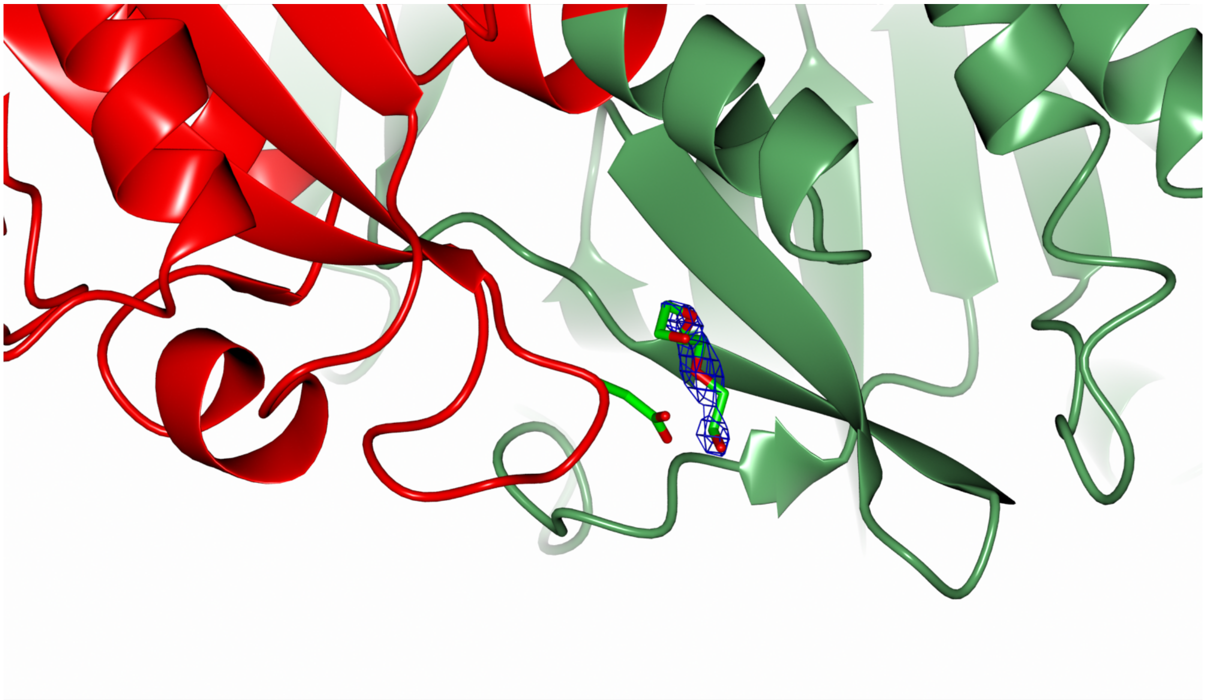


Figure 2.9: PEG Binding at hinge. This figure shows L153D PEP structure hinge region with PEG bound with experimental $2F_o$-F_c electron density maps (rendered at 1.5σ) in blue mesh. The density is cropped for clarity. The N- and C-terminal domains are coloured in red and green respectively. The various ligands and L153D mutation are coloured by atom type, with carbon as green, oxygen as red, and nitrogen as blue.

2.5 Summary of Results

Based on the known biochemical information of PEPCK, two causes of the observed hysteresis are possible. The first, and most likely, is the mutagenized loop is now preferentially stable in the disordered solvent exposed conformation, leading to favouring the global closed state. Throughout turnover, this hinge loop now samples more “normal” conformations effectively resetting itself, shifting its equilibrium closer to the proper wild-type open-closed states. This resetting changes in the energetic landscape leading to a more active enzyme as it can more readily traverse its catalytic cycle without prematurely closing. The second, could be a slow but tight binding allosteric regulator leading to a change in the energetic landscape.

The accompanied kinetic data can be rationalized with the presented structural data. The reduction in k_{cat}, specifically in the OAA decarboxylating reaction is likely due to a stalling of the enzyme PEP/GDP release. In contrast to this significant stalling, the modest reduction is k_{cat} for the PEP carboxylation reaction may be due to a destabilization of GTP-OAA state, and the equilibrium shift towards the enolate intermediary state. Clearly, based on the difference in k_{cat} between the two reactions, the equilibrium shift to stabilize the closed GDP-PEP state is shifted further than the GTP-OAA and enolate states. The increase in catalytic efficiency for PEP may be due to a stabilizing of the GDP bound state induce a conformational change that is with an increase affinity for PEP as the GDP crystal having PEP bound in a closed conformation. The reduction in catalytic efficiency in GTP may be due to a destabilization of the C-terminal domain, although the preferential closure once GDP and PEP bind may offset this penalty for GDP.

The structural evidence suggests that the unidirectional hysteresis may be caused by the mutants ability to close upon PEP and GDP binding leading to a low energy thermodynamic well. Under experimental condition, saturating quantities of PEP and GDP are present, which may induce premature closure before CO_2 can bind and the enzyme can turnover. This would manifest as a large reduction of the observed rate. After turnover, the enzyme has reset the coupling of this hinge loop with the transition to the closed state upon ligand bind. That is, the enzyme now the active form which remains open long enough for CO_2 to bind, and therefore turnover correctly. In the OAA decarboxylating reaction, this stalling occurs potentially after turnover, manifesting as a large reduction in k_{cat}, but will not manifest as burst kinetics (inactivation over time) as PEP is removed from the experiment through the coupling enzyme.

The allosteric preposition is plausible, as seen in the structural data PEG molecules can bind in this hinge region, preferentially when the enzyme is closed. Although PEG is not found in the kinetic assays, other molecules may be regulating this hinge region by interacting in a similar fashion to PEG. Although this allosteric binding should be studied further, because not every closed structure has PEG bound suggests that this binding may be not as important as other structural changes going on throughout the enzyme.

In closing, this single point mutant has conferred drastically different kinetics as well as structural differences that are unseen in the WT enzyme. This highlights how the single isozyme specific differences can cause significant changes in structural and function, showing the complexity of the energetic landscape. Further experiments, specifically addressing the transition between inactive and active conformations are necessary to be sure of the interpretations.

3.0 Thermal Adaptions of PEPCK

3.1 Introduction

In order to better understand the diversity of protein sequences and their corresponding energy landscapes, thermally adapted isozymes have long been studied as they must operate under vastly different environmental parameters while maintaining the same chemical outcome. In different thermal regimes, enzymes must adapt in order to maintain fulfilling cellular needs under differing thermal energy availability. Much of the focus on these studies has been on understanding the molecular mechanism that allow thermophilic enzymes to maintain their folded state at such high temperatures.[87] Mutational studies in attempts to create a thermal adapted enzymes from a mesophilic scaffold have been disappointing, indicating that specific residue changes are not enough to change the thermal dependency of the enzyme.[88,89] Understanding how the global features of the enzyme contribute to the function under different thermal conditions, although more convoluted, may lead to better outcomes.

With respect to flexibility and dynamics, thermally adapted proteins can be categorized as existing on a spectrum, where psychrophilic enzymes are generally proposed to be more flexible and the thermophilic counterparts are more rigid at a given temperature.[90] It was shown by Petsko, that the flexibility and dynamics of a given enzyme are conserved at the temperatures (~T_{opt}) the enzymes exist at *in vivo*.[91] This suggests that proteins adapt to maintain similar levels of kinetic flux and conformational motion, and therefore energetic landscape, at the temperatures that they have evolved to operate at. Psychrophiles, with their increase in flexibility generally have an accompanied decrease in thermostability, in contrast to rigid thermophilic enzymes that must maintain their tertiary structure at high temperatures where unfolding is more likely.[90] Thermophilic enzymes adapt for increased rigidity by increasing hydrophobic core packing, increased salt-bridge content, decreased interactions with solvent at the protein-solvent interface, as well as an increased number of prolines, increased π-stacking and cation interactions, and increased disulfide content.(review in 87) Initially, a two-state model for thermal denaturation was proposed. This classical model suggests that an enzyme-catalyzed rate will increase with temperature based on the Arrhenius equation, until an optimum has been reached. After this optimum, the enzyme thermally denatures and the activity drops rapidly. A new three-state model, known as the equilibrium model, has been proposed.[92,93] In this model, above the T_{opt} the enzyme does not initially denature but rather is folded and becomes reversibly less active with increasing temperature (with a K_{eq} associated with the equilibrium between folded active and folded inactive states). At sufficiently high enough temperatures, thermal denaturation occurs and the enzyme is irreversible inactivated. Thermal denaturation has been attributed to increased vibrations of atoms, disrupting intermolecular bond network and ultimately a collapse of the tertiary structure. Cold denaturation is much less understood. The proposed mechanism of cold denaturation is thought to be one where as temperatures decrease and surface waters becomes more ordered and function as a clathrate cage. This resultant removal of important protein hydration leads the protein to unfold, fueled by a favourable increase in entropy of the system.[94–97]

The inherent flexibility and rigidy of psychro- and thermophilic enzymes has functional consequences as well. The increased flexibility of psychrophiles gives them an increase in catalytic prowess at low temperatures over their rigid counterparts.[90] This flexibility allows these enzyme to use less thermal energy to transverse barriers of its free-energy landscape transitioning through substrate bound, transition, and product release states. Each free-energy barrier to move the enzyme-ligand system through this trajectory must be overcome through stochastic thermal motions. Low temperature flexibility adaptations help offset the decrease in thermal energy available by allowing these transitions to happen more readily, that is, by reducing the enthalpy of activation.[98] For thermophilic enzymes, the increased rigidity prevents denaturation, but also is tuned to utilize the amount of thermal energy available to traverse these structural transitions to maintain flux. These thermophiles are proposed to adapt their entropic barriers as the enthalpic ones are easily overcome with the thermal energy of their environments. Lowering of the entropy barriers is proposed to occur through a reduction in the sheer number of conformational states accessible. This change in the driving force in reducing the free-energy requirements of catalysis are known as the enthalpy-entropy compensation of adaptation.[99]

Functional and structural work detailing the conformational changes of PEPCK catalysis suggest that PEPCK may be a good system to further understand the thermal adaptations of dynamics, and thus the associated free-energy landscapes. For PEPCK to operate efficiently, there must be a series of concerted conformational changes of the R- and P-loop upon substrate binding with the maintained closure of the Ω-loop lid long enough for the chemical steps to take place.[38] It could be theorized that at high temperatures, both the intrinsic rates of chemistry and the entropic favourability of the disordered open state of the lid will increase simultaneously. Based on this hypothesis, one might speculate that PEPCKs must evolve the nature of this lid to tune the balance between the closing and opening of this lid with intrinsic response of the chemical reaction to temperature. For example, at low temperatures where chemistry is slow, the lid must stay closed long enough for the chemical steps to occur but not so long as to hinder product release. Enzymes under different temperature selection may lower or raise free-energy barriers, and their respective entropic/enthalpic contributions, associated with particular motions to maintain proper flux. To begin to gain some insight into how PEPCKs balance conformational states and the associated free-energy barriers, three thermally adapted isozymes of GTP-dependent PEPCK were studied. These enzymes encompass a broad range of temperatures and were selected based upon changes to known important residues in Ω-loop dynamics. Specifically the Ω-loop was examined as it is unlike the other functionally important loops that have been characterized, as it transitions conformation states without interacting with substrates/products. Thus, it was hypothesized that its sequence would be less constrained by the nature of the chemical reaction and more sensitive to changes necessary for function under different temperatures. Based on this criteria, enzymes from three sources were selected: *Polaromonas naphthalenivorans* a psychrophile (T_{opt}, 20°C)[100], *Rattus norvegicus* the mesophile (T_{opt}, 37°C) and the standard structural and functional model of PEPCK in our laboratory, and finally, *Hungateiclostridium thermocellum* a thermophile (T_{opt}, 60°C).[101]

3.2 Methods

3.2.1 Construct Design

The GTP-PEPCK (EC: 4.1.1.32) protein sequence from *Rattus norvegicus* was described in 2.2.1. *Polaromonas naphthalenivorans (*referred to as polarPEPCK)*(*WP_011799537.1) and *Hungateiclostridium thermocellum* (referred to as clostPEPCK) (*Clostridium thermocellum*)(A3DJE3) were codon optimized for *Escherichia coli* and synthesized by GenScript (See Section 7.2). GenScript cloned the separate genes into the pESUMO-Star (Kan) vector (LifeSensors) between the BSA1 and XHO1 sites so that the final expressed protein had the SUMO 6xHIS fusion tag at the N-terminus. Upon cleavage of the fused protein, the resultant protein would have no additional amino acids.

3.2.2 Expression

The expression of all enzymes was carried out in a fashion identical that described in 2.2.2.

3.2.3 Purification

All isozymes of PEPCK were purified in the same manner, with the exception that clostPEPCK was purified at room temperature, rather than 4°C. The purification protocol was identical to the procedure described in 2.2.3.

3.2.4 Kinetic Characterizations

To determine each kinetic constant, all other substrates were held constant and the assays were completed in duplicate. For the Arrhenius plots, saturating concentrations of all substrates were used. The addition of glycerol to increase the viscosity of the reaction was varied dependent on temperature to maintain 2.4 centipoise at each temperature.[102] All assays were completed in pre-equilibrated kinetic assay mixes at a final volume of 1mL with the use of a temperature controller. The temperatures were verified with a thermocouple for accuracy. A CaryUV100 spectrophotometer was used to monitor the reduction of NADH to NAD^+ in the coupled assays at 340nm.

As previously shown, OAA can spontaneous decarboxylate to pyruvate in the presence of metals.[103] To measure this background rate the standard assays were used without the addition the PEPCK enzyme. All rates have been normalized for this background rate depending on the metal mix used.

OAA + GTP → PEP + GDP + CO_2

The standard assay mix was composed of 100mM HEPES pH 7.5, 10mM DTT, 300μM NADH, 0.5mM GTP, 2mM $MgCl_2$ (4:1 nucleotide:metal ratio), 100μM $MnCl_2$, 500μM OAA, 1mM ADP, 50μg pyruvate kinase (PK), 5U lactate dehydrogenase (LDH), and 2.5μg of GTP PEPCK. Reactions were started with the addition of OAA.

OAA + GTP → Pyruvate

The standard assay mix described above in the OAA→PEP section was used with the modified to not include the coupling enzyme pyruvate kinase in order to measure the enzyme-catalytzed decarboxylation of OAA.

PEP +GDP + CO_2 → OAA + GTP

The standard assay mix was composed of 100mM HEPES pH 7.5, 10mM DTT, 300μM NADH, 1mM GDP , 4mM $MgCl_2$ (4:1 nucleotide:metal ratio), 100μM $MnCl_2$, 10mM PEP, 50mM $KHCO_3^-$ (bubbled with dry ice), 10U of MDH and 2.5μg of GTP PEPCK. Reactions were started upon the addition of GTP PEPCK.

3.2.5 Kinetic Data Analysis and Equations

Michaelis-Menten and Arrhenius plots were derived from plots created in SigmaPlot11 EnzymeKinetics module using equation 3.1 and equation 3.2 respectively. Inflection temperatures of each Arrhenius plot for the biphasic changes in E_a were determined based on visual inspection and the highest R^2 value for the fitted line.

$$v=\frac{V_{max,app}[S]}{K_{m,app}+[S]} \tag{3.1}$$

$$k=A\,e^{\frac{-Ea}{RT}} \tag{3.2}$$

3.2.6 Crystallization

PolarPEPCK cocrystals with GTP and oxalate were grown at 20mg/mL through vapour diffusion at 4°C in drops supplemented with 2mM oxalate and 2mM GTP. The mother liquor condition was 100mM TRIS-Cl at pH 8.0, 0.8M $LiCl_2$ and 30-40% PEG 8000. Drops were set up in 2:4, 3:3, 4:2 μL ratios of protein:mother liquor. Crystals were harvested, and soaked for 1 hour in 100mM TRIS-Cl at pH 8.0, 0.8M $LiCl_2$ and 38% PEG 8000, 10% PEG 400, and 1mM of both oxalate and GTP. Crystals were cryocooled by direct immersion into liquid nitrogen.

Data was collected at the Canadian Light Source in Saskatoon, Saskatchewan on the 08B1-1 beamline using a Rayonix MX300HE detector. Data was indexed and scaled in HKL-2000[79] and the final statistics data are presented in Table 3.1.

Three additional structures were collected under the same conditions but with 100μM $MnCl_2$ either co-crystallized or soaked during the cryoprotection. In the first structure $MnCl_2$ (1mM) was soaked into the crystal, forming Mn^{2+}(M1)-GDP-oxalate polarPEPECK. The second structure was co-crystallized with $MnCl_{2,}$ (100μM) forming Mn^{2+}(M1)-GTP-oxalate polarPEPCK. The third and final structure was a co-crystal with Mn^{2+} (1mM) and Mg^{2+} (5mM) forming Mn^{2+} (M1) – Mg^{2+} (M2) – GTP-oxalate polarPEPCK. This data was collected on the University of Waterloo's home source rotating copper anode diffractometer with a Raxis IV++ image plate detector. Data were indexed and scaled in HKL-2000[79] and the final statistics are presented in Table 3.1.

3.2.7 Structure Determination

The highest resolution data set collected from the synchrotron was solved by merging two datasets of the same crystal, with the first dataset being a short exposure low-resolution pass and the second a longer exposed higher-resolution pass. This first structure, GDP-oxalate polarPEPCK, was solved through molecular replacement using the previously solved mtbPEPCK (PDB# 4R43[33]) and MOLREP[80] in the CCP4 package.[81] Iterative rounds real-space refinements and additions of water molecules and heteroatoms were carried out in COOT[82], followed by refinement using REFMAC5.[83] Final validation was carried out using MolProbity, http://molprobity.biochem.duke.edu/index.php, and final model statistics are presented in Table 3.1.[84]

The additional metal-bound structures were solved solved in an identical fashion to that outlined above expected the initial GDP-oxalate polarPEPCK structure was used as the starting model for the molecular replacement in MOLREP.[80,81] Final model statistics are present in Table 3.1.

Table 3.1: Crystallographic data table for the three polarPEPCK structures.

	GDP-Oxalate polarPEPCK	**Mn^{2+}-GDP-Oxalate polarPEPCK**
Wavelength (Å)	1	1.54
Resolution range (Å)	56.6 - 1.41 (1.46 - 1.41)	32.7 - 2.14 (2.22 - 2.14)
Space group	P 21 21 21	P 21 21 21
Unit cell (Å)	68.121 85.988 113.232 90 90 90	68.318 85.914 114.281 90 90 90
Total reflections	8083610	1465173
Unique reflections	127745 (12020)	37233 (3369)
Multiplicity	12.7 (5.0)	11.1 (8.5)
Completeness (%)	99.5 (94.7)	99.1 (91.4)
Mean I/sigma(I)	18.6 (1.58)	11.6 (1.9)
Wilson B-factor (Å^2)	17.5	34.4
R-merge	0.14 (0.98)	0.20 (1.07)
R-meas	0.14 (1.08)	0.21 (1.14)
R-pim	0.03 (0.46)	0.06 (0.38)
CC1/2	0.91 (0.56)	0.91 (0.63)
Reflections used in refinement	127741 (12020)	37233 (3369)
Reflections used for R-free	6395 (612)	1842 (158)
R-work	0.159 (0.266)	0.186 (0.267)
R-free	0.178 (0.271)	0.245 (0.304)
Number of non-hydrogen atom	5522	4953
macromolecules	4920	4759
ligands	63	48
solvent	539	146
Protein residues	614	614
RMS(bonds)	0.015	0.014
RMS(angles)	1.88	1.81
Ramachandran favored (%)	97.9	97.2
Ramachandran allowed (%)	2.12	2.45
Ramachandran outliers (%)	0	0.33
Rotamer outliers (%)	0.39	1.21
Clashscore	2.75	2.54
Average B-factor	20.7	34.8
macromolecules	19.7	34.9
ligands	27.9	35
solvent	29.1	32.2

	Mn^{2+}-GTP-Oxalate polarPEPCK	**Mn^{2+}-Mg^{2+}-GTP-Oxalate polarPEPCK**
Wavelength (Å)	1.54	1.54
Resolution range (Å)	30.9 - 1.79 (1.85 - 1.79)	30.52 - 2.001 (2.072 - 2.001)
Space group	P 21 21 21	P 21 21 21
Unit cell (Å)	68.383 85.742 113.686 90 90 90	68.489 85.557 112.857 90 90 90
Total reflections	1707490	1665108
Unique reflections	60604 (4360)	45085 (4145)
Multiplicity	11.5 (5.7)	13.1 (6.7)
Completeness (%)	95.2 (69.4)	99.1 (91.9)
Mean I/sigma(I)	24.1 (2.5)	21.2 (2.2)
Wilson B-factor (Å^2)	26	31.63
R-merge	0.10 (0.61)	0.12 (0.64)
R-meas	0.10 (0.68)	0.12 (0.69)
R-pim	0.03 (0.27)	0.03 (0.24)
CC1/2	0.96 (0.80)	0.96 (0.82)
Reflections used in refinement	60603 (4360)	45085 (4145)
Reflections used for R-free	3037 (227)	2277 (219)
R-work	0.168 (0.292)	0.1740 (0.2648)
R-free	0.198 (0.335)	0.2120 (0.3076)
Number of non-hydrogen atom	5179	5167
macromolecules	4794	4746
ligands	62	66
solvent	323	355
Protein residues	614	614
RMS(bonds)	0.015	0.015
RMS(angles)	1.87	1.9
Ramachandran favored (%)	98.2	97.4
Ramachandran allowed (%)	1.8	2.6
Ramachandran outliers (%)	0	0
Rotamer outliers (%)	1	2
Clashscore	1.68	4.8
Average B-factor	26.5	32.8
macromolecules	26.2	32.2
ligands	32.1	36.6
solvent	30.2	40.1

3.3 Results

3.3.1 Enzyme inactivation

As it is well established that thermophiles are more resistant to chemical denaturants and heat, inactivation time-courses were completed for the three isozymes (Figure 3.1).[104] For the polar- and rcPEPCK, they were completely inactive when incubated for 10 minutes at 55°C. ClostPEPCK however was most resistant at this temperature and retained partial activity until 50 minutes. Surprisingly, clostPEPCK was most sensitive to moderate (37°C) and cold temperatures (4°C). The observation of cold denaturation was not surprising but the moderate temperatures was not expected. *Hungateiclostridium thermocellum* besides being a thermophile, is also anaerobic, and the inactivation may be due to an oxidation event. PolarPEPCK is more resistant to inactivation than its mesophilic counterpart at modest temperatures, which is interesting but very challenging to determine a causal mechanism. Fortunately, the enzymes are stable and were stored at 4°C during the other kinetic experiments making the collections accurate and reproducible.

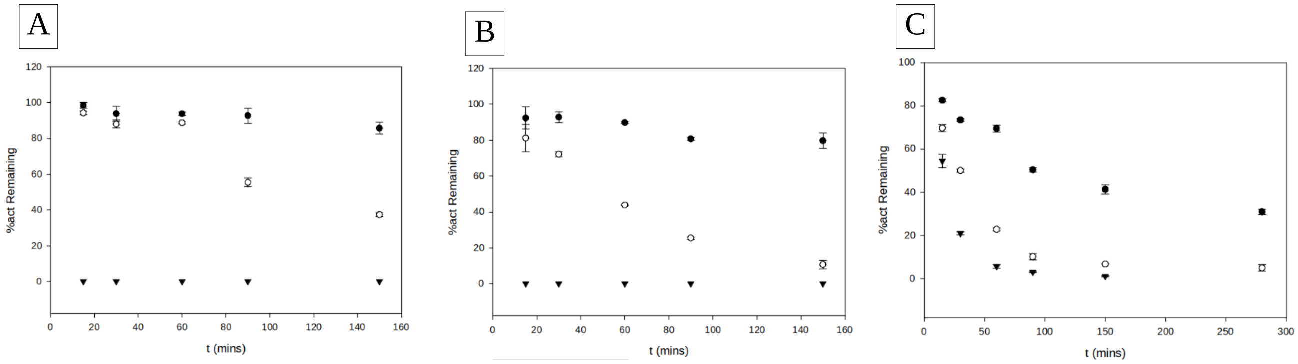


Figure 3.1: Inactivation of thermally adapted PEPCK. This figure shows a time course for A) polarPEPCK, B)rcPEPCK, and C)clostPEPCK inactivation when incubated at 55°C (closed triangles), 37°C (open circles), and 4°C (closed circles).

3.3.2 Arrhenius plots - Overview

To further understand the functional properties of the three thermal isozymes, Arrhenius plots were completed for each the PEP carboxylation and OAA decarboxylation reaction directions (Figure 3.2, Table 3.2). Unfortunately, a relatively narrow temperature window (~5-40°C) was obtained for the OAA carboxylation reaction as at high temperature the spontaneous decarboxylation of OAA to pyruvate began to outcompete the enzymatic rates for conversion of OAA to PEP. Viscogens were also added to the reaction mixtures at constant centipoise values to determine if the k_{cat} measured was being controlled by diffusive events. RcPEPCK, in the PEP carboxylating reaction direction, with viscogen at high temperatures has a very low R^2 value (0.28) due to a low number of data points used but ultimately the fit best described the data. PolarPEPCK after 40°C began to become less-active, but was not observed to undergo complete inactivation (Figure 3.2). During the inactivation time-course, polarPEPCK was stable at 37°C and during collection at these temperatures above 40°C, the rate is stable. Because of the thermal degradation of the coupling enzyme malate dehydrogenase, the suspected temperatures necessary to inactive the meso- and thermophilic enzymes was unable to be experimentally achieved.

3.3.3 Arrhenius plots - PEP carboxylation reaction at low temperature

Arrhenius plots can be used to understand the activation energy for the rate-determining step controlling k_{cat} (that is, after or including the committed step). Typically, Arrhenius plots are described to be linear, but non-conforming trends have also been seen.[105] As shown in the PEP carboxylation reaction, polarPEPCK exhibits typical behaviour, with no slope changes across until a maximum threshold was reached, whereas the mesophilic and thermophilic enzymes both have slight downward curvature (Figure 3.2). The observed curvature in Arrhenius plots have been hypothesized to have different mechanistic origins,[105] but most generally this curvature can be interpreted as either a change in the activation energy or the rate limiting step. More specifically, a previous event which was rate limiting at low temperatures has been replaced by a different event as temperature increases.

Based upon the slopes of the plots, the activation energies are 83, 100, 130kJ/mol for the polar-, rc, and clostPEPCKs respectively. Based upon these numbers a trend is observed in which increasing thermophilicity is correlated with an increase in the observed energy required to achieve the transition state. When comparing the derived pre-exponential factors across the low temperature regime of the three tested enzymes, the same trend of increasing A with thermophilicity is also observed. Accordingly, when comparing the low temperature E_a or pre-exponential factor for the PEP carboxylation for all enzymes, viscogen has no effect (Table 3.2).

3.3.4 Arrhenius plots - PEP carboxylation reaction at high temperature

For each of the three enzymes, there is a change in the slope of their Arrhenius plots after a particular break-point (Figure 3.2). For the polarPEPCK, the enzyme transitions to becoming less active, whereas in the meso- and thermophilic PEPCKs both E_a and A decrease. Although there are not many data points for the inactivation stage of polarPEPCK, it does appear that the addition of viscogen to polarPEPCKs influences only A and not E_a in contrast to the higher temperature variants, viscogen seems to lower both the E_a and A.

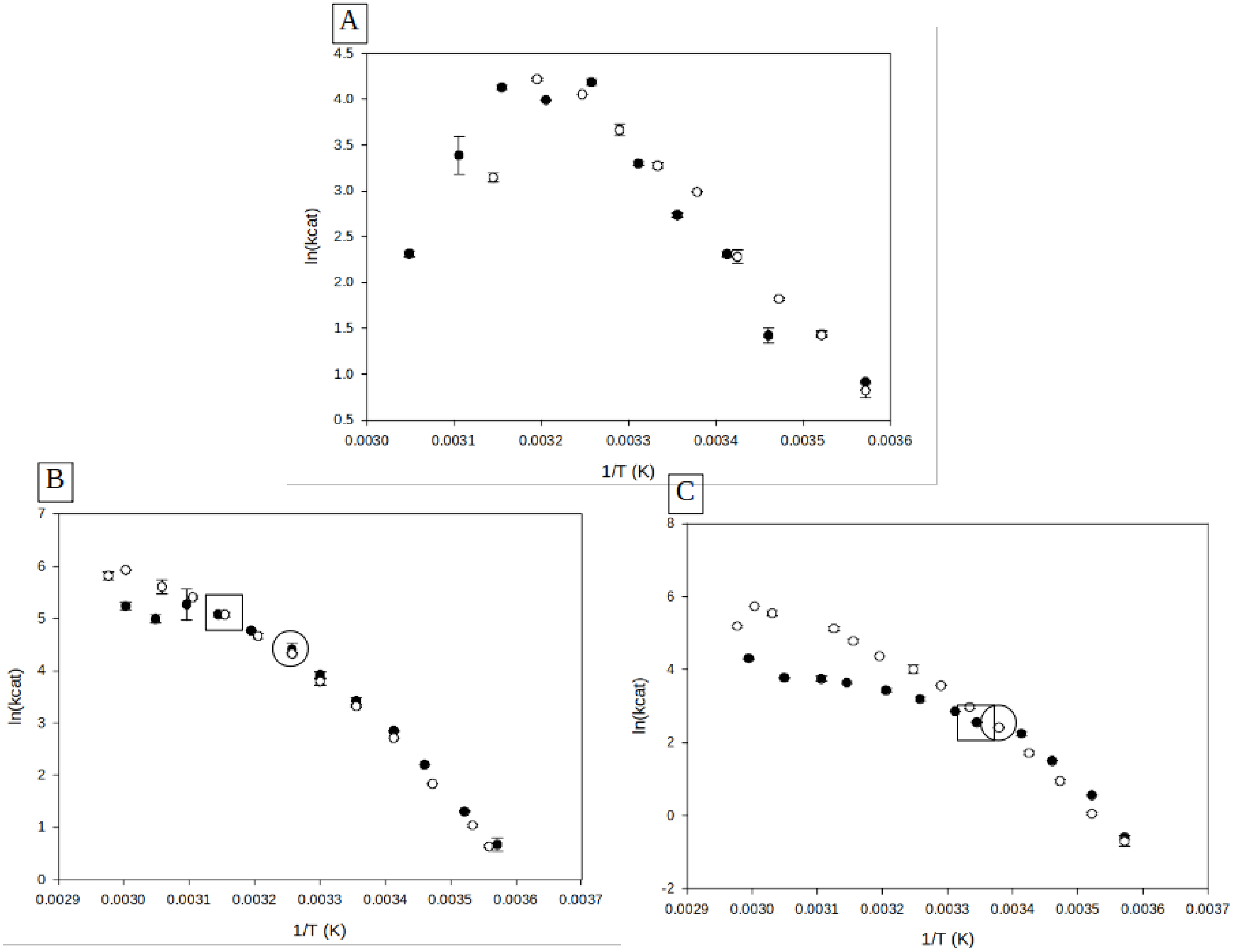


Figure 3.2: Arrhenius plots for the PEP decarboxylation of thermal isozymes of PEPCK. This figure shows the Arrhenius plots for A) polarPEPCK, B)rcPEPCK, and C)clostPEPCK for both the reaction without viscogen (open circles) and with viscogen (closed circles). In panels B and C, the circle and square represents the inflection point for the two measured slopes in table 3.2 for the plot without viscogen and with viscogen respectively.

3.3.5 Arrhenius plots - OAA decarboxylation reaction

Over the temperature range measured, the decarboxylation rates for the three enzymes were linear (Figure 3.3). The E_a for all three enzymes are lower than the PEP carboxylation reaction, suggesting that the OAA decarboxylation reaction is more favourable with than the reverse reaction, and may be measuring a different process. This is in line with the enzymes *in vivo* preferences that have been described.[43] Through inspection of the raw data, viscogen addition has no change on either rc- or clostPEPCK (Figure 3.3). However the fit data shows that upon calculating the Arrhenius parameters there are differences (Table 3.2) however it is likely that these differences are attributed to errors within the associated measurements rather than mechanistic differences. PolarPEPCK on the other hand does appear to maintain the E_a but raise A with addition of viscogen.

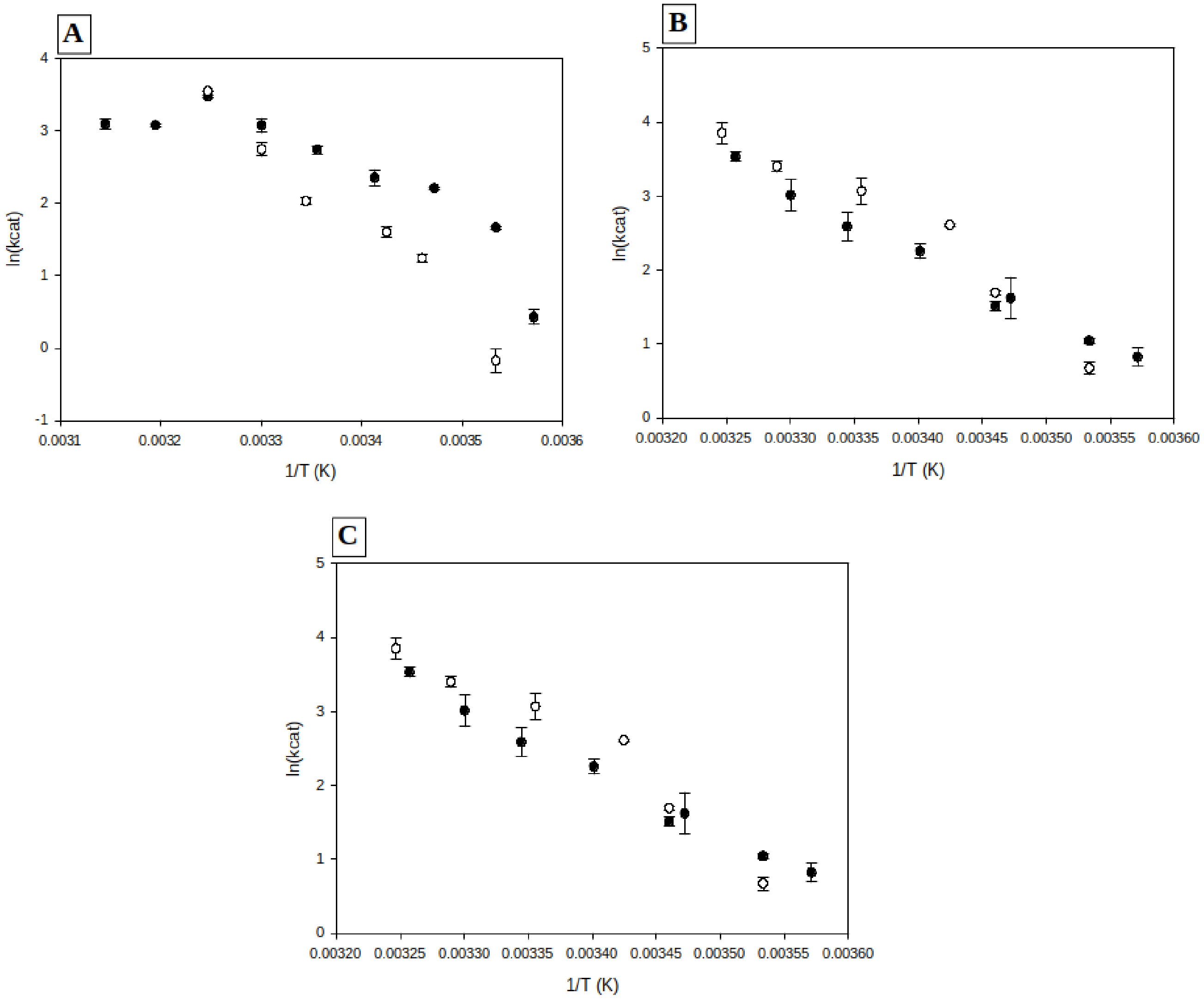


Figure 3.3: Arrhenius plots for the OAA decarboxylation of three thermal isozymes of PEPCK. This figure shows the Arrhenius plots for A) polarPEPCK, B)rcPEPCK, and C)clostPEPCK for both the reaction without viscogen (closed circles) and with viscogen (open circles).

Table 3.2: Derived variables from the Arrhenius plots for thermal isozymes of PEPCK.

PEP + CO_2 + GDP → OAA + GTP	**Slope**	**E_a (kJ/Mol)**	**A (s^{-1})**	**R^2**
polarPEPCK	-10 000 ± 840	83 ± 7	37 ± 2.8	0.99
polarPEPCK + Viscogen	-9400 ± 420	78 ± 3.5	34 ± 1.4	0.95
rcPEPCK (low temperature)	-12 000 ± 440	100 ± 3.7	44 ± 1.5	0.99
rcPEPCK (high temperature)	-5700 ± 500	47 ± 4.2	23 ± 1.5	0.95
rcPEPCK + Viscogen (low temperature)	-11 000 ± 570	92 ± 4.7	40 ± 1.9	0.98
rcPEPCK + Viscogen (low temperature)	-1789 ± 1477	15 ± 12	11 ± 4.5	0.28
clostPEPCK (low temperature)	-16 000 ± 410	130 ± 3.4	56 ± 1.4	0.99
clostPEPCK (high temperature)	-9800 ± 440	82 ± 3.7	36 ± 1.4	0.99
clostPEPCK + Viscogen (low temperature)	-18 000 ± 1200	150 ± 10	63 ± 4.2	0.99
clostPEPCK + Viscogen (low temperature)	-5100 ± 350	42 ± 3	20 ± 1.1	0.97

OAA + GTP → PEP + CO_2 + GDP				
polarPEPCK	-7100 ± 510	59 ± 4.2	26 ± 1.7	0.96
polarPEPCK + Viscogen	-11 000 ± 1000	91 ± 8.3	39 ± 3.5	0.95
rcPEPCK	-8600 ± 360	71 ± 3	31 ± 1.2	0.96
rcPEPCK + Viscogen	-11 000 ± 1300	88 ± 11	38 ± 4.5	0.93
clostPEPCK	-8000 ± 1200	66 ± 10	29 ± 4.2	0.89
clostPEPCK + Viscogen	-12000 ± 1200	98 ± 10	42 ± 3.9	0.96

3.3.6 Michaelis-Menten Parameters of Thermal Isozymes

The Arrhenius plots were generated using saturating substrate concentrations, thus removing information about the any event prior to the first irreversible step. The k_{cat}/K_m ratio, otherwise known as the catalytic efficiency, can give information about the steps leading up to and including the irreversible step. Below the break-point the k_{cat}/K_m ratio is the same for each rc- and clostPEPCK at the two temperatures obtained (Table 3.3). However, above this breakpoint, the k_{cat}/K_m ratio increases, suggesting that the steps leading up to the irreversible step are becoming more favourable. In contrast, in polarPEPCK at elevated temperatures in which is observed to reversibly inactivating the k_{cat}/K_m decreases (Table 3.3).

Table 3.3: Michaelis-Menten parameters for thermal isozymes of PEPCK at different temperatures.

PEP + CO_2 + GDP → OAA + GTP

Enzyme	K_m (µM) PEP	k_{cat} (s^{-1}) PEP	k_{cat}/K_m ($M^{-1}s^{-1}$) PEP
15 (°C)			
polarPEPCK	810 ± 87	5.8 ± 0.19	7.2×10^3
rcPEPCK	120 ± 8.7	6.0 ± 0.01	5.0×10^4
clostPEPCK	130 ± 14	1.9 ± 0.01	1.5×10^4
25 (°C)			
polarPEPCK	2500 ± 650	23 ± 1.4	9.2×10^3
rcPEPCK	200 ± 23	13 ± 0.34	6.5×10^4
clostPEPCK	540 ± 60	9.1 +/- 0.3	1.7×10^4
37 (°C)			
polarPEPCK	3400 ± 290	92 ± 3.2	2.7×10^4
rcPEPCK	330 ± 34	71 ± 1.7	2.2×10^5
clostPEPCK	510 ± 37	46 ± 1.1	9.0×10^4
45 (°C)			
polarPEPCK	5900 ± 1000	47 ± 3.0	8.0×10^3
55 (°C)			
rcPEPCK	830 ± 110	230 ± 7.1	2.8×10^5
clostPEPCK	870 ± 70	200 ± 4.3	2.3×10^5

3.3.7 Substrate Binding of *Polaromonas naphthalenivorans* PEPCK

To further understand the molecular details of the catalytic mechanism, structural studies were completed on the psychro- and thermophilic enzymes. ClostPEPCK was not amenable to crystallization under conditions that were explored. PolarPEPCK however, was successfully crystallized and four structures were solved of the closed state with oxalate, GDP/GTP, and/or the M1 and M2 metals. The series of structures was stimulated as it was observed in the first structure solved of polarPEPCK, that there was an unexpected binding conformation of oxalate, and GDP was present instead of the co-crystallized GTP (Figure 3.4A). As shown, oxalate is found in an staggered conformation, no metals (M1 nor M2) are present, and GDP is bound. In the second structure determined, Mn^{2+} was soaked in to the identically prepared crystals after crystallization (Figure 3.4B). This metal could diffuse into the closed active site, and did bind the M1 metal stabilizing the coordinated form of oxalate as seen previously.[62] In these crystals, GDP was found to be the nucleotide present in the active site. In an attempt to obtain the final Mn^{2+}-Mn^{2+}-oxalate-GTP complex, low concentrations of Mn^{2+} were co-crystallized and soaked. This co-crystal structure did have Mn^{2+} in the M1 position, accompanied with oxalate and now GTP. Surprisingly, the M2 metal is still not observed to be bound in concert with the GTP nucleotide. A final structure where high concentrations of Mn^{2+} and Mg^{2+} were cocrystallized alongside GTP and oxalate yielded the expected fully occupied active site as previously shown.[62]

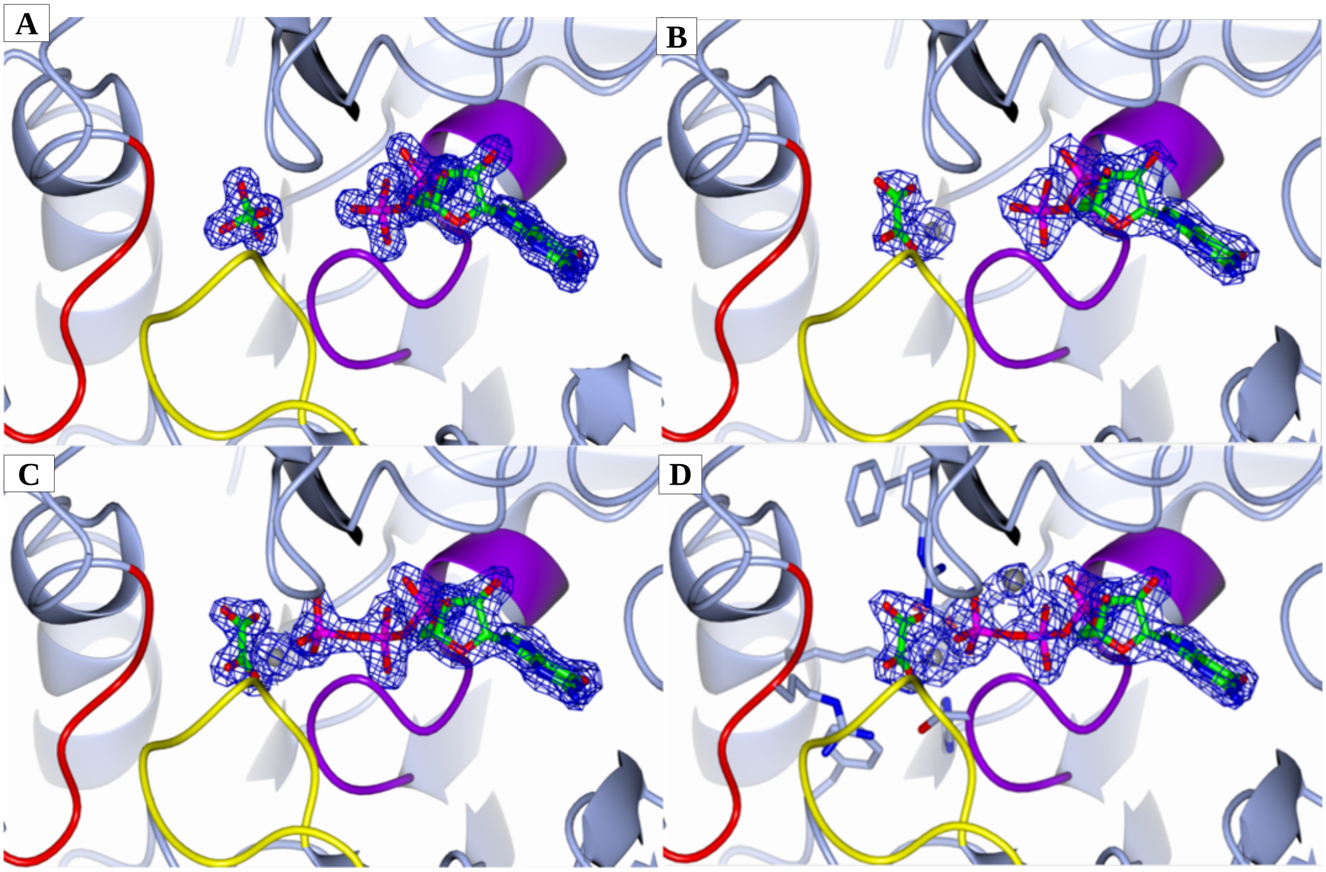


Figure 3.4: Ligand states for polarPEPCK structures. This figure shows the various ligands found in the structures solved of polarPEPCK. Panel A represents the structure without Mn^{2+} cocrystallized or soaked. This structure had oxalate in a staggered conformation with GDP. Panel B represents the structure where Mn^{2+} is soaked into the crystal. This structure has Mn^{2+} at the M1 site with oxalate complexing with the metal, and GDP. Panel C represents the structure where Mn^{2+} was cocrystallized. This structure has Mn^{2+} at the M1 site, oxalate coordinating with the metal, and GTP. Panel D represents the structure where both Mg^{2+} and Mn^{2+} were cocrystallized with GTP and oxalate. This structure shows the fully occupied active site with Mn^{2+} at the M1 site, Mg^{2+} at the M2 site, and oxalate and GTP bound to these metals. The experimental $2F_o$-F_c electron density maps (rendered at 1.7σ for ligands, 3.5σ for metals) are shown in blue mesh. The density is cropped for clarity. The R-, P-, and Ω-loop are shown in red, purple and yellow respectively. The various ligands are coloured by atom type, with carbon as green, oxygen as red, nitrogen as blue, phosphorous as purple, and Mn^{2+} as grey.

3.3.8 Ω-Loop Closure - Latch Stabilization

The goal of crystallizing the closed conformation of polarPEPCK was to determine how a psychrophilic PEPCK has adapted to change the mechanism of closure to its low-temperature environment. To examine this process, two areas of the lid were focused upon based upon prior work on rcPEPCK (Figure 3.5). First, the latch, accounting for the loop itself which is shown to be important in maintaining a closed state (Figure 3.5A). The second region important for the operation of the lid is the hinge region, which has been shown to undergo conformational changes to fold the lid over the active site (Figure 3.5B).

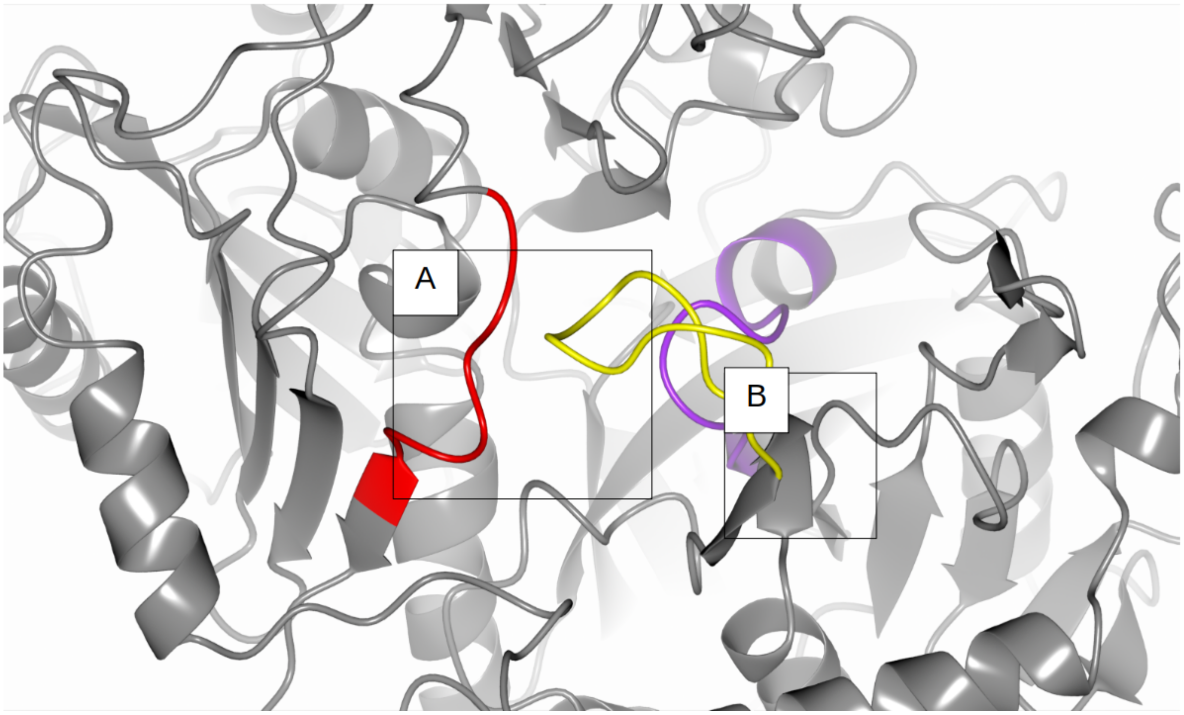


Figure 3.5: Two Ω-loop closure sites. The overall structure is rcPEPCK bound with Mn^{2+}(M1)-Mn^{2+}(M2)-oxalate-GTP (PDB ID# 3DT2). Panel A represents the area of oxalate binding, and the latch region of the Ω-loop that is responsible for closure. Panel B represents the hinge region of the Ω-loop which is necessary for closure. The R-, P-, and Ω-loop are coloured in red, purple and yellow respectively.

All four polarPEPCK structures were essentially identical besides with the expection of the ligands found in the active site. It is unsurprisingly that even in cases were the metals are absent and oxalate is in the staggered conformation the enzyme is closed. For the lid to close, both the R- and P-loop must be ordered (upon ligand association) to remove steric hindrances of the lid. As seen from these observed structures, the metal binding appears to no influence on these active site loop transitions as long as the substrate requirements are filled. Because of this, the two structures used for this comparison between the psychro- and mesophilic PEPCKs are the cocrystallized polarPEPCK structure, with M1-oxalate-GTP bound, and rcPECK M1-M2-oxalate-GTP bound (PDB ID# 3DT2) (Figure 3.6). Both the residues responsible for substrate (oxalate) recognition (Figure 3.5A and B) and the stabilizing forces required for latch closure are examined (Figure 3.6C and D). Unsurprisingly, it is observed that the substrate recognition is the same between the two isozymes. Upon binding of oxalate and GTP, both the R- and P-loop move towards in active site. Nε of R87 guanadine of the R-loop forms a salt-bridge with the carboxylic acid of oxalate. The P-loop upon rearrangement, brings the hydroxyl of S286 into proximity to form a H-bond with this carboxylate moiety. Together these movements allow the R-loop to effectively pair with the Ω-loop upon closure.

There appears to be more extensive contact between the latch and the adjacent R-loop in the mesophilic enzyme (Figure 3.5C and D). First, H470 has forms a salt-bridge with E89 as E469 exhibits two hydrogen bonds with the backbone amide of E89 and the hydroxyl of S90. Furthermore, two backbone H bonds, between the R87 carbonyl and the amide of E468, and between the A467 carbonyl and the amide of R87s are the final stabilizing forces. In total, this amounts to 5 interactions stabilizing the latch of the lid closed. In comparison, the psychrophilic lid has 3 stabilizing contributions. The first is at the latch of lid between the backbone T466 amide and the carbonyl of R84. The second, is actually two hydrogen bonds shared between R473 from the hinge region of the lid coordinating with the both oxygens of the carboxylate of E86.

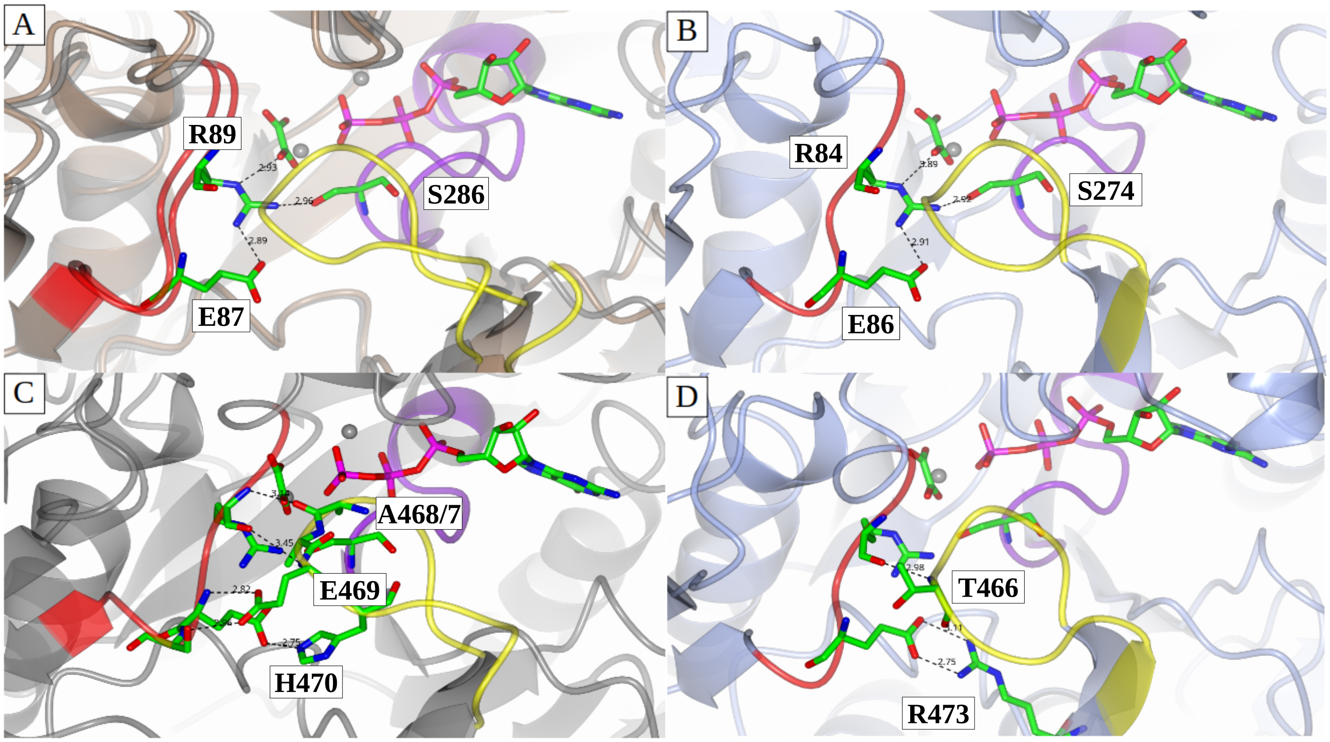


Figure 3.6: Latch region of Ω-loop in closed rc- and polarPEPCK. Panel A shows holo-rcPEPCK (PDB#2QEW) in brown and closed-rcPEPCK (PDB#3DT2) in grey while showing the interactions responsible for substrate (oxalate) recognition. Panel B shows the closed-polarPEPCK (Mn^{2+}-oxalate-GTP) in ice blue showing the interactions responsible for substrate (oxalate) recognition. Panel C shows the interactions necessary for maintaining lid closure in the closed-rcPEPCK state. Panel D shows the interactions necessary for maintaining lid closure in the closed-polarPEPCK state. The R-, P-, and Ω-loop are coloured in red, purple and yellow respectively. Atoms are coloured by type with carbon in green, oxygen in red, nitrogen in blue, phosphorous in purple, manganese in grey.

3.3.9 Ω-Loop Closure - Asymmetric Hinge Closure

The hinge region was previously identified as being an important dynamic element in the closure of the Ω-loop (Figure 3.7).[37] A comparison between rcPEPCK open and closed state transition (Figure 3.7A and B) and the psychrophilic closed state (Figure 3.7C) is shown. Unfortunately, this crystal forms does not allow for structural determination of holo polarPEPCK, so only one end point is known. When comparing the rcPEPCK open and closed state, it can be seen that the N-terminal β sheet hinge region is unchanging. The N-terminal hinge is fixed by one H-bonds between R461 and the adjacent E588 with an additional stabilizing force from the non-hinge R437 to hold E588 in place. The C-terminal hinge is the variable region, where H477 electrostatically interacts with E591. E591 rotates to an “in” position, closing this interaction distance. In the closed polarPEPCK structure there are no interactions from the N-terminal region, although the non hinge hydrogen bond between the conserved R434 and E590 is retained. The C-terminal hinge region has the only contact between the hinge and surrounding structure, between R474 and the aforementioned E590.

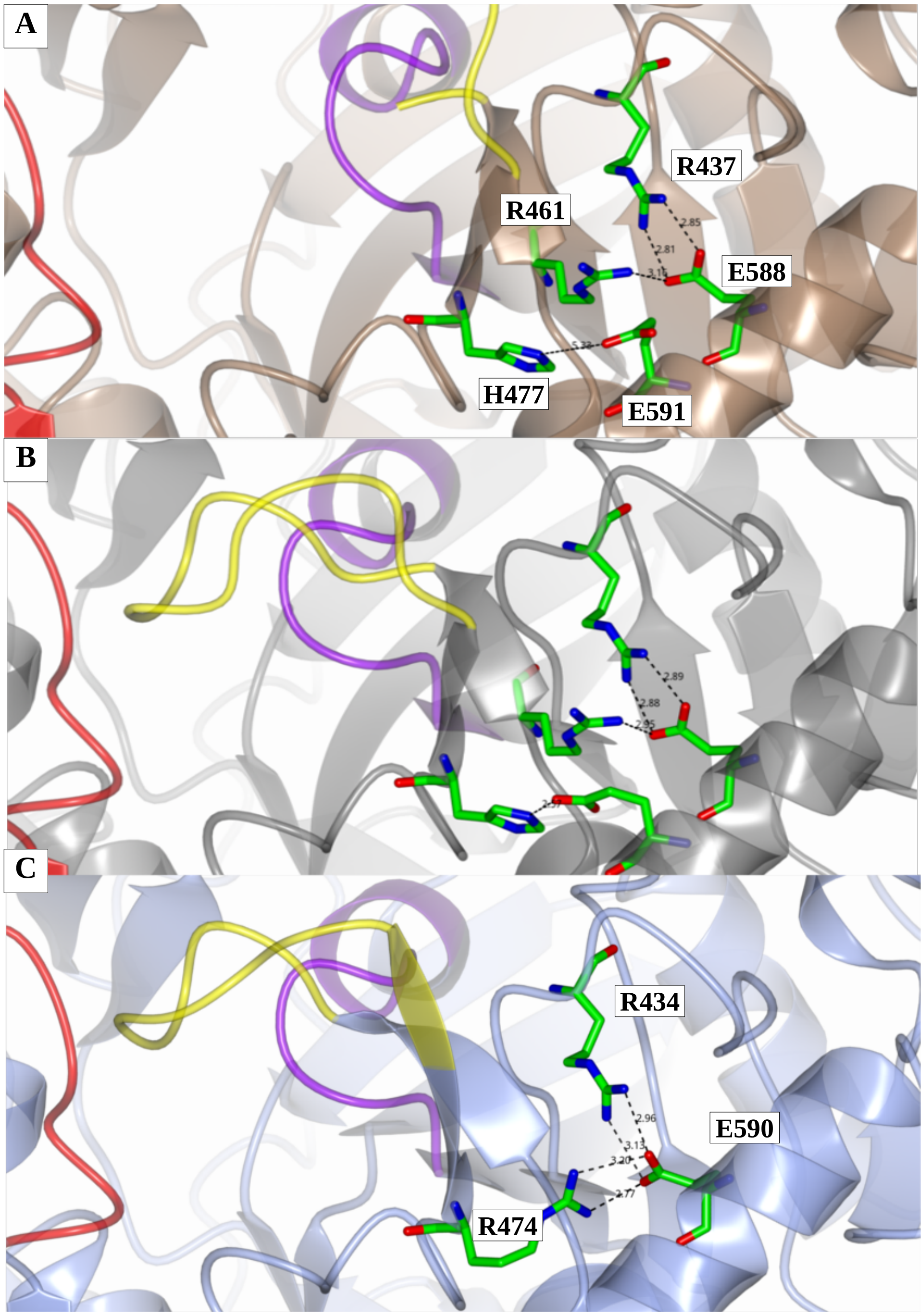


Figure 3.7: Hinge region of Ω-loop in closed rc- and polarPEPCK. Panel A shows holo-rcPEPCK (PDB ID# 2QEW) hinge region unanchored. Panel B shows the closed-rcPEPCK (PDB ID# 3DT2) in grey showing the interactions responsible for hinge anchoring necessary for closure. Panel C shows the closed polarPEPCK interactions responsible for hinge anchoring necessary for closure. The R-, P-, and Ω-loop are coloured in red, purple and yellow respectively. Atoms are coloured by type with carbon in green, oxygen in red, nitrogen in blue, phosphorous in purple, manganese in grey.

3.4 Discussion

Before interpreting the results, a generalized discussion on the mechanistic meaning of the Michaelis-Menten, Arrhenius parameters, and viscogen effects may be beneficial. These parameters will be discussed in the context of the catalytic cycle of PEPCK. PEPCK catalysis can be summarized into a few elementary steps. First, the lid must be open to allow substrate to bind. Second, the lid must close. Third, the first chemical step (OAA decarboxylation) and the formation of the enolate, representing the first irreversible step, must occur. Together it is postulated that these three steps contribute to the observed $k_{cat}/K_{m.}$ Fourth, the second chemical step occurs (enolate phosphorylation). Finally, the lid must reopen for product release. Steps 3 and onwards, would contribute the $k_{cat,}$ but the observed rate constant would be dominated by the slowest, or rate determining step. Arrhenius parameters interpretations are not as clear cut, as they were formulated for non-enzymatic chemical reactions. $E_{a,}$ the activation energy is approximately the enthalpy of activation ($\Delta H^{\ddagger}$) while the second and final kinetic parameter discussed with the pre-exponential factor, A, can be thought to as a proxy for activation entropy ($\Delta S^{\ddagger}$). This value is an extrapolation of the systems behaviour at infinite T, thus acting as a barrier(E_a)-less reaction or that the energetic cost is so much smaller than the energy available. This removes the contribution of the enthalpy of activation from the total free-energy cost, leaving only entropic barriers. Viscogens effectively increase the solvent friction, which will impact any diffusive properties through solvent. These diffusive properties include conformational dynamics with α-fluctuations (motions coupled to bulk solvent) being more affected than β-fluctuations (hydration shell). α-fluctuations will control large conformational changes such a domain motions and surface loop motions, while β-fluctuations control internal dynamics such as rotamer motions.[106]

3.4.1 Mechanistic Interpretation of Kinetic Adaptations of PEPCKs

In the PEP carboxylation reaction at low temperature, increasing thermophilicity leads to an increase in E_a ($\Delta H^{\ddagger}$) which is in agreement with the literature (reviewed in [99]). Thermolabile enzymes achieve higher turnover rates at lower temperatures by lowering the enthalpy of activation concomitantly with their apparent increase in flexibility. The number of interactions that must be broken and formed, and the resultant free-energy difference, is lower for a more flexible enzyme as there are less interactions stabilizing its conformation than a thermostable, rigid enzyme. To maintain structural integrity at high temperatures more intermolecular interactions must hold the tertiary structure of the enzyme. This rate constant is also unaffected by viscogen, suggesting that conformational changes are not the primary driving force in determination of the rate. Previous work has suggested that for rcPEPCK, the rate determining step is the phosphoryl transfer in both chemical reaction directions, which would be less effected by increase in solvent friction.[38] At high temperatures, the slope has changed. For polarPEPCK, the enzyme begins to lose activity with increasing temperature. The shape of the Arrhenius plot is reminiscent of the aforementioned equilibrium model, as the enzyme is reversibly inactivated. During the time-course of the experiments at these high temperatures, the rates were not fluctuating suggesting that the enzyme was not becoming thermally denatured with time. Due to experimental constraints with the mesophilic and thermophilic variants, higher temperatures were not achieved to determine if they under a similar inactivation. Furthermore, with the few data points present beyond the maximal activation for polarPEPCK, it does appear that viscogen presence does further reduce activity, through the pre-exponential factor. This suggests that the inactivation of the enzyme is through disrupting entropy driven conformational changes associated with turnover, such as the balancing between closure of the Ω-loop with turnover. The rc- and clostPEPCK suffer similar reaction changes with increasing temperature, where after the inflection temperature the slopes decreases leading to both a lowering of E_a and A. The new friction-dependency suggests that in fact the rate determining step has changed from the aforementioned phosphorylation to a conformational one. It is likely that the lid dynamics have been perturbed. The viscogen data suggests that at high temperature the lid opening step reduces E_a (favourable reduction in enthalpy of activation) and A (disfavourable reduction of enthalpy of activation) thermodynamic parameters. This suggests that at high temperatures the barriers to disrupt bonds to reopen is more favourable (enthalpy based) and entropy now dictates turnover. With entropy driving the flux, at high temperatures the lid is most likely to remain open. What may be occurring is that at very high temperatures, the “irreversible” step is not chemical now as the propensity for this to occur is high, but rather the lid ordering is now vastly less likely to occur. Essentially, upon lid closure, the phosphorylation will always occur as the thermal energy present creates a “barrierless” reaction.

The Michaelis-Menten parameters derived with varying temperature (Table 3.3) were compiled to see the change in the catalytic efficiency (k_{cat}/K_m) with temperature. After the inflection point, polarPEPCK catalytic efficiency decreased. At these high temperatures, the lid will be most favourably disordered. In line with the lower k_{cat}/K_m, representing binding steps before the first chemical one, the inactivation may be due to an inability for the lid to order itself closed long enough for the phosphorylation step to occur. In this, the enzyme will release bound substrates before commitment through its catalytic cycle leading to the observed reduction in k_{cat}/K_m. This behaviour has been seen in another Ω-loop hinge mutant, H477R. Although the origin of the inability for the lid to order is different, a similar outcome was achieved where it was proposed that the enzyme capture of the substrate (and closure) was slower than the release of OAA. This manifested as many futile encounters occurring before successful transition through the reaction coordinate.[37] For the higher temperature variants, after the break-point catalytic efficiency improves. This improved catalytic efficiency is counter to the Arrhenius data suggesting at high temperatures lid ordering is the rate-determining step, as one would suspect that as in the case of polarPEPCK substrate would diffuse out of the active site. It may be that the higher temperature variants remained bound as they await lid closure, as opposed to the polarPEPCK where they may diffuse out. There is still some discussion as to the effect of thermophilicity on substrate binding, where some often thought because psychrophilic enzymes typically have a higher K_m at a given temperature they are unable to bind their substrates as tightly.[107,108] This is inaccurate in the sense that the catalytic efficiency, and not the K_m, is more indicative of the binding steps. It was proposed that as temperature increases for a psychrophile, the active site binding loops are more flexible and thus cannot bind the substrate. Later, it was computationally shown that the active site flexibility is nearly identical across the thermally adapted enzymes, and perhaps the global flexibility may be dictating changes in binding efficiency.[109110] It may be that polarPEPCK reduction in catalytic efficiancy is not due to the active site loops flexibility, in line with the conservation of the R- and P- loops and oxalate-GTP structures presented here, but is rather a decoupling of the global dynamics and ligand binding. Under this model, the higher temperature variants would better be adapted to maintain their global dynamics under the temperatures tested and would not suffer from a decrease in catalytic efficiency.

In the OAA decarboxylation reaction, the narrow temperature range does not allow a comparison between low and high temperatures. For the meso- and thermophilic PEPCKs, viscogen had no apparent effect on the Arrhenius plots. This suggests, as previously stated, the phosphorylation steps are rate limitings for those two variants.[38] PolarPEPCK however, does appear to be influenced by the presence of viscogen, leading to an increase in E_a as well as A. This suggests that viscogen presences helps facilitate turnover through entropic means. This would suggest that the polarPEPCK enzyme k_{cat} is at least partially controlled by lid reopening in contrast to its counterparts studied here. What may be occurring, but is highly speculative, is that under viscogen presence, the non-productive lid-reopening step has been slowed which allows a higher proportion of chemical steps to be completed creating a more efficient enzyme.

3.4.2 Mechanistic Interpretation of Structural Adaptations of PEPCKs

Without the open polarPEPCK structure the mechanistic interpretation of the transition of this lid from open to closed state is speculative, although the data does suggest that there may be a conserved mechanism of closure between the psychro- and mesophilic PEPCKs studied. To begin, the ligation state of the three structures lacking one or both metals were unsurprising. RcPEPCK is co-crystallized with Mn^{2+} as it induces nucleation of the crystals, so all structures solved of this isozyme has a metal bound to the M1 and M2 site (if nucleotide is present). Of the other PEPCKs solved, 3 structures, 2 from *Mycobacterium tuberculosis* and one from *Thermus thermophilus*, are missing metals when bound with either nucleotides or a ligand occupying the OAA/PEP site (PDB ID# 4WPU, 4RCG, 2PC9). What this suggests is that the substrates of PEPCK can bind without metals, but from kinetic experiments the enzyme cannot turn over without these metal requirements (experiments not shown). Most interestingly, the fact that *Thermus thermophilus* PEPCK doesn’t bind the M2 metal when GTP is bound suggests that this is not a temperature adaptation, but rather an issue of metals availability.

Based on the polarPEPCK structures, there does appear to be cooperativity between the M1 metal presence selecting for which nucleotide will bind. When no metals are present through the experiment, GDP binds preferentially. When the M1 metal is soaked in after the fact, GDP is still bound, suggesting that GDP will bind prior to the metal presence and cannot be freely exchanged. This is unsurprisingly as the structure is closed and the lid is closed, making it likely very hard to completely reopen for the bulkier substrate release. Back soak experiments, where the crystals were left in ligand free cryoprotectant for ~24 hours, did not allow the structure to relax and oxalate, and GDP were still observed to be found in the active site. Alternatively, when the metal is co-crystallized, GTP is selected for. This is the first evidence that the M1 metal is playing a role is nucleotide binding. A proposed mechanism for M1 selectivity of the nucleotide species is likely based on the M1 cationic charge stabilizing the γ-phosphate of GTP, whereas in GDP this large anionic charge is too far away for the M1 to stabilize it.

The maintenance of the closed lid of PEPCK is important for the efficient turnover of the enzyme. To maintain this closure, PEPCKs employ two areas of the lid, denoted the latch and hinge. The latch, which consists of the main loop, interacts with the R-loop. As shown, the mesophilic enzyme utilized more interactions (5 vs 3) than its psychrophilic counterpart. Of these interactions, there is one conserved backbone-backbone hydrogen bond. The other stabilizing contribution of the psychrophile is actually not from the latch region to R-loop, but rather is an arginine from the hinge, that extends across away from the hinge and towards to the R-loop. One could hypothesize that this interaction would be significantly controlled through entropic means. Arginine side-chains are highly entropic as they have many freely rotatable bonds. The probability of this arginine to be in this extended conformation is likely quite low, meaning that for this interaction to be sustained the side chain must adopt a lowly populated conformer.

The second structural feature of the lid that is required for the folding of the lid over the active site is the hinge region which consists of the entry/exit N- and C-terminal β sheets of the Ω-loop and the adjacent α helix. The rcPEPCK isozyme has been extensively studied to determine how this hinge changes during closure.[37] In brief, this hinge is asymmetrically anchored through the N-terminus, as shown by the unchanging electrostatic interactions between open and closed states. In contrast, the C-terminal hinge is changing upon closure, where a new salt bridge is formed. This asymmetry appears to be conserved in the psychrophilic enzyme, where the lack of N-terminal interactions suggests that in the open state, there would still be an absence of stabilizing forces. The only interaction holding the hinge in this closed conformation occurs from the C-terminus, by one electrostatic interaction. The sheer number of interactions maintaining the closed state that simply speaking, higher temperatures require the enzyme to employ more stabilizing forces to keep it closed. This pertains to both the hinge, as well as the latch regions. To offset the influence of temperature driving unfavourable entropically driven disorder, more enthalpic forces must be used to lower the free-energy of the conformational state. Although there are no thermophilic PEPCK structures of the GTP-dependent isoform, an ATP utilizing PEPCK from *Thermus thermophilus* has been deposited in the PDB as part of the RIKEN Structural Genomics/Proteomics Initiative (RSGI). This structure (PDB ID# 2PC9) has ATP bound, but is in an open state. Due to the lid being open, the latch interactions cannot be discerned but some information on hinge dynamics can be interpolated. The hinge region of the *Thermus thermophilus* PEPCK is predominately composed of hydrophobic and π-interactions as opposed to the general H-bonding and electrostatic stabilizing from the lower temperature variants (Figure 3.8). Any suggestion as to how this hinge operates would be mostly speculative beyond this high level observation but is consistent in adaptation in this dynamic elements with the above analysis.

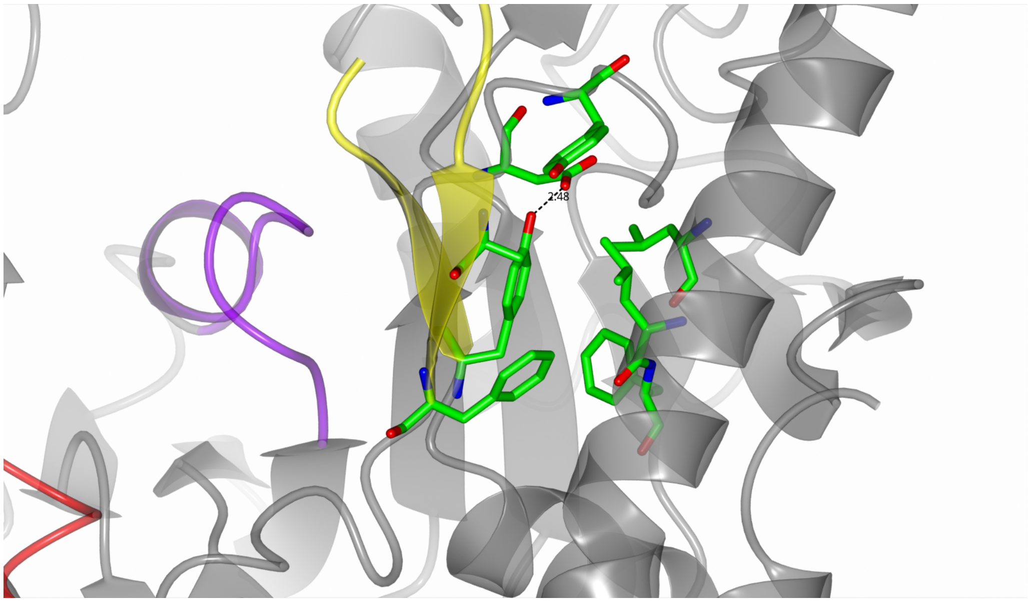


Figure 3.8: Hinge region of Ω-loop in open Thermus thermophilus PEPCK (PDB#2PC9). A shown, the hinge region is predominately hydrophobic and π-interactions. 1 Hydrogen-bond is shown with the black dotted line. The R-, P-, and Ω-loop are coloured in red, purple and yellow respectively. Atoms are coloured by type with carbon in green, oxygen in red, nitrogen in blue, phosphorous in purple, manganese in grey.

Hydrophobic bonds are shown to increase in strength as temperature coincidentally increases.[111] The proposed mechanism of action is that at higher temperatures, water is driven away from enthalpic H-bonds and entropy dominates. In this setting, water is effectively removed from water H-bond networks and into bulk solvent, thus strengthening the hydrophobic interactions. As water network are driven towards bulk solvent, this would also reduce hydrophilic interactions such as the obvious H-bonds, as well as ionic interactions. As *Thermus thermophilus* (T_{opt}~68°C) and other thermophilic organisms must live under these failing water networks, they may employ more hydrophobic interactions in order to dynamically transition through their reaction coordinate. This may explain the hydrophobic nature of the hinge region of *Thermus thermophilus* PEPCK, and the obvious bias of electrostatic and H-bond interactions used by the lower temperature isozymes described.

3.5 Summary of Results

This account is a detailed approach to understanding the thermal adaptation of enzymes in the context of a dynamic, mechanistic understanding, rather than the typical thermostability experiments. It appears that the general design principles in creating thermozymes are fairly well established, but the specific changes are, like much of protein biochemistry, not well understood. With further experimentation and validation, PEPCK could be included in the few enzymes that are well studied for their thermal adaptations (see [112] for review of thermal adaptation study model enzymes).

These new results show PEPCKs lowering $E_{a,}$ while increasing A with increasing thermophilicity, suggesting that thermophiles lower the entropic barriers whereas psychrophiles lower the enthalpic ones. PEPCKs also show a non-linear Arrhenius plot for the PEP carboxylation reaction, where at low temperatures chemical steps are rate limiting, changing at high temperature to lid-closure. PolarPEPCK specifically undergoes a reversible inactivation with increasing temperatures, while suffering a decrease in catalytic efficiency may be due to an increase in global dynamic fluctuations preventing the ligand bound state. In contrast, the meso- and thermophilic enzymes have an increased in catalytic efficiency where their rigidity may not allow for the aforementioned decoupling at the temperatures tested. In the OAA decarboxylation reaction, for the temperature window measured, the rate limiting step seems to be in-line with the previously suggested chemical steps, although polarPEPCK appears to be somewhat controlled by dynamic events. Interestingly, viscogen presences helps polarPEPCKs activity, suggesting slowing lid reopening is beneficial. The likely, but untested, conclusion is that upon slowing lid reopening more chemical events can go through to completion.

Structural evidence suggests that the dynamic elements (Ω-loop) responsible for catalysis, but not substrate binding, are adapted. The structural evidence shows that the psychrophilic PEPCK has lowered the number of enthalpically driven salt/H-bond interactions to maintain the closed state of the enzyme, while thermophilic PEPCKs may use hydrophobic interactions for these dynamic switches. This is once again, in line with the enthalpy-entropy compensation described above.

More models are required if we are to understand, and to utilize first principles in creating thermally adapted enzymes for whichever application is needed. Directed evolution seems to be the most probable way to adapt a molecular scaffold to a new temperature regime[113–115], but this does not answer the fundamental mechanistic questions. Furthermore, the most exciting addition to the investigating of thermally adapted enzymes is the understanding that the protein-water interface may be the source of adaptation to a temperature regime.[110] These advances and using more well characterized model enzymes will allow us to be able to determine what the true characteristics for thermal adaptations are as opposed to unique strategies employed by specific enzymes. These studies must not only probe the thermostability and low resolution thermodynamic trends, but also specific changes the mechanisms of dynamic changes enzymes must undergo to be proficient.

4.0 Biochemical, Structural, and Kinetic Characterization of PP_i-PEPCK from *Propionibacterium freudenreichii*

4.1 Introduction

Studying the differences between mutants or isozymes of PEPCK gives a method to understand subtle changes in the dynamics and the corresponding free-energy landscape. The more changes added, one could imagine a more drastic adjustment of its landscape. Furthermore, adding additional motifs and domains to the PEPCK scaffold may impart new substrate requirements, quaternary state changes, allosteric regulation and or other functionalities. Fortunately, a third PP_i dependent isoform of PEPCK is quite divergent from its nucleotide counterparts in both primary sequence and structure, which may offer insight into the relationship between these gross changes and function. This form of PEPCK can offer a new system to probe these structural rearrangements on their consequences while maintaining a level of control through a similar chemical outcome.

In the late 1960s and into the early 1970s, PP_i PEPCK (then known as phosphoenolpyruvate carboxytransphosphorylase) from *Propionibacterium freudenreichii* (*Pf*PEPCK) was studied in a series of papers. These biochemical investigations began to shed light on many details of this enzymes biochemical/physical properties including its metal dependencies, kinetic constants, CO_2 utilization, and quaternary state changes upon ligand binding.[61,116–120] Based on the prior work, the most obvious differences between the PP_i- and nucleotide dependent PEPCKs were determined to be that PP_i dependent PEPCK utilizes pyrophosphate as its phosphoryl substrate, it has a bias for the traditionally "reverse" PEP to OAA reaction, and molecular weight size difference being almost twice the size of nucleotide dependent PEPCKs (~127kDa for PP_i-, 70kDa for ATP/GTP-PEPCK). In general, PP_i dependent PEPCK predominately works through the carboxylation of PEP to form OAA, while also having the ability to perform a pyruvate kinase like activity reminiscent of the side-reaction of the nucleotide dependent PEPCKs described above.[118] These studies also suggested that like the nucleotide-dependent enzymes, PP_i PEPCK required divalent metal cations for its chemical reaction, with Co^{2+} observed to be the most activating species.[121,122] In the 1968 paper investigating if CO_2 or bicarbonate was being used as the substrate for carboxylation of PEP it was shown through both radiochemical and spectrophotometric approaches that CO_2 was the most likely substrate.[116] The initial characterization of PP_i PEPCK evaluated quaternary structure changes and its downstream regulation of the enzymes activity. Although quaternary state changes have not been found within GTP dependent PEPCKs and all enzymes characterized to date have been shown to be monomers, ATP dependent PEPCKs have been found to exist as multimers. *Trypanosoma bruci* and *Trypanosoma cruzi* PEPCK both have been shown to form dimers,[34,119] while even higher order structures have been reported, with *Saccharomyces cerevisiae* PEPCK forming tetramers and *Urochloa panicoides* forming hexamers.[124,125] The 1970 investigations showed that OAA, malate, and fumarate, all substrates of the TCA cycle downstream of where PEPCK would operate in the PEP to OAA direction, induce dimerization and subsequent inactivation of the enzyme.[120] This indicates a product inhibition mechanism operating to regulate PP_i dependent PEPCK activity. In addition to this inactivating dimer, a tetramer was also described, although active and short-lived it would dissociate into active dimers, however no mechanism for tetrameter formation was deduced. Based on these prior studies, a general scheme for the dynamic quaternary state changes was proposed (Figure 4.1).

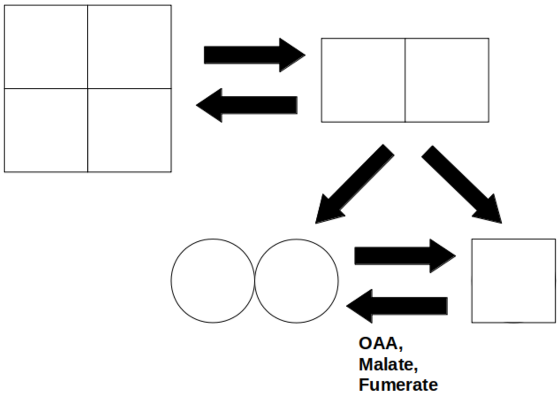


Figure 4.1: Proposed quaternary structure changes for PP_i dependent PEPCK. Adapted from O'Brien and Wood 1974.[120] Squares indicate active forms of the enzyme, while circles indicate inactivated forms.

In 2015, Chiba et al reintroduced PP_i dependent PEPCK to the research community by trying to understand the evolutionary link between it and nucleotide dependent PEPCKs.[58] With such little sequence conservation between PP_i dependent PEPCK and the nucleotide dependent PEPCKs, it was concluded that these two classes of enzymes likely arose from divergent sources. This paper also indicated that PP_i dependent PEPCK from *Entamobae histolytica*, forms heterotrimers from 3 different isoforms found in this parasite. Most recently, Chiba et al published the first crystal structure of a PP_i PEPCK from *Actinomyces israelii (Ai*PEPCK*)* and re-evaluated their previous conclusions as the active site architecture and catalytic residues between PP_i and nucleotide dependent PEPCKs are widely conserved (those responsible for M1 metal, PEP/OAA, and phosphate binding).[31] The Chiba study divided the *Ai*PEPCK structure into four different lobes and N and C terminal core domains where the active site is situated. It was shown that *Ai*PEPCKs employ a histidine which occludes nucleotides from binding to the phosphate binding site. Finally, the observed dimer interface was described.[31] Although this structure was very informative in adding to the understanding PP_i PEPCKs, the lack of updated biochemical data, specifically for *Ai*PEPCK, leaves many questions open. With the wealth of biochemical data from our labs studies on the rcPEPCK that was unknown to the 1970 researchers studying PP_i PEPCKs, this study aims to update the current understanding of this class of enzyme. In this study, I used the originally characterized PP_i PEPCK from *Propionibacterium freudenreichii* (*Pf*PEPCK) and completed a comprehensive investigation of this enzyme's structural and functional properties. All of the new information presented below, coupled with our current knowledge of nucleotide dependent PEPCKs, has allowed to discern the few similarities and many differences of this enzyme when compared to nucleotide dependent PEPCKs. Further, this foundation of biochemical data for the PP_i PEPCK, will allow future studies to probe interesting questions with regards to the structure/function of PP_i PEPCKs as well as its relationship to the nucleotide dependent PEPCKs.

4.2 Methods

4.2.1 Construct Design

The PP_i PEPCK (EC. 4.1.1.38) protein sequence from *Propionibacterium freudenreichii (*WP_013160152.1) and *Actinomyces israelii (*WP_043560275.1) were codon optimized for *Escherichia coli* and synthesized by GenScript (see Section 7.3). GenScript cloned these gene individually into the pESUMO-Star (Kan) vector (LifeSensors) between the Bsa1 and Xho1 sites so that the final expressed proteins had the SUMO 6xHIS fusion tag at the N-terminus. Upon cleavage of the fused protein, the resultant proteins have no additional amino acids.

4.2.2 Expression

The expression and purification for both *Pf*PEPCK and *Ai*PEPCK were identical and are outlined below. Four flasks with 50mL (supplemented with 50ug/mL of kanamycin) of LB media was inoculated and grown overnight at 37°C from glycerol stocks of PP_i PEPCK (BL21-DE3). These cells were centrifuged for 5 minutes at 5000G and the resultant pellets were resuspended into four flasks with 1L of ZYP-5052 media and 50ug/mL of kanamycin.[78] These cells were allowed to grow for ~20 hours at 20°C before they were harvested through centrifugation for 10 minutes at 5000G. The cell pellets were resuspended in prechilled B1 (Table 4.1).

Table 4.1: Buffer composition for protein purification for PP_i PEPCKs

Buffer 1 (B1)	50mM TRIS-OH (pH 8.0 at RoomTemperature) + 10mM Imidazole + 300mM NaCl + 2mM TCEP
Buffer 2 (B2)	50mM TRIS-OH (pH 8.0 at RoomTemperature) + 300mM Imidazole + 2mM TCEP
Buffer 3 (B3)	50mM TRIS-OH (pH 8.0 at RoomTemperature) + 2mM TCEP
Buffer 4 (B4)	50mM TRIS-OH (pH 8.0 at RoomTemperature) + 1M NaCl + 2mM TCEP

4.2.3 Purification

Once the cells were resolubilized in B1, they were passed through a prechilled french press twice at 1000PSI. All subsequent steps were completed at 4°C. The lysed cells were then centrifuged at 12000G for 30-45 minutes. The supernatant was incubated with nickel-NTA resin for 1 hour. This resin was then rinsed with B1 until the flow-through reached an A_{280} below 0.1. B2 was then passed over the resin to elute the bound protein into 10mL fractions. These fractions were pooled and concentrated to approximately 5mL. The protein was then passed down a P6-DG column pre-equilibrated in B3. Now buffer exchanged, the protein was left overnight to stir with ~1mg of SUMO protease (purified as described in[39]) to cleave the 6xHIS SUMO fusion domain from the PP_i PEPCK. This digested protein sample was then incubated with cleaned nickel-NTA resin equilibrated in B3, for 1 hour. The resin was washed with B3 where the flow-through containing protein was collected in 10mL fractions. This now SUMO fusion free sample of PP_i PEPCK was applied to HighQ anion exchange resin pre-equilibrated in B3 running on a Bio-Rad fPLC. A linear gradient was applied overnight from B3-B4 where the elution was collected. Peaks active for PP_i PEPCK and clean via SDS-PAGE analysis were then pooled and concentrated. One final buffer exchange with P6-DG resin into B3 was completed. The final purified protein was concentrated to 20mg/mL using $E_{280\ 1\%}$ extinction coefficients of 10.9 and 11.4 (as determined with ExPASy ProtParam tool) for *Pf*- and *Ai*PEPCK respectively. The protein was frozen in 30uL drops by immersion in liquid nitrogen and stored at -80°C.

4.2.4 Analytical Size-Exclusion Chromatography

Size-exclusion chromatography was utilized to determine the molecular weights of the ligand induced quaternary changes. An ENrich SEC650 10x300 column from Bio-Rad was used. 100µL injections at

5mg/mL were run down the pre-equilibrated column at 1mL/min. The standard buffer was 25mM HEPES pH 7.5 and 10mM DTT, with ligands added at varying concentrations to supplement the buffer.

4.2.5 Crystallization

“Dimeric” crystals (furthermore called malate-crystal) were initially grown at 10mg/mL through vapour diffusion at 25°C. The mother liquor condition consisted of 100mM HEPES pH 7.5, 200mM $MgCl_2$, 10% ethylene glycol, 7.5% isopropanol, 10-20% PEG 2000. Drops were set up with 2:4, 3:3, 4:2 μL ratios of protein:mother liquor with 10mM of malate in each drop. These initial crystals were unsuitable for diffraction and were harvested for seeds by collecting all crystals (with a plate morphology) from one drop into a stabilization solution consistent with the mother liquor condition with the addition of 10% PEG 400. A dilution series was created, and 0.5uL of these seeds were added to a fresh set of drops 24 hours after equilibration started. Crystals with a plate morphology grew with significant thickness and were harvested back into the stabilizing solution and soaked for 1 hour in the presence of 20mM malate and 1mM $FeCl_3$. Crystals were cryocooled by direct immersion into liquid nitrogen.

“Monomeric” crystals (furthermore called holo-crystal) were grown at 20mg/mL through vapour diffusion at 25°C. The mother liquor condition was 100mM Bis-Tris-Propane pH 8.0, 100mM NaCl, 100mM sodium formate, and 6-16% PEG 4000. Drops were set up in 2:4, 3:3, 4:2 μL ratios of protein:mother liquor. Crystals were harvested, and soaked for 1 hour in 100mM Bis-Tris-Propane at pH 8.0, 100mM NaCl, 100mM sodium formate, 20%PEG 2000, 10% PEG 400, and 1mM $FeCl_3$ cryoprotectant. Crystals were cryocooled by direct immersion into liquid nitrogen.

Data were collected at the Canadian Light Source in Saskatoon, Saskatchewan on the 08B1-1 beamline using a Rayonix MX300HE detector. Data were indexed and scaled in HKL-2000[79] with final statistics present in Table 4.2.

4.2.6 Structure Determination

The first structure solved was the higher resolution malate-crystal form through molecular replacement with the previously solved *Ai*PEPCK (PDB ID# 6K31). The model used was pruned in CHAINSAW and then used with MOLREP in the CCP4 package.[80,81] Iterative rounds of real-space refinements and addition of water molecules and heteroatoms were carried out in COOT[82] and REFMAC5.[83] Automatic local restraints were utilized throughout the refinement process. Final MolProbity statistics were collected with the online server at http://molprobity.biochem.duke.edu/index.php.[84]

The second structure, the ‘holo-crystal’, was solved using molecular replacement with molecule A of the malate structure and solved as described above. Final model statistics are present in Table 4.2.

Table 4.2: Crystallographic data table for the holo- and malate-crystal *Pf*PEPCK structures.

	Holo-Pf-PEPCK	**Malate-Pf-PEPCK**
Wavelength	1	1
Resolution range	67.7 - 3.2 (3.31 - 3.2)	74.8 - 2.65 (2.75 - 2.65)
Space group	P 21 21 21	P 1
Unit cell	140.466 212.189 253.293 90 90 90	134.131 146.952 151.094 82.702 83.999 70.798
Total reflections	5764402	3620682
Unique reflections	124714 (12315)	304269 (29413)
Multiplicity	13.1 (11.3)	3.8 (3.7)
Completeness (%)	99.6 (99.5)	97.1 (93.8)
Mean I/sigma(I)	16.3 (1.3)	10.4 (2)
Wilson B-factor	101.89	46.73
R-merge	0.14 (1.95)	0.118 (0.757)
R-meas	0.15 (2.04)	0.137 (0.884)
R-pim	0.04 (0.86)	0.07 (0.453)
CC1/2	0.91 (0.51)	0.93 (0.72)
Reflections used in refinement	124688 (12315)	304220 (29413)
Reflections used for R-free	6206 (643)	15255 (1531)
R-work	0.272 (0.352)	0.217 (0.316)
R-free	0.29 (0.374)	0.234 (0.354)
Number of non-hydrogen atom	34809	72238
macromolecules	34799	70969
ligands	4	8
solvent	6	1261
Protein residues	4391	8948
RMS(bonds)	0.013	0.013
RMS(angles)	1.64	1.66
Ramachandran favored (%)	97.1	97
Ramachandran allowed (%)	2.75	2.87
Ramachandran outliers (%)	0.18	0.17
Rotamer outliers (%)	0.67	1.81
Clashscore	1.71	2.32
Average B-factor	134	54.9
macromolecules	134	55.2
ligands	106	19.4
solvent	64.4	38.1

4.2.7 Kinetics

To determine each kinetic constant, all other substrates were held constant and the assays were completed in duplicate. All assays were completed at 25°C in pre-equilibrated kinetic assay mixes at a final volume of 1mL with the use of a temperature controller. A CaryUV100 spectrophotometer was used to monitor the reduction of NADH to NAD^+ in the coupled assays at 340nm.

As previously shown, OAA can spontaneous decarboxylate to pyruvate in the presence of metals.[103] To measure this background rate the standard assays were used with the exception that PP_i PEPCK was not present in the assay mix. All rates have been normalized for this background rate depending on the metal mix used unless otherwise specified.

$OAA + PP_i \rightarrow PEP + PO_4^- + CO_2$

The standard assay mix for OAA decarboxylation was composed of 100mM HEPES pH 7.5, 10mM DTT, 300µM NADH, 250µM $PP_{i,}$ $_1$mM $MgCl_2$ (4:1 PP_i:Metal ratio), 100µM $FeCl_3$ (or other M1 metal substitute), 500µM OAA, 1mM ADP, 50µg PK, 5U LDH, and 2.5µg of PP_i PEPCK. Reactions were started upon addition of OAA.

$OAA + PP_i \rightarrow$ Pyruvate

The standard assay mix described above in the OAA→PEP section was used with the exception that pyruvate kinase was not present in order to measure the enzymatic production of pyruvate from OAA. 6µg of PP_i PEPCK were used for these experiments.

PEP + PO_4^- + CO_2 → OAA

The standard assay mix for PEP carboxylation was composed of 100mM HEPES pH 7.5, 10mM DTT, 300µM NADH, 10mM $PO_{4,}^{-}$, 40mM $MgCl_2$ (4:1 PO_4^-:Metal ratio), 100µM $FeCl_3$ (or other M1 metal substitute), 1mM PEP, 50mM $KHCO_{3-}$ (bubbled with dry ice), 10U of MDH and 2.5µg of PP_i PEPCK. Reactions were started with the addition of PP_i PEPCK.

PEP + PP_i → Pyruvate

The standard assay mix described above in the PEP → OAA section was used for the pyruvate-kinase like reaction with the exception that LDH was added in substitution of MDH.

IC_{50} Determination for both OAA → PEP and PEP → OAA reactions

To determine the IC_{50} for various characterized inhibitors of PEPCK, the assays were completed under the same conditions as above, with the only exceptions being a concentration of 50µM OAA and 100µM PEP were used in the respective reactions.

4.2.8 Kinetic Data Analysis and Equations

Michaelis-Menten and substrate inhibition constants were derived from plots created in SigmaPlot11 EnzymeKinetics module using equation 4.1 and equation 4.2 respectively.

K_i values determined from the full progress curves were derived from the kinetic constants derived from Equation 4.1 when replotted on graphs k_{cat}/K_m vs. [I] and $1/k_{cat}$ vs. [I] to determine the K_i of both competitive and non-competitive inhibition, using equations 4.3 and 4.4, respectively.

IC_{50} values were plotted with GraphPad Prism8 using equation 4.5. K_i values were determined from these IC_{50} values using equation 4.6.

$$v=\frac{V_{max,app}[S]}{K_{m,app}+[S]} \tag{4.1}$$

$$v=\frac{k_{cat}[S]}{K_m+[S]+\frac{[S_2]}{K_i}} \tag{4.2}$$

$$\frac{K_{m.app}}{V_{max,app}}=\frac{K_m}{V_{max}K_i}[I]+\frac{K_m}{V_{max}} \tag{4.3}$$

$$\frac{1}{V_{max,app}}=\frac{1}{V_{max}K_i}[I]+\frac{1}{V_{max}} \tag{4.4}$$

$$y=min\frac{max-min}{1+10^{x-logIC_{50}}} \tag{4.5}$$

$$K_i=\frac{IC_{50}}{1+\frac{[S]}{K_m}} \tag{4.6}$$

4.3 Results and Discussion

4.3.1 Ligand induced quaternary state changes through size-exclusion chromatography

The previous biochemical assessments completed in the 1970s evaluated the quaternary state changes under ligand influence by sedimentation.[126] These state changes were then monitored for activity in which the dimer arising from monomers was found to be inhibitory. This dimerization was induced by OAA, malate, and fumarate. All substrates and malate were tested to confirm these findings (Figure 4.2, Table 4.3). Surprisingly, PEP at 10mM was found to shift the monomer closer towards the dimer form. As discussed at length later, PEP was observed to elicit substrate inhibition at sufficiently high concentrations. This new data indicates that in addition to malate, OAA, and fumarate, PEP can contribute to the inhibitory dimerization. This quaternary regulation from malate, OAA, and fumarate is likely a feedback mechanism to shut off flux through PP_i PEPCK as these are all downstream metabolites to PP_i PEPCK. In this scenario, PEP-induced dimerization is counter-intuitive as the organism would likely want to maintain flux under circumstances of high PEP and may be artifactual, showing some level of promiscuity in this dimer-inducing binding site. Finally, *Ai*PEPCK was tested for quaternary state changes under various conditions. *Ai*PEPCK was found to be a dimer when no ligands are present[31] and was unable, under all conditions tested, to form a monomer. This may indicate that *Ai*PEPCK is constitutively dimeric. With the environmental conditions and known kinetic consequences determined for the formation of the dimer, high resolution crystal structures were determined to hopefully lend insight into structural transitions between the monomer and dimer.

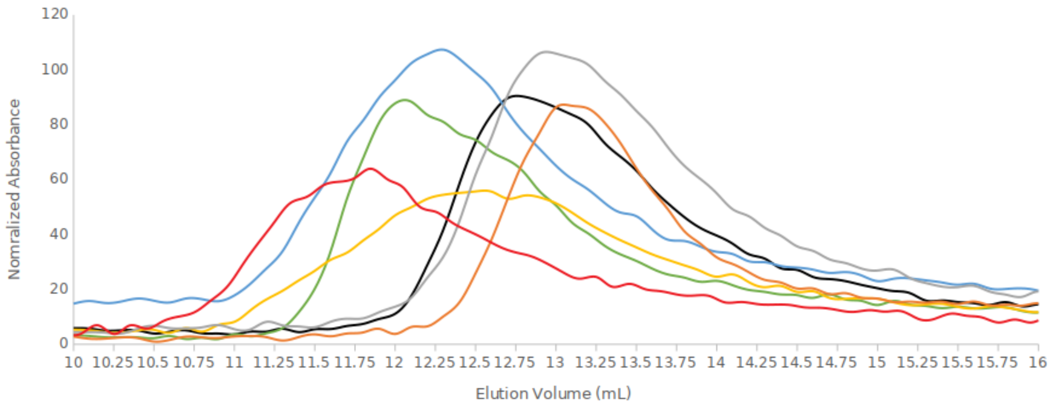


Figure 4.2: Size exclusion traces from various runs. Based on previously ran standards 13.4mL is the elution volume denoting approximately 127kDa and 12.2mL is approximately 250kDa. Red – *Ai*PEPCK with no ligands, Blue – *Pf*PEPCK 1mM OAA, Green – *Pf*PEPCK 10mM malate, Yellow – *Pf*PEPCK 10mM PEP, Grey – *Pf*PEPCK 1mM PEP, *Pf*PEPCK10mM PP_i – Black, Orange – *Pf*PEPCK with no ligands.

Table 4.3: Estimated quaternary states based on size-exclusion molecular weight estimation for both PP_i PEPCK from *P. Freudenreichii* and *A. israel.*

Size-Exclusion Molecular Weight Estimates

Ligands	Quaternary State for *P. freudenreichii*	Quaternary State for *A. israelii*
No ligands present	Monomer	Dimer
1mM PEP	Monomer	Dimer
10mM PEP	(Shift to) Dimer	
10mM PO_4^-	Monomer	
100mM $KHCO_3^-$	Monomer	
1mM OAA	Dimer	Dimer
10mM PP_i	Monomer	
10mM Malate	Dimer	Dimer
1mM βSP	Dimer	
150mM NaCl		Dimer

4.3.2 Crystal Structures

With the quaternary state changes now fully characterized for *Pf*- and *Ai*PEPCKs, the atomic details are necessary to understand how the dimer prevents the enzyme from catalyzing the conversion of PEP to OAA. To do so, two crystal forms of the *Pf*PEPCK were solved, one without any ligands present and one with 10mM malate co-crystallized to induce the dimer formation. Both crystal forms were a series of dimers, with the malate-structure having 8 molecules in the ASU making 4 sets of dimers, while the holo-structure had 4 molecules in the ASU corresponding with 2 sets of dimers. The interface surface areas, and dimer RMSDs after global alignment are described in Table 7.3.1. As previously mentioned, the coordinating active site residues responsible for coordinating substrates/metals of the previously solved *Ai*PEPCK enzyme, and now *Pf*PEPCK, are completely conserved compared to the previously solved GTP- and ATP-dependent PEPCKs and therefore will not be discussed below.

4.3.3 All Crystals Induce Dimerization

Upon inspection of both the holo- and malate-crystals, they are in fact sets of dimers similar to what has been described previously (Figure 4.3).[31] This observation suggests that this dimer is in fact a conserved quaternary structure of PP_i PEPCKs (at least for those which forms dimers), and crystallizing this enzyme may not be amenable to see the holo-enzyme as it preferentially forms a stable dimer *in crystallo*. This does indicate that the dimer is a local free-energy minima, as there is some propensity to form in absence of the inducing ligands. This does not rule out the possibility that other crystal forms could yield different conformations, or potentially a new co-crystal form of PEP and PO_4^- (or their analogues) may allow us to see the Michaelis-Menten complex. As seen from the RMSD values (Table 7.3.1), the relative positions of the dimers are essentially the same between both of the *Pf*PEPCK structures solved. There are two particular regions that vary across the different monomeric units. The first region, is a helix (residues 31-51) that changes register in comparison to the adjacent structural elements. The second variable region is a domain spanning residues 490-607 that is found to be in different conformations between all structures and will be further elaborated on below. Like with the variable helix, this domain can be found in three general positions (see Figure 4.4B and C for the domain and helix respectively). Each monomer can be classified based on these positions of the variable helix and domain. It does appear that there may be some coupling between this helix and the conformation of the domain.   For instance, monomers (chains ADEFG) from the malate-crystal have the same position of both the helix and domain. The second grouping is with chains BCG and AD of the malate- and holo-crystals, while the third group is chains B and C of the holo-crystals. For simplicity, the structural alignments of all dimer pairs are not shown.

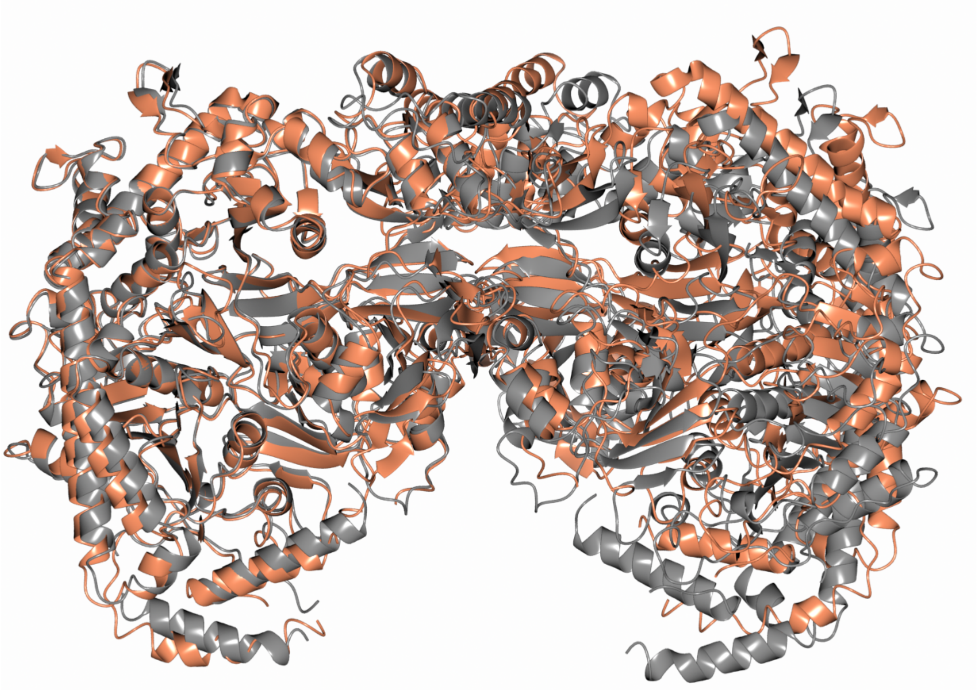


Figure 4.3: Alignment of malate-crystal dimer (grey, chains A-F) and *Ai*PEPCK dimer (coral, chains A-A') anchored during superposition with chain A.

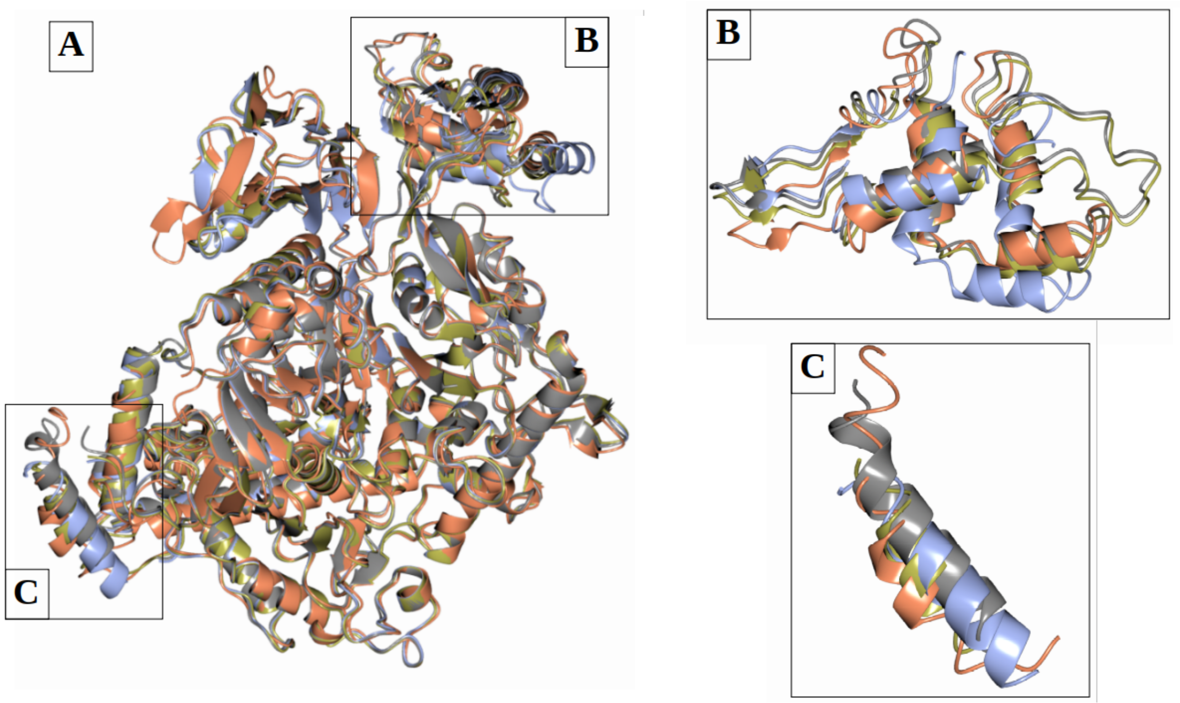


Figure 4.4: Representative alignment of various monomers of PP$_i$ PEPCK. Panel B shows the domain-500, while panel C shows helix 31-51. Chains are coloured as holo-A – ice blue, holo-C – gold, malate-A – grey, *Ai*PEPCK A – coral.

4.3.4 Relative Conformation of Domain-500

As described above, the vast majority of the enzyme structure is conserved among dimer pairs. When aligning the individual monomers of the two structures, the core of the enzyme is unchanged (Figure 4.4). As mentioned, one domain spanning residues 490 to 608 (of the previously described lobe 2[31]) was variable and in most part, the most difficult area to model into the electron density maps (Figure 4.5). This domain, furthermore referred to as domain-500, contains one of the loops responsible for the dimer interface interactions (loop 590-600). Although there is a plasticity to the location of this domain, the overall folded structure of the domain and the interface are conserved. This indicates that the relative position changes of this domain is, at least with the current information, inconsequential to the dimer formation as long as the interface is maintained. The relative dynamic nature of this domain may be important for sampling many different locations to correctly insert this loop into its contacting pocket of its mate, allowing for dimer formation. This domain may “sense” for its partner and establish the first contact in the dimer, after which the second interface can properly align. Further studies evaluating this particular domain, as well as its coupling to helix 31-51 may be important in understanding the dimerization phenomenon.

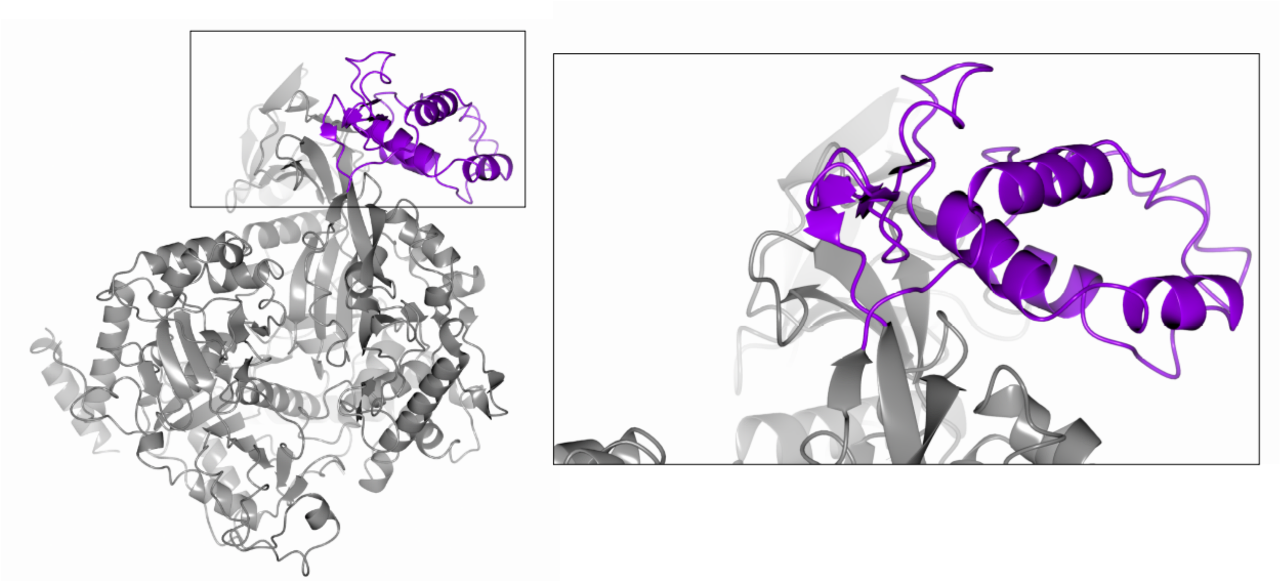


Figure 4.5: Malate-crystal monomer (chain A: grey) with the domain-500 shown in purple. The entry and exit points of this domain are very close to one another as it begins from a 2-strand β-sheet.

4.3.6 Dimer Interfaces

Although the dimer structures are conserved between the *Ai* and *Pf* PEPCK isozymes, the interface interactions are variable. These differences may be important in uncovering the nature of the constitutive dimer form of the *Ai*PEPCK in comparison to the ligand induce form of the *Pf*PEPCK. Chiba et al previously described the *Ai*PEPCK interfaces and will be used as a reference in the comparison.[31] The malate dimer from chains A-F was used as it is the best modelled dimer, although when comparing every dimer pair from both malate-crystal and holo-crystal these interfaces are generally well modelled and congruent.

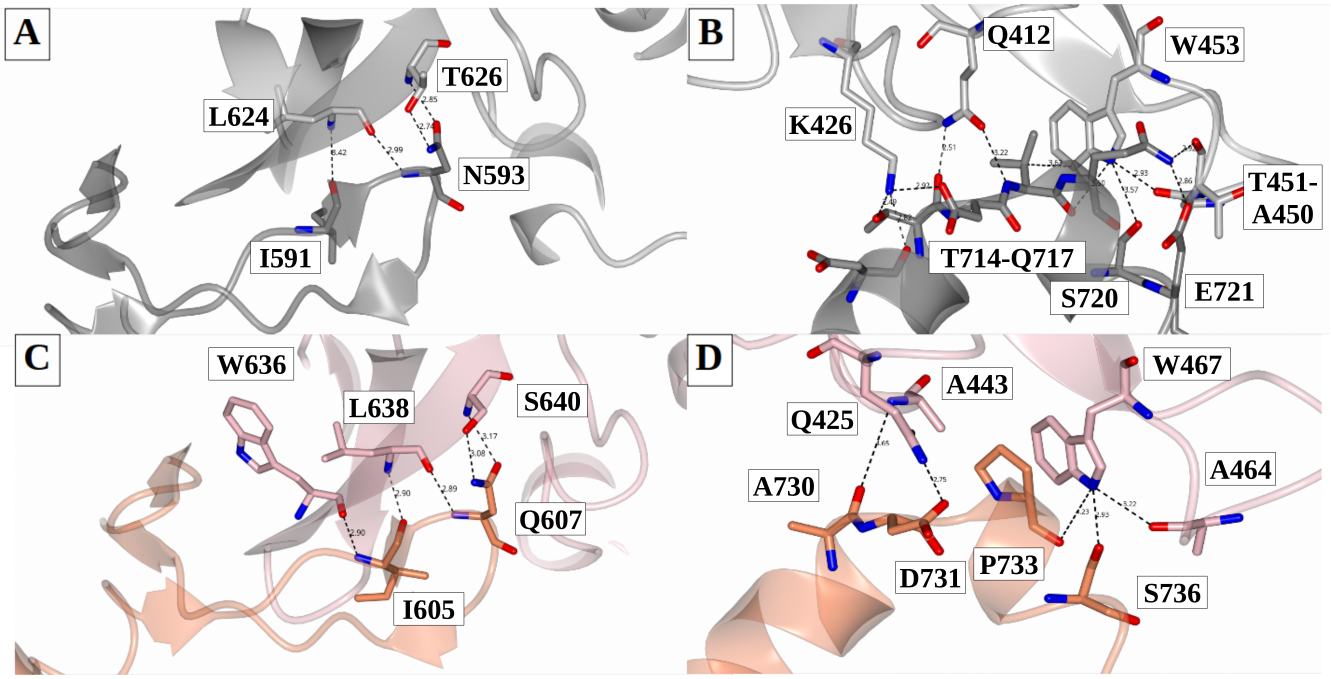


Figure 4.6: Dimer interface interactions for both *Pf*PEPCK and *Ai*PEPCK. Panel A and B are for interface 1 and 2 respectively from the malate-crystal (dimer A-F). Monomer of chain A is represented in dark grey, while chain F is in light grey. Panel C and D are representations of interface 1 and 2 respectively from *Ai*PEPCK (dimer A-A'). Monomer of chain A is represented in coral, while chain A' is in like pink. Atoms are coloured by type with carbon in chain colour, nitrogen as blue, and oxygen as red. Interaction distances are labelled and represented as black dashed lines.

Interface 1 is the corresponding dimer surface of this previously described domain-500. A loop (590-600 *Pf*PEPCK numbering) containing interacting residues is shown (Figure 4.6A and C). This loop inserts to interact with its dimer mate through mostly conserved interactions, although the backbone H-bond of the amide of I591and the carbonyl of W622 (*Pf*PEPCK numbering) is missing in the *Pf*PEPCK interface (Figure 4.6C). Although this tryptophan is conserved, the positions of the backbone amide and carbonyl increased the distance from 2.9Å in the *Ai*PEPCK structure to 4.7Å in *Pf*PEPCK. In brief, 2 backbone hydrogen bonds stabilize this interface, the first of which is between the carbonyl of I591 and the amide of L624, while the second is between the carbonyl of L624 and the amide of N593. The side chain of N593 also forms two hydrogen bonds with the backbone amide and carbonyl of T626 (S640 in *Ai*PEPCK).

Interface 2 is more variable between the two isozymes (Figure 4.6B and D). Although previously described elsewhere[31], *Ai*PEPCKs interface consists of a backbone hydrogen bond between the amide of A443 and the carbonyl of A730, a hydrogen bond between the terminal amide of Q425 and carboxylic acid of D731, and finally a tridentate hydrogen bond network between the backbone carbonyl of P733, the hydroxyl of S736, and the carbonyl of A450 all coordinating with the indole of W467 (Figure 4.6B). Panel C shows the interface for *Pf*PEPCK. First, the *Ai*PEPCK A443-A730 backbone H-bond is removed for the side chain amine of K426 forming three H-bond network with the hydroxyl of T714, the carbonyl of D711, and the carboxylic acid of D715. Secondly, *Ai*PEPCKs Q425 H-bond with D731 has been replaced for two H-bonds between D412 and the backbone of carbonyl of T714 and the amide of V716. Thirdly, the tridentate H-bond network of W453 is conserved although the proline is now V716. In addition to these two H-bonds, it is likely that the V716 sidechain may form a hydrophobic interaction with W453 phenyl ring. Finally, a bidentate H-bond network has been introduced between the side-chain amide of Q717, carboxylic acid of E721 and the backbone carbonyl of T415.

In total, the *Ai*PEPCK dimer is stabilized through a total of 9 H-bonds (5 at interface 1, and 4 at interface 2), while *Pf*PEPCK is stabilized through a total of 12 H-bonds (4 at interface 1 and 8 at interface 2) with an additional hydrophobic interaction at interface 2. This is somewhat surprising as it could have been hypothesized that the constitutive dimeric *Ai*PEPCK would have more stabilizing forces to maintain the quaternary structure. Therefore, it is hypothesized that the basis for the constitutive dimeric form of the *Ai*PEPCK is likely influenced by some other factors yet to be determined. One potential reason would be its primary sequence differences lock this enzyme into a conformation that allows the dimer to form after translation, whereas the *Pf*PEPCK must have a ligand present to switch it from the unknown dynamic holo-state to its rigid stable dimer. Unfortunately, with the current information, the conformation of the holo-state of the *Pf*PEPCK is unknown and therefore it cannot be infer what the dynamic transition from monomeric to dimeric might look like.

4.3.7 Active Site Loops

In the previous paper first describing the structure of PP_i PEPCK, it was observed that the catalytic residues between both the PP_i- and GTP-PEPCKs were completely conserved.[31] Although not surprising, as the same substrates are being used (and products produced) a comparison of the important active site loops may be insightful to understand the mechanistic details of this enzyme. It does appear that the three active site loops previously characterized are structurally conserved (Figure 4.7), while the sequences are quite different (Figure 7.3.1). It is not entirely unreasonable to assume there would be some molecular rearrangements of these active site loops upon ligand binding, transitioning from an "open" state to a chemical competent "closed" one, or in contrast to activation, upon dimer formation creating a impotent conformation. This transition is well characterized in rcPEPCK, where the loops undergo a transition from a disordered state to an ordered state as described earlier. In contrast, in the three PP_i PEPCK structures, all of these loops are well modelled indicating that they are not disordered. This suggests that the dimeric state may rigidify the catalytic loop, preventing proper coordination of ligands for the carboxylation of PEP. The kinetic data presented later are consistent with this interpretation, although it does appear that the dimer will allow for OAA decarboxylation. A co-crystal structure of the dimer and OAA/βSP would be very useful in understanding how this enzyme in its dimeric form can facilitate chemistry of one of the two reactions. Both the R- and P-loop look to be in a conformation that is similar to that of both the open holo-state or closed-state in the rat cytosolic enzyme (Figure 4.7A). The Ω-loop however is in a new, rather unexpected conformation where it looks to be pinned back away from the active site (Figure 4.7B). Typically, unless PEPCK is closed upon ligand binding, or perturbed through mutations, the Ω loop is typically disordered in its open state where in all three dimer structures of PP_i PEPCK the Ω-loop is ordered. Upon further investigation, there are four interactions between the Ω loop and the surrounding periphery. The first interaction is a H-bond between the N-terminal hinge β-sheet carbonyl of M842 of the loop and guanidine of R808. As described in detail previously, a mutation study of rat cytosolic PEPCK investigating the mechanism of lid closure showed that these β sheet hinge regions are important.[37] In brief, the N-terminal hinge is anchored throughout the open to closed transition, whereas the C-terminal hinge acts as a switch to allow for a folding event and subsequent closure of the active site. Without a catalytically competent closed state, we cannot determine that this interaction is similar to the anchoring as seen in the nucleotide forms. The rest of the interactions stabilizing the pinned conformation of the Ω-loop are throughout the loop. The first is between the amine of K845 and the hydroxyl of S472 of the P-loop. This may indicate a level of cooperativity between PP_i/PO_4^-

binding inducing a conformational change of the P-loop and the closure of the Ω loop, similarly see in the nucleotide dependent forms.[29] The final two interactions are H-bonds between the backbone amide and carbonyl of A852 and the hydroxyl of T998 and the guanidine of R1001 respectively.

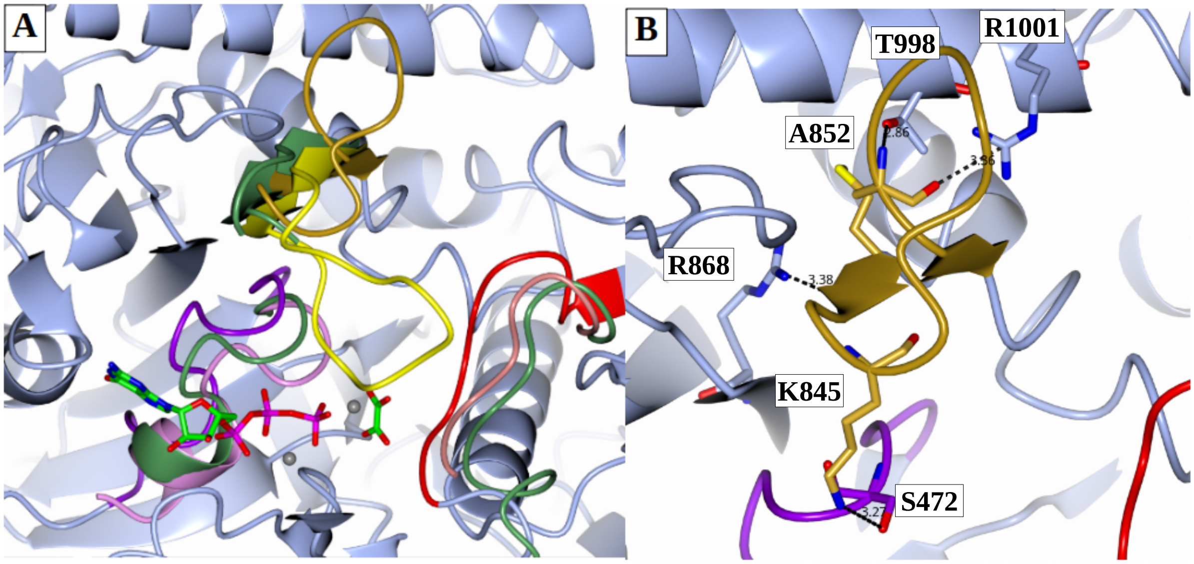


Figure 4.7: Structural Alignment of rc and malate-*Pf*PEPCK active site loops. Panel A - structural alignment of malate-*Pf*PEPCK and rcPEPCK complex with GTP and intermediate analogue oxalate (PDB ID# 3DT2), and the holo structure of rcPEPCK (PDB ID# 2QEW). Malate-*Pf*PEPCK loops are coloured as followed, Ω-loop gold, P-loop purple, R-loop red. RcPEPCK loops are coloured as followed, Ω-loop yellow (closed conformation), P-loop pink, R-loop salmon. The holo structure loops are show in green. Panel B – shows the molecular interactions between the Ω-loop and its surrounding residues. All other atoms are coloured by type; carbon is green, nitrogen is blue, red is oxygen, purple is phosphorous, grey is manganese.

4.3.8 Malate Binding Site

Although malate was both co-crystallized and soaked into the malate-crystals, no malate was found in the structure. This was an unexpected observation, as malate was present throughout the crystallization experiment and cryo-protection to ensure its presence. The current data suggests the active site is likely not the binding site, as there are no accessibility constraints providing evidence as to why malate would not be present. There are two alternative hypotheses as to why malate was not bound. The first, and least likely option, is that the concentration of malate was not high enough to occupy its binding site. This is unlikely as gel filtration data suggests that 10mM is enough to induce complete dimerization, although the kinetic data suggests 20mM is only modestly higher (1.5 times) than the expected K_i. Using a higher affinity inhibitor such as OAA or βSP may be more amenable. The second hypothesis is that the dimerization occurs in a two step fashion, where malate binds to the monomeric form and inducing a dimer-competent conformer. Upon dimer formation, the binding efficiency of malate is decreased and malate is not needed to maintain the dimer.

4.3.9 Insights into Tetramer Formation

In the 1972, the tetramer of *Pf*PEPCK was first determined and was further characterized.[126] This tetramer was active, and dissociated into active dimers but was short lived. An electron micrograph was collected to show the tetrameric form of the enzyme but it appears that it had to be fixed with glutaraldehyde prior to imaging.[126] To date, there has been no evidence that this tetrameric form can be induced from monomers. What is most likely is that this tetramer is an artifact of resolubilizing the crystals which were produced as a final purification step in the initial biochemical characterizations. Supporting this later idea, upon investigating the three crystallographic forms and their lattice arrangements, a common tetrameric species is observed (Figure 4.8). It may be possible that the crystallization of these enzymes not only preferentially selects for the dimer, but these dimers can stack with one another to form a stable tetramer. This tetramer may be transiently stable upon the breakdown of the lattice during resolubilization but ultimately, may be artifactual to the crystallization process.

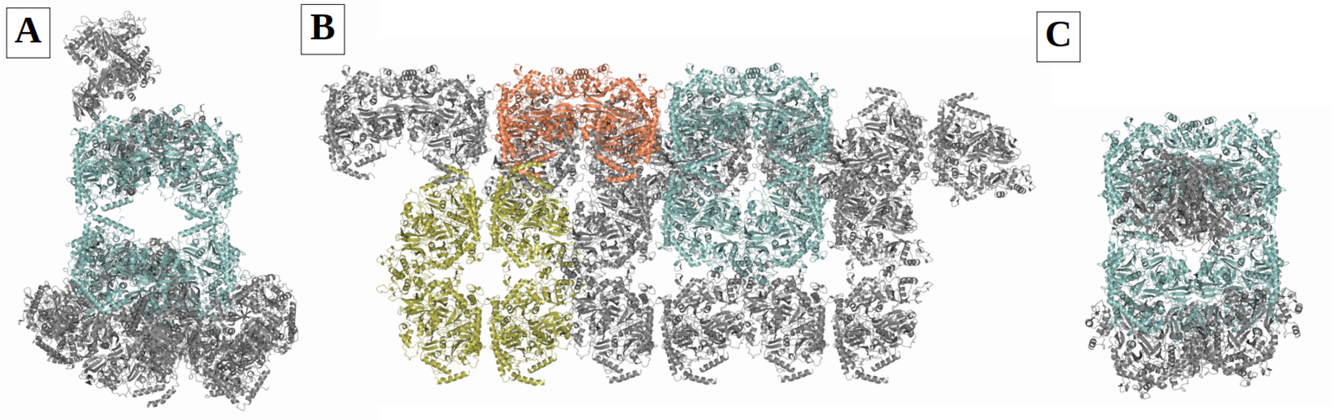


Figure 4.8: Crystal lattice tetramer. Panel A – holo-crystal lattice. Panel B – malate-crystal lattice. Panel C – *Ai*PEPCK crystal lattice. The observed tetramer is in green, while the other monomeric units are in grey. The dimeric form of PP_i PEPCK is shown in coral, and an alternative tetramer found in the malate-crystal is shown in gold although this tetramer is not found in the other crystal forms.

4.3.10 *Actinomyces israelii* PEPCK kinetics

In light of the new structural information deposited (PDB ID# 6K31), *Ai*PEPCK was expressed and purified to compare its kinetic properties with that of *Pf*PEPCK. Under all conditions tested, it was observed that this enzyme was inactive. This is in agreement with the size-exclusion data presented earlier (Figure 4.2 and Table 4.3) indicating that *Ai*PEPCK is always found to be dimeric and therefore it may be a possibility that the role for this protein is non-enzymatic. Furthermore, the metabolic role of PEPCK is likely fulfilled by the two other GTP-dependent isozymes found in *Actinomyces israelii.*

4.3.11 Metal activation of PP_i PEPCK

The initial biochemical characterization of *Pf*PEPCK examined metal activation.[61] They indicated that Co^{2+} was the most activating cation for the both the OAA- and PEP-forming reactions being approximately 3 times more activating than Mn^{2+} (or Zn^{2+} which is 2 times more activating than Mn^{2+}). In addition, for the pyruvate kinase reaction, Cu^{2+} was 2 times and 6 times more activating than Co^{2+} and Mn^{2+} respectively. When these metal activation experiments were revisited, similar results were obtained (Table 4.4). Mn^{2+} and Zn^{2+} are both similarly activating for the OAA-producing reaction - however Co^{2+} could not be effectively tested as it precipitated heavily in the reaction mix and no activity was measured once partially solubilized. Unfortunately for the 1969 study, Fe^{2+} was not tested, which in the current studies was observed to be the most activating cation for *Pf*PEPCK. This is the first indication that Fe^{2+} as being the true M1 metal cofactor for any PEPCK. Fe^{2+} has been tested with nucleotide dependent PEPCKs, but it has never been shown to be most activating among divalent metal cations when the reaction is supplemented with Mg^{2+}.[33] As iron can be found in a 3+ or 2+ state, the true electronic state of iron is unknown. Reducing agents appear to be vital for the maximal activity of the enzyme (data not shown) in both these studies and those previous, suggesting that Fe^{2+} may be the preferred oxidation state.[117] Further x-ray absorption near edge spectroscopy (XANES) experiments should be completed to confirm these initial findings.

Table 4.4: Metal activity specific activity measurements. Reactions were completed under the PEP → OAA standard reaction conditions. Cobalt could not be measured as cobalt precipitated heavily under the reducing condition.

Metal Activation of PEP → OAA Reaction

Metal	Specific Activity (μmolmin^{-1}mg^{-1})
$FeCl_3$	14 ± 0.5
$CaCl_2$	1.9 ± 0.2
$CoCl_2$	ND
$MgCl_2$	1.8 ± 0.2
$MnCl_2$	4.8 ± 0.3
$NiCl_2$	1.1 ± 0.1
$CuCl_2$	4.2 ± 0.2
$ZnCl_2$	4.7 ± 0.1

4.3.12 PEP carboxylation reaction

The finding that Fe^{2+} is is the most activating M1 metal makes the previous kinetic data and Michaelis-Menten parameters more difficult to compare as cobalt was used in their standard assay reactions. From the previous studies, the PEP carboxylation reaction had a specific activity approximately 5 times higher than that of the OAA decarboxylation reaction.[117] This is in agreement with this study as the PEP carboxylation reaction is approximately an order of magnitude more active than the reverse OAA decarboxylation reaction (Table 4.5 and 4.6). Functionally, *Propionibacterium freudenreichii* has been characterized as a carbon fixing organism through PP_i PEPCK activity, in line with the preferred PEP carboxylating reaction measured.[127–129] This clear bias towards PEP carboxylation has not been seen in GTP/ATP dependent PEPCKs, as all isozymes studied have relative rates are much closer to another other and all showing a preference for PEP synthesis. Further kinetic characterizations with different PP_i PEPCKs will indicate if this catalytic bias is specific to this isozyme or more ubiquitous among this class of enzyme. Although in most cases GTP/ATP PEPCKs catalyze the conversion of OAA to PEP more favourably as seen in both *in vivo* and *in vitro* data, GTP/ATP PEPCKs potentially preferring the PEP to OAA reaction have been noted. Of the studies where both reactions were studied, Case et al noted that the GTP dependent PEPCK from *Mycobacterium smegmatis* catalyzed the PEP carboxylation reaction with a specific activity of 43μmolmin^{-1}mg^{-1} in contrast to the observed 12.3μmolmin^{-1}mg^{-1} in the opposing direction. When evaluating the k_{cat}/K_m ratio for these reactions for the PEP carboxylating ($0.97x10^5$ $M^{-1}s^{-1}$) and OAA decarboxylation ($37x10^5$ $M^{-1}s^{-1}$) reaction OAA binds tighter by a factor of approximately 30.[24] With this significantly different catalytic efficiency for the two reactions, *M. Smegmatis* PEPCK is still likely a predominate OAA decarboxylating PEPCK. Even in this case, the opposing kinetic parameters exemplify that GTP/ATP dependent PEPCKs flux is more bidirectional than that of PP_i PEPCK. The lack of significant kinetic differences among the ATP/GTP dependent PEPCKs studied indicate that rcPEPCK, the best studied isozyme, is a good proxy for comparing PP_i and nucleotide utilizing PEPCKs. Henceforth, the main comparative enzyme discussed will be the rat cytosolic form.

The reported PEP K_m (74µM) here is almost an order of magnitude smaller than that previously reported (K_m=530µM).[126] This original higher K_m is more in line with the values reported for nucleotide PEPCKs with the rat cytosolic enzyme K_m being 294µM.[39] The previously reported values of PO_4^- K_m (K_m = 1.25mM) are approximately one half of that here (K_m = 2.3mM) but still generally quite high. The seemingly high K_m, and resulting lower k_{cat}/K_m of PO_4^- is not likely to be consequential in the cellular context . Due to its metabolic importance, inorganic phosphate/polyphosphates are found to be in high concentration *in vivo*. An estimation of inorganic phosphate concentrations of 50mM based on elemental composition of *Saccharomyces cerevisiae* was proposed by van Eunen et al 2009, and other literature suggests cellular concentrations between 10-75mM.[130–133] This indicates that the the excess inorganic phosphate in the cellular environment would saturate the enzyme. The CO_2 concentrations recorded below were derived based on the conversion to estimate CO_2 from bicarbonate.[134] The previous study of *Pf*PEPCK recorded a CO_2 K_m of approximately 550µM (at pH 7.8), four fold higher than the updated value (K_m = 130µM). Referring back to the literature on GTP/ATP dependent PEPCK, $KHCO_3^-$ (or $NaHCO_3^-$) values are typically listed, but by estimating CO_2 concentrations (assuming pH 7.5), typically CO_2 K_m values would range between 0.4 and 2mM with many of the recordings for the low end being attributed to parasitic isozymes of PEPCK. One remarkable exception to this is the ATP dependent PEPCK from *Trypansoma cruzi* where the reported K_m of HCO_3^- is 25.7µM (CO_2 = 2.8µM).[21] When evaluating all of these K_m, k_{cat} and k_{cat}/K_m ratios, all evidence is in agreement that this enzyme is working unidirectionally to facilitate PEP carboxylation and CO_2 fixation, as opposed to nucleotide dependent PEPCKs bidirectional nature. A low K_m (and subsequent higher $k_{ca}t/K_m$ ratio) indicate the PEP and CO_2 binds quite tightly the enzyme, allowing flux through the enzyme during low concentration of both PEP and CO_2.

Table 4.5: Kinetic constants derived from PP_i-PEPCKs conversion of PEP to OAA. Substrate inhibition constants are bolded for PEP. All background rates have been subtracted.

$PEP + CO_2 + PO_4^- \rightarrow OAA + PP_i$

K_m (µM)			k_{cat} (s^{-1})			k_{cat}/K_m ($M^{-1}s^{-1}$)		
$MgCl_2$	$MnCl_2$	$FeCl_3$	$MgCl_2$	$MnCl_2$	$FeCl_3$	$MgCl_2$	$MnCl_2$	$FeCl_3$
5.2 ± 0.7	0.07 ± 0.03	26 ± 3.1	4.7 ± 0.02	6.3 ± 0.2	45 ± 1.5	9.0×10^5	9.0×10^7	1.7×10^6

K_m (µM)			k_{cat} (s^{-1})			k_{cat}/K_m ($M^{-1}s^{-1}$)		
PEP	PO_4^-	CO_2	PEP	PO_4^-	CO_2	PEP	PO_4^-	CO_2
74 ± 9.0	2300 ± 310	130 ± 15	63 ± 2.3	52 ± 2.7	47 ± 1.3	8.5×10^5	2.2×10^4	3.6×10^5
K_i = 6300 +/- 920								

4.3.13 OAA decarboxylation reaction

Although the flux through PP_i PEPCK is very likely to be through the PEP carboxylation reaction there is still some propensity for the decarboxylation reaction (Table 4.6). The previously reported K_m value of OAA is approximately 470µM, an order of magnitude higher than this new documented value (K_m = 42µM).[117] This new reported K_m value is similar to the K_m of rcPEPCK (K_m = 52µM).[39] The pyrophosphate K_m reported here (K_m = 5.5µM) is 40 times lower than the first account (K_m = 440µM).[117] The k_{cat}/K_m ratio reported indicate that *in vivo*, there will likely be some flux through the OAA decarboxylation reaction as the binding efficiency of both sets of substrates are similar.

Table 4.6: Kinetic constants derived from PP_i-PEPCKs conversion of OAA to PEP. Substrate inhibition constants are bolded under K_i for PP_i. Background rates have been subtracted.

$OAA + PP_i \rightarrow PEP + PO_4^- + CO_2$

K_m (µM)		k_{cat} (s^{-1})		k_{cat}/K_m ($M^{-1}s^{-1}$)	
OAA	PP_i	OAA	PP_i	OAA	PP_i
42 ± 9.8	5.5 ± 1.4 **K_i = 1300 ± 510**	4.9 ± 0.4	4.9 ± 0.4	1.5×10^6	1.4×10^5

4.3.14 Pyruvate kinase and OAA → pyruvate reactions

The initial biochemical characterization of PP_i PEPCK showed two specific reactions could be catalyzed by the enzyme while using PEP under different conditions. First, the PEP carboxylation reaction is the predominate reaction and this is in line with the new kinetic data presented here using the true metal cofactor of the enzyme, Fe^{2+}. The second, half-reaction occurs in the absence of CO_2 where PEP is initially dephosphorylated, and subsequently protonated to form pyruvate.[61] CO_2 was found to be a potent inhibitor of the pyrvuate kinase reaction ($K_i = 130\mu M$) suggesting that pyruvate is formed when the PEP conversion to OAA is arrested at some point in its reaction before carboxylation (Table 4.7). The pyruvate kinase reaction, appears to be stimulated under a non-reducing environment as well. As previously mentioned, GTP/ATP PEPCKs have been shown to form pyruvate although through the conversion of OAA to PEP, while there is no evidence of a pyruvate-kinase like reaction. This conversion can take place using the diphosphate nucleotide (half-reaction) to prevent final phosphorylation after decarboxylation, and there is also evidence that OAA + NTP can form pyruvate albeit at a much lower rate.[40]

Upon first look, the data showing pyruvate formation is in agreement with the previous studies in that pyruvate is formed (in both types of reactions) more readily under non-reducing conditions. For the pyruvate-kinase like reaction, the earlier work reported 1.5 and 1.8μmol of pyruvate produced per $min^{-1}mg^{-1}$ under reducing and non reducing conditions respectively. The reported turnover rates under non-reducing conditions are approximately 5 times higher ($8\mu mol in^{-1} mg^{-1}$) and on par under reducing conditions ($2.5\mu mol in^{-1} mg^{-1}$) from what has been previously reported.[61] The previously reported K_m values under non-reducing conditions for PEP and PO_4^- are 40μM and 660μM, respectively. The values reported here are very similar to those of the past and those of the PEP carboxylation reaction (Table 4.7).[61]

When comparing the formation of pyruvate from OAA under non reducing conditions, pyruvate accounts for 80% of the flux as the OAA to PEP as the k_{cat} (non-reducing) reported has not had the pyruvate formation reaction subtracted (Table 4.7). This shows that under non-reducing condition, OAA decarboxylation to PEP is drastically reduced. The data presented in the pyruvate formation from OAA under standard (reducing) assay conditions can be compared with rcPECK. This shows that all PEPCKs can produce pyruvate through a side reaction when OAA decarboxylation reaction is not completed. When comparing this side reaction k_{cat}, *Pf*PEPCK produces about 1 molecule of pyruvate per second, where rcPEPCK is producing 0.4 molecules per second. This pyruvate production in rcPEPCK has been attributed to inappropriate lid-reopening when the reactive enolate species after decarboxylation is produced, resulting in protonation by solvent. Although these mechanistic details are missing for PP_i PEPCKs, the pyruvate kinase reaction (from PEP to pyruvate) protonation was shown to be non-specific and achiral[119], suggesting a solvent effect similar to that of rcPEPCK.[135] The stimulation by non-reducing conditions is hard to interpret, as redox changes can have large consequences whether through metal, cysteine/methionine redox changes. Any of these consequences could be imagined to change the chemical reaction kinetics whether through chemical steps or dynamics of the enzyme itself.

Table 4.7: Kinetic constants derived from *Pf*PEPCKs pyruvate-kinase like reaction. Substrate inhibition was measured as shown bolded under the K_i for PEP.

PEP + PO_4^- → Pyruvate (Non-reducing)

K_m (µM)		k_{cat} (s^{-1})		k_{cat}/K_m ($M^{-1}s^{-1}$)	
PEP	PO_4^-	PEP	PO_4^-	PEP	PO_4^-
64 ± 4.7	1000 ± 240	8 ± 0.2	7.0 ± 0.6	1.3×10^5	7×10^3

PEP + PO_4^- → Pyruvate (Reducing)

K_m (µM)		k_{cat} (s^{-1})		k_{cat}/K_m ($M^{-1}s^{-1}$)	
PEP	PO_4^-	PEP	PO_4^-	PEP	PO_4^-
8.5 ± 1.3 **K_i = 3300 ± 2200**	160 ± 37	2.5 ± 0.02	1.5 ± 0.08	3.0×10^5	9.4×10^3

OAA + PP_i → Pyruvate (Non-reducing)

k_{cat} (s^{-1}) @500µM
OAA
2.0 ± 0.04

OAA + PP_i → Pyruvate (Reducing)

k_{cat} (s^{-1}) @500µM
OAA
1.2 ± 0.05

OAA + PP_i --> PEP + PO_4^- + CO_2 (Non-reducing)

k_{cat} (s^{-1}) @500uM
OAA
2.5 ± 0.05

4.3.15 Inhibition of PP_i PEPCK by analogues

The wealth of biochemical data for GTP/ATP PEPCKs has found many selective inhibitors of nucleotide dependent PEPCKs. Representative inhibitory mimics of all substrates, intermediates, and products have been previously characterized.[63] Surprisingly, due to the conservation of the active site residues, all of these previously characterized inhibitors were found to be ineffective inhibitors of *Pf*PEPCK with the exception of βSP which is an analogue of OAA (Table 4.8). It is unclear however, if this inhibition is due to a competitive mechanism or the dimerization previously shown by βSP (Table 4.3).[126] It is likely that in the OAA decarboxylating direction it is due to competitive effects. Unfortunately βSP can only be used to inhibit this reaction as it is also a substrate for malate dehydrogenase, the coupling enzyme used in the PEP carboxylation assay. Phosphonoformate, shown to be another OAA analogue in rcPEPCK (PDB ID# 2RK8), has an observed K_i of 2.4mM in PP_i PEPCK whereas against rcPEPCK its an order of magnitude more potent (K_i = 230μM).[63] Oxalate is a potent inhibitor of PEPCK (K_i = 89μM) against the PEP carboxylation reaction for GTP/ATP PEPCKs.[63] This potency is attributed to the understanding that these PEPCKs undergo a two step mechanism, creating an enolate intermediate which oxalate mimics. Interesting, it appears that oxalate does not inhibit PP_i PEPCK, indicating that the actual chemical mechanisms may be different the nucleotide and phosphate utilizing enzymes, or more likely, the dynamics changes and binding that must occur during turnover are different enough to weaken the binding of the enolate intermediate. This raises the possibility that the transitions this form of PEPCK undergo may be quite unique compared to those of the the nucleotide PEPCKs. Two small molecules have been structural and kinetically tested as PEP analogues, sulfoacetate and phosphoglycolate (sulfoacetate = PDB ID# 2RKE, phosphoglycolate = PDB ID# 2RKA).[63] When comparing both of these inhibitors in rcPEPCK, sulfoacetate is in the outer-shell geometry whereas phosphoglycolate is found in both this outer-shell geometry and directly interacting with the M1 metal, in a proposed catalytically productive state. This catalytically productive PEP-GDP complex has not been seen in previous studies, until now as shown in Figure 2.6 with the L153D mutant. When evaluating the inhibition of the PEP analogues between rc and *Pf*PEPCK, phosphoglycolate has the same K_i of 1mM while SA is much less potent against PP_i PEPCKs at a K_i of 3.6mM (vs 83μM against rcPEPCK). This may indicate a particular bias for how PEP binds to the active site of PP_i PEPCK, coordinating with the M1 metal rather than binding in a outer-shell in the release state. This observation is in line with the kinetic bias *Pf*PEPCK has for the PEP carboxylation reaction. 3-Mercaptopincolinic acid (MPA) was tested as it has been traditionally characterized as a selective hallmark inhibitor of GTP/ATP PEPCK. Recently, MPA inhibition was shown to be a mixed between a competitive component in the active site as it coordinated with the M1 Mn^{2+} and a non-competitive component as it was also found to be bind behind the P-loop in an allosteric pocket.[65] Based on the lack of inhibition by other PEPCK inhibitors, it was surprising to observe that upon first glance, MPA surprisingly was a potent inhibitor of *Pf*PEPCK. However, further analysis showing a lack of inhibition when using Mn^{2+} as the activating cation demonstrated that consistent with its being a known chelating agent of divalent cations, this inhibition was due to a non-specific chelation effect and not reversible active site binding. Finally, pyruvate was tested as an inhibitor of *Pf*PEPCK and it was observed that it did not inhibit. This is significant because this can inform us on the nature of the pyruvate-kinase reaction. The lack of binding of pyruvate, which would be expected to be seen (even weakly) if it was a product of an enzyme catalyzed reaction, and the non-stereospecific protonation leads one to believe that it is likely a solvent effect when the chemical reaction is not completed.

Table 4.8: Inhibition data for *Pf*PEPCK against the previously studied inhibitors of PEPCK. M1 metals are indicated in brackets, if not specified it is $FeCl_3$. N/a denotes that these inhibitors were tested but IC_{50} values could not be derived as they were such poor inhibitors. The plotted data can be found in S1-4.

IC_{50} PEP --> OAA

Inhibitor	IC_{50} (mM)	R^2	K_i (mM)
MPA ($FeCl_3$)	0.2	0.99	0.085
MPA ($MnCl_2$)	N/a	N/a	N/a
Malate ($FeCl_3$)	8.3	0.98	3.5
Malate ($MnCl_2$)	4.2	0.98	1.8
Pyruvate	N/a	N/a	N/a
Oxalate ($FeCl_3$)	4.2	0.98	1.8
Oxalate ($MnCl_2$)	2.9	0.98	1.2
Sulfoacetate	8.5	0.98	3.6
PGA	2.4	0.87	1
Phosphoformate	5.5	0.99	2.4

IC_{50} OAA --> PEP

Inhibitor	IC_{50} (mM)	R^2	K_i (mM)
B-sulfopyruvate	0.1	0.86	0.046
Malate	1.6	0.83	0.47

IC_{50} PEP-->> Pyruvate

Inhibitor	IC_{50} (mM)	R^2	K_i (mM)
CO_2	0.54	0.99	0.13
Malate	8.1	0.99	3.5

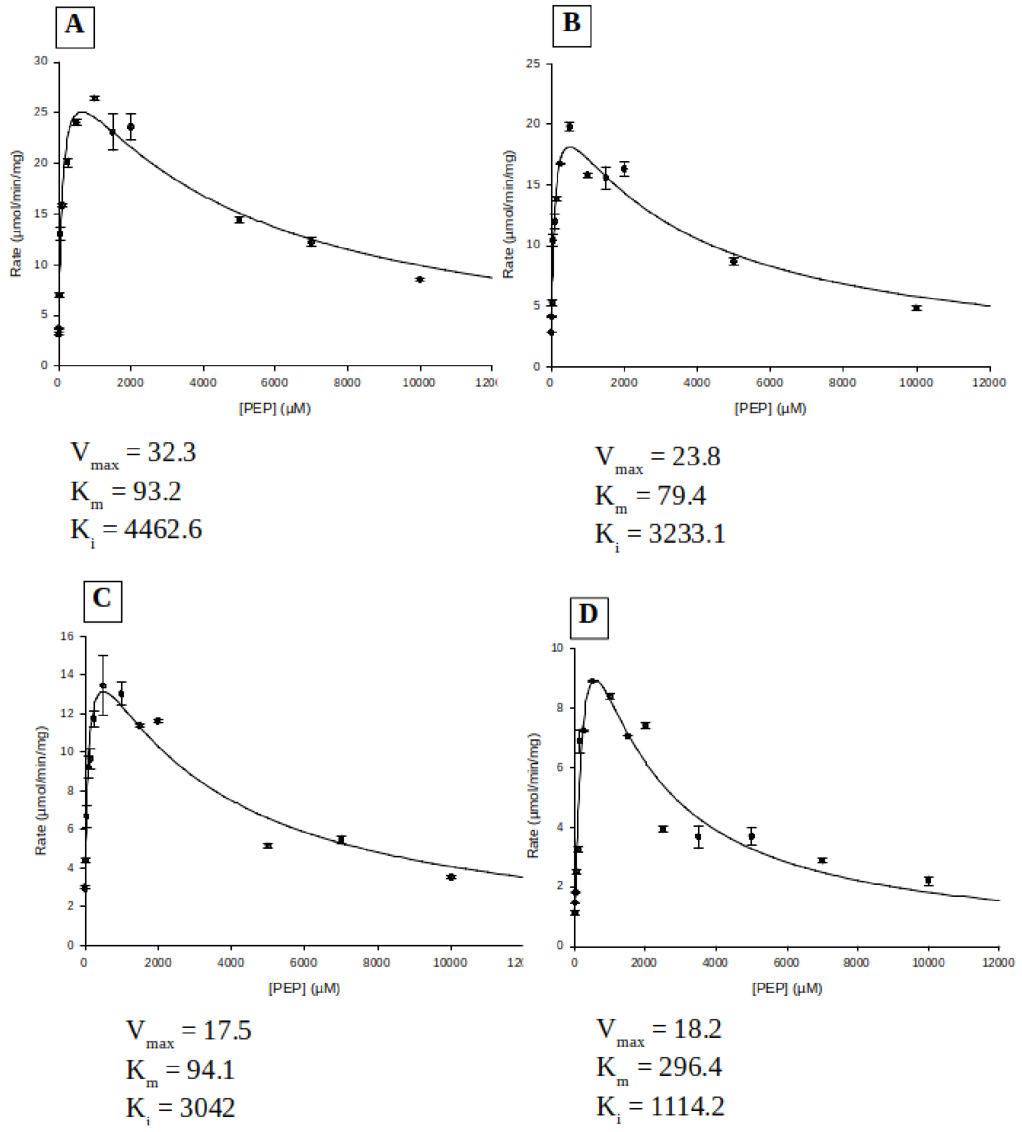
4.3.16 Substrate Inhibition by PEP and PP_i

Interestingly and in contrast to rcPEPCK and other nucleotide dependent PEPCKs, there is a substrate inhibition effect seen at high concentrations of either PEP (Table 4.5 and 4.7) or PP_i (Table 4.6)(see Figure 7.3.6 for progress curves). The only substrate inhibition previously noted in literature is with high concentrations of metals and bicarbonate, which may be due to a general salt effect or an allosteric effect at the newly discovered anion-binding allosteric pocket.[85,136] The PEP-induced substrate inhibition was observed for both the pyruvate kinase reaction and PEP carboxylation reaction. Therefore, the mechanism of inhibition must be altering both the carboxylation and pyruvate-kinase reactions, and with the size-exclusion quaternary state change data, suggests its origin is due to dimerization. In contrast, PP_i elicits an apparent substrate inhibition in the OAA decarboxylating direction which looks to be stronger than that of PEP (for the PEP carboxylating reaction) but does not lead to dimerization as in the case of PEP (Figure 4.2, Table 4.3). This suggests the PP_i substrate inhibition may be a true competitive mechanism in which it competes with the anionic OAA substrate, or alternatively PP_i may chelate the active site metals as proposed in the previous investigations.[61]

4.3.17 Inhibition of *Pf*PEPCK by malate

Due to the previous studies highlighting malate functioning as an inhibitor, and possibly a regulator *in vivo,* of PP_i PEPCK, specifically through the induction of dimerization, a more comprehensive look into the details of its potency was completed.[117,126] Malate has an IC_{50} value of 8.3mM (K_i of 3.5mM) (Table 4.8) which is a comparable to the results in the 1966 paper from Lochmuller, Wood, and Davis as determined by extrapolating their IC_{50} plots in Figure 9 of their publication.[117] This IC_{50} is very similar to the IC_{50} found in the pyruvate kinase reaction. The similarity in these two reactions is the use of PEP as the substrate. When contrasting this to the OAA decarboxylation reaction, the inhibition of malate is much stronger (IC_{50} = 1.6mM, $K_i = 470\mu M$). From previous kinetic data OAA does not cause substrate inhibition so the dimer is unlikely to affect OAA decarboxylation rates when compared to the monomer. A reasonable suggestion therefore is that malate can out-compete the weaker binding OAA and is therefore is observed to be more inhibitory in this direction.

As IC_{50} values and their subsequently derived K_i values can be unreliable, full inhibition progress curves were completed and the kinetic constants were re-plotted to gain more mechanistic insight into malate inhibition (Figure 4.9). What is readily noticed is that the addition of increasing amount of malate further enhances the apparent substrate inhibition effects seen through suppression of the k_{cat} (V_{max}). This substrate inhibition effect at the highest concentration of malate gave an aberrant trend in the re-plot (as seen in Figure 4.10), where the predominate effect switched from a k_{cat} to k_{cat}/K_m effect as the malate concentration increased. This aberrant trend is likely because the substrate inhibition has become so strong that is has suppressed the V_{max} so that the Michaelis-Menten parameters cannot be accurately inferred. Because of this, the K_i value found has left this last concentration out, leaving the same K_i for each k_{cat} and k_{cat}/K_m. This data suggests that the dimer both has a competitive as well as non-competitive effect. A possible mechanism based on prevents both PEP binding, as well as turnover of the enzyme, likely through the locking of a catalytically incompetent conformation.



A

$V_{max} = 32.3$
$K_m = 93.2$
$K_i = 4462.6$

B

$V_{max} = 23.8$
$K_m = 79.4$
$K_i = 3233.1$

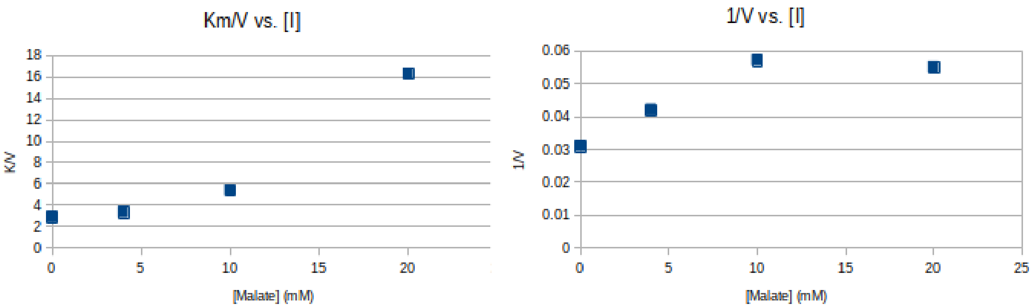
C

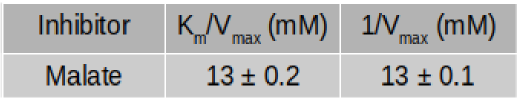
$V_{max} = 17.5$
$K_m = 94.1$
$K_i = 3042$

D

$V_{max} = 18.2$
$K_m = 296.4$
$K_i = 1114.2$

Figure 4.9: Individually progress curves fit the substrate inhibition model for malate inhibition against the PEP-forming reaction. Concentrations of malate are as follows A) 0mM, B) 4mM, C) 10mM, D) 20mM.





Inhibitor	K_m/V_{max} (mM)	$1/V_{max}$ (mM)
Malate	13 ± 0.2	13 ± 0.1

Figure 4.10: Replots from the malate inhibition progress curves. As highlighted above, the highest malate concentration point was left out of the calculation for the $1/V_{max}$ effect as the V_{max} substrate inhibition effect dominated, suppressing the true V_{max} in Figure 4.9.

4.4 Summary of findings

The initial biochemical characterizations that were completed in the late 60's, early 70's brought to light many interesting behaviours of the PP_i PEPCKs. With the wealth of information on GTP/ATP PEPCKs, a comparison was due to further understand the similarities and differences between these two classifications of PEPCK. This study answers the fundamental questions surrounding the basic functions of the enzyme, while also highlighting what differences there are between the two systems. This study supplies vital information to allow *Pf*PEPCK to be used as a model system for PP_i PEPCK.

Unfortunately, the crystallographic structures did not yield the information that was hoped for, as crystals obtained in both the presences or absences of malate yielded the inactive dimeric state of *Pf*PEPCK. Likely, the formation of the crystalline lattice is best accomplished when the monomers are coaxed into their dimeric form, leading to a low energy conformation. This does not rule out that other crystal forms, and co-crystallization experiments will not offer the holo-, substrate-, and product-bound complexes, but other structural tools might be more fortuitous where crystallization is not amenable. What can be concluded from this study in conjunction with the previous structural assessment is that the dimer structure is conserved across the two isozymes tested, the dimer is stabilized by two interfaces, and the domain-500 and a N-terminal helix are dynamic compared to the rest of the enzyme although the functional consequences of this conformational flexibility are unknown. It does appear that there is a coupling in the positional change in register of the N-terminal helix and the location of domain-500. Furthermore, the malate binding site is still unknown and without further structural data of the other complexes of the enzyme it is hard to say why the dimeric form is inhibitory. The tetrameric state of the enzyme is likely an artifact of the crystallization used to purify the enzyme in the initial studies, although confirmation of this is necessary to be sure. And finally, the mechanistic details of the confrontational changes through its reaction cycle are unknown.

This revisit of the biochemical aspects of PP_i PEPCK has allowed a few interesting conclusions, and further lines of questioning. First, the PP_i PEPCK M1 metal co-factor appears to be reduced Fe^{2+}, the first of its kind. This may be in-line with *Propionibacterium freudenreichii* oxygen sensitivity as a facultative anaerobe. Secondly, it has a low K_m and subsequently a very high k_{cat}/K_m for $CO_{2.,}$ suggesting an evolutionary adaptation as this enzymes host organism is known to be a carbon fixer through the use of PP_i PEPCK. This observation with the rest of the kinetic data suggests that PP_i PEPCK is a unidirectional enzyme functioning primarily in the carboxylation of PEP, where nucleotide utilizing PEPCKs are more-or-less bidirectional and primarily operate in OAA decarboxylation *in vivo*. Third, the supposed pyruvate-kinase reaction is in all likelihood in fact a non-enzymatic consequence of the PEP carboxylation not going to completion and the intermediate is then protonated by solvent to form pyruvate. Fourth, inhibition data suggests that the dynamic changes and the resultant free-energy costs through turnover may be significantly different for the PP_i class of PEPCK. This also opens the avenue for therapeutic targeting as humans do not have a PP_i PEPCK, whereas some pathogens (ex. *Entamoeba histolytica*) do. Fifth, the dimerization and subsequent inactivation previously shown from OAA and malate is in fact be the same mechanism causing substrate inhibition of PEP, while the substrate inhibition by PP_i is likely through a competition with substrates or a consequence of chelation of the M1 and M2 metals. Finally, based on this kinetic data, the dimer likely can still bind OAA for the decarboxylation reaction as there is no substrate inhibition of OAA, but binding and turnover of PEP is much harder as seen in the k_{cat}/K_m and k_{cat} effect.

In closing, this study has produced the most comprehensive biochemical study on a PP_i PEPCK. Having the initial kinetic and structural data in hand can now allow us to further probe the details of this new, but old, form of PEPCK. A comprehensive biochemical characterization of other PP_i utilizing

PEPCKs should be completed to see what characteristics are ubiquitous with this form of PEPCK or unique to this specific isozyme.

5.0 Concluding Remarks and Future Directions

The hysteretic mutant is a clear example of drastic modification of the free-energy landscape with only a single change. The observed structural results showed that this allosteric loop governs catalytically relevant global changes through dynamic switching between a solvent buried and solvent exposed conformation. In brief, by creating a more hydrophilic hinge loop this enzyme preferentially adopts a closed state, prematurely closing in particular circumstances and stalling the enzymes ability to function. This clear example adds to the growing host of literature regarding the importance of all areas of the enzyme in allowing enzymes to traverse through its free-energy landscape and catalytic cycle. These results suggests considering entire protein sequence/structure in our efforts to optimize and produce new enzymes. Furthermore, the structural data in conjunction with previous knowledge detailing the series of steps in this enzyme catalytic cycle offers a plausible explanation to the observed Michaelis-Menten parameters for this mutant. Although this work offers a proposal in regards to the cause of the observed hystersis, it still does not determine how this mutant PEPCK alleviates the hysteresis as it turns over. This re-establishment of the proper coupling of this hinge loop with global dynamics seems plausible but more evidence is required to be sure of these claims as well as offer a mechanism to how PEPCKs free-energy landscape is evolving. Further work in investigating the structural and functional changes between the inactive form, and the final activated form would be very insightful albeit extremely challenging. Kinetic observations of the Michaelis-Menten parameters determined are after the transition from inactive to active state would also offer insights into the functional consequences of perturbing this loop. It is by all means likely that the k_{cat}/K_m ratio would be different for each of the two structures. As PEP/βSP seemingly destabilized the C-terminal domains and CO_2 does not bind in the GDP-PEP structure, it suggests that the inactive form k_{cat}/K_m for all substrates is dramatically altered from those of the active form and what has been reported. Although these experiments would be very challenging, if completed would offer an exquisite example of the importance of the total dynamics of this model system.

PEPCKs thermal adaptation is probably the least surprising results presented, as all of the results are in line with the theories established within literature. Because of the pre-existing knowledge on this enzyme system, this study offers very specific reasoning into the functional changes that enzymes are pressured into adopting as they evolve to different thermal climes. Specifically, the kinetic data suggesting that at high temperatures, where the enzyme is still increasing in activity, enzyme dynamic perturbation leads not to a loss in substrate capture but rather its ability to efficiently traverse its free-energy landscape. With this, at sufficiently high enough temperatures when enzymes because reversibly inactivated, the data suggests that the entropically driven dynamics prevent substrate association. This data does not however suggest if this loss of association is due to local or global causes. The most interesting aspect of this study is the structural work detailing specific Ω-loop adaptations that allow for balance between open and closed states. This data detailing the mechanistic changes in the latch and hinge regions of PEPCK are quite obvious in hindsight (for example, less interactions are required at lower temperatures) but cement design rationale that could be used in protein design efforts. An interesting study may be swapping the residues responsible for the lid dynamics onto a different protein background, that is, taking the components of the mesophilic lid and placing them on the psychrophilic background. This may offer insight, much like in the hysteretic study, how global and local dynamics work together to produce an a functional optimal enzyme. A GTP-dependent thermophilic and hyperthermophilic variant would be a great addition in understanding some of these molecular details, and may lend insight into the wider area of dynamic adaptation with temperature.

The PP_i PEPCK study is a very significant step forward in the beginning of a research area that will better understand this isozyme. A great many outstanding questions from this work can offer an initial avenue for continuing these studies. First, cryoEM may be an opportunity to study the closed active form of the monomer, although the lower resolution may be problematic to see the nuances of the changes. Alternatively, careful co-crystal structures may also show the high resolution form of the active conformation. Second, the malate binding site is still unknown and a kinetic investigation of the potential 2-step mechanism proposed above may lend insight into this binding site. Third, structural characterizations of other isozymes, particularly those that may form trimers may be insightful into other ways this scaffold has adapted. Fourth, with this data presented here and with data found elsewhere[31], the N-terminal helix is clearly dynamic and variable across isozymes and may be important in the facilitation of the dimer interface. This dynamic interplay between dimer formation and this helix may also offer a path to understand why some PPi PEPCKs are active, and others are constitutively dimeric and inactive. This once again reaffirms the previous idea that as biochemists we must evaluate all parts of the enzyme, as distant areas of a protein influence each other. Finally, based on the inhibition data, there may be some exciting revelations in regards to the dynamic transitions this enzyme undergoes in contrast to the nucleotide counterparts. This data also suggests potential functional differences that could be exploited in structure-based drug design, if the molecular details are uncovered in how this form of PEPCK operates. It may be possible to create select inhibitors of pathogenic PP_i PEPCKs, while leaving human GTP dependent PEPCK relatively unaffected. This homage to the initial characterization of PPi PEPCKs will hopefully invigorate new studies to further complete the story of PEPCK, especially regarding this neglected form.

After so many years of studies on PEPCK, new and interesting results are still forthcoming, with many questions rising to the surface. These studies, along with all previous work, indicate that this enzyme is suitable in understanding the relationship between energetics, dynamics, and function. As more careful studies are completed using this enzyme, it may harken its addition into the pantheon of model systems that have been long studied to undercover the origins of function and dynamics of proteins.

References

1. Koshland DE. Application of a Theory of Enzyme Specificity to Protein Synthesis. *PNAS*. 1958;44(2):98-104.
2. Fischer E. Einfluss der Configuration auf die Wirkung der Enzyme. *Ber Dtsch Chem Ges*. 1894;27(192):2985-2993.
3. McCammon JA, Karplus M. Internal motions of antibody molecules. *Nature*. 1977;268:765-766. doi:10.1038/268765a0
4. Karplus M, McCammon JA. Molecular dynamics simulations of biomolecules. *Nat Struct Biol*. 2002;9(9):646-652. doi:10.1038/nsb0902-646
5. Frauenfelder H, Petsko GA. Structural dynamics of liganded myoglobin. *Biophys J*. 1980;32:465-483. doi:10.1016/S0006-3495(80)84984-8
6. Utter MF, Kurahashi K. Mechanism of action of Oxalacetic Carboxylase. *J Biol Chem*. 1954;207:821-841.
7. Utter MF, Kurahashi K, Rose IA. Some properties of oxalacetic carboxylase. 1954;207:803-819. doi:10.1007/978-94-015-8354-1_3
8. Utter MF, Kurahashi K. Purification of Oxalacetic Carboxylase from Chicken Liver. *J Biol Chem*. 1954;207:787-802.
9. Stark R, Guebre-Egziabher F, Zhao X, et al. A role for mitochondrial phosphoenolpyruvate carboxykinase (PEPCK-M) in the regulation of hepatic gluconeogenesis. *J Biol Chem*. 2014;289(11):7257-7263. doi:10.1074/jbc.C113.544759
10. Stark R, Kibbey RG. The mitochondrial isoform of phosphoenolpyruvate carboxykinase (PEPCK-M) and glucose homeostasis: Has it been overlooked? *Biochim Biophys Acta*. 2014;1840(4):1313-1330. doi:10.1016/j.bbagen.2013.10.033
11. Hanson RW. Thematic Minireview Series: A Perspective on the Biology of Phosphoenolpyruvate Carboxykinase 55 Years After Its Discovery. *J Biol Chem*. 2009;284(40):27021-27023. doi:10.1074/jbc.R109.040519
12. Montal ED, Dewi R, Bhalla K, et al. PEPCK Coordinates the Regulation of Central Carbon Metabolism to Promote Cancer Cell Growth. *Mol Cell*. 2015;60(4):571-583. doi:doi:10.1016/j.molcel.2015.09.025
13. Marrero J, Rhee KY, Schnappinger D, Pethe K, Ehrt S. Gluconeogenic carbon flow of tricarboxylic acid cycle intermediates is critical for Mycobacterium tuberculosis to establish and maintain infection. *Proc Natl Acad Sci U S A*. 2010;107(21):9819-9824. doi:10.1073/pnas.1000715107

14. Stark R, Pasquel F, Turcu A, et al. Phosphoenolpyruvate Cycling via Mitochondrial Phosphoenolpyruvate Carboxykinase Links Anaplerosis and Mitochondrial GTP with Insulin Secretion. *J Biol Chem*. 2009;284(39):26578-26590. doi:10.1074/jbc.M109.011775

15. Hakimi P, Yang J, Casadesus G, et al. Overexpression of the cytosolic form of phosphoenolpyruvate carboxykinase (GTP) in skeletal muscle repatterns energy metabolism in the mouse. *J Biol Chem*. 2007;282(45):32844-32855. doi:M706127200 [pii] 10.1074/jbc.M706127200

16. Yuan Y, Hakimi P, Kao C, et al. Reciprocal changes in phosphoenolpyruvate carboxykinase and pyruvate kinase with age are a determinant of aging in Caenorhabditis elegans. *J Biol Chem*. 2016;291(3):1307-1319. doi:10.1074/jbc.M115.691766

17. Das B, Tandon V, Saxena JK, Joshi S, Singh AR. Purification and characterization of phosphoenolpyruvate carboxykinase from Raillietina echinobothrida, a cestode parasite of the domestic fowl. *Parasitology*. 2013;140:136-146. doi:10.1017/S0031182012001254

18. Hidalgo J, Latorre P, Carrodeguas JA, Velázquez-Campoy A, Sancho J, López-Buesa P. Inhibition of Pig Phosphoenolpyruvate Carboxykinase Isoenzymes by 3-Mercaptopicolinic Acid and Novel Inhibitors. *PLoS One*. 2016;11(7):1-17. doi:10.1371/journal.pone.0159002

19. Hammond KD, Balinsky D. Kinetic studies on phosphoenolpyruvate carboxykinase purified from the mitochondrial and cytosol fractions of monkey liver. *Int J Biochem*. 1978;9:199-211. doi:10.1016/0020-711X(78)90150-7

20. Holten D, Nordalie R. Comparative Studies of Catalytic Properties of Guinea Pig Liver Intra- and Extramitochondrial Phosphoenolpyruvate Carboxykinase. *Biochemistry*. 1965;4(4):723-731.

21. Cymeryng C, Cazzulo JJ, Cannata JBJ. Phosphoenolpyruvate carboxykinase from Trypanosome cruzi. Purification and physicochemical and kinetic properties. *Mol Biochem Parasitol*. 1995;73:91-101.

22. Penny D., Kornecki E. Activities, Intracellular Localization and Kinetic Properties of Phosphoenolypyruvate Carboxykinase, Pyruvate Kinase, and Malate Dehydrogenase in Turtle (Pseudemys scripta elegans) Liver, Heart and Skeletal Muscles. *Comp Biochem Physiol*. 1973;46(B):405-415.

23. Jabalquinto AM, Cardemil E. The kinetic mechanism of yeast phosphoenolpyruvate carboxykinase. *Biochim Biophys Acta*. 1993;1161:85-90. doi:10.1016/0167-4838(93)90200-B

24. Case CL, Concar EM, Boswell KL, Mukhopadhyay B. Roles of Asp75, Asp78, and Glu83 of GTP-dependent phosphoenolpyruvate carboxykinase from Mycobacterium smegmatis. *J Biol Chem*. 2006;281(51):39262-39272. doi:10.1074/jbc.M602591200

25. Wilkes J, Cornish RA, Mettrick DF. Purification and properties of phosphoenolpyruvate carboxykinase from Ascaris suum. *Int J Parasitol*. 1982;12(2-3):163-171. doi:10.1016/0020-7519(82)90012-1

26. Podkovyrovi SM, Zeikus JG. Purification and characterization of phosphoenolpyruvate carboxykinase, a catabolic CO2-fixing enzyme, from Anaerobiospirillum succiniciproducens. *J Gen Microbiol*. 1993;139:223-228. doi:10.1099/00221287-139-2-223

27. Harlocker SL, Kapper MA, Greenwalt DE, Bishop SH. Phosphoenolpyruvate Carboxykinase from Ribbed Mussel Gill Tissue: Reactivity With Metal Ions, kinetics, and Action of 3 Mercaptopicolinic Acid. *J Exp Zool*. 1991;257:285-298. doi:10.1002/jez.1402570302

28. Holyoak T, Sullivan SM, Nowak T. Structural Insight into the Mechanism of PEPCK Catalysis. *Biochemistry*. 2006;45(27):8254-8263. doi:10.1021/bi060269g

29. Sullivan SM, Holyoak T. Structures of rat cytosolic PEPCK: Insight into the mechanism of phosphorylation and decarboxylation of oxaloacetic acid. *Biochemistry*. 2007;46(35):10078-10088. doi:10.1021/bi701038x

30. Dunten P, Belunis C, Crowther R, et al. Crystal Structure of Human Cytosolic Phosphoenolpyruvate Carboxykinase Reveals a New GTP-binding Site. *J Mol Biol*. 2002;316:257-264. doi:10.1006/jmbi.2001.5364

31. Chiba Y, Miyakawa T, Shimane Y, Takai K, Tanokura M, Nozaki T. Structural comparisons of phosphoenolpyruvate carboxykinases reveal the evolutionary trajectories of these phosphodiester energy conversion enzymes. *J Biol Chem*. 2019;294(50):19269-19278. doi:10.1074/jbc.RA119.010920

32. Aich S, Prasad L, Delbaere LTJ. Structure of a GTP-dependent Bacterial PEP-carboxykinase from Corynebacterium glutamicum. *Int J Biochem Cell Biol*. 2008;40:1597-1603. doi:10.1016/j.biocel.2007.12.002

33. Machová I, Snášel J, Dostál J, et al. Structural and Functional Studies of Phosphoenolpyruvate Carboxykinase from Mycobacterium tuberculosis. *PLoS One*. 2015;10(3):e0120682. doi:10.1371/journal.pone.0120682

34. Trapani S, Linss J, Goldenberg S, Fischer H, Craievich AF, Oliva G. Crystal structure of the Dimeric Phosphoenolpyruvate Carboxykinase (PEPCK) from Trypanosoma cruzi at 2 Å Resolution. *J Mol Biol*. 2001;313:1059-1072. doi:10.1006/jmbi.2001.5093

35. Sudom AM, Prasad L, Goldie H, Delbaere LTJ. The Phosphoryl-transfer Mechanism of Escherichia coli Phosphoenolpyruvate Carboxykinase From the Use of AlF3. *J Mol Biol*. 2001;314:83-92. doi:10.1006/jmbi.2001.5120

36. Cotelesage JJH, Prasad L, Zeikus JG, Laivenieks M, Delbaere LTJ. Crystal structure of Anaerobiospirillum succiniciproducens PEP carboxykinase reveals an important active site loop. *Int J Biochem Cell Biol*. 2005;37:1829-1837. doi:10.1016/j.biocel.2005.03.008

37. Cui DS, Broom A, McLeod MJ, Meiering EM, Holyoak T. Asymmetric Anchoring is Required for Efficient Ω-Loop Opening and Closing in Cytosolic Phosphoenolpyruvate Carboxykinase. *Biochemistry*. 2017;56(15):2106-2115. doi:10.1021/acs.biochem.7b00178

38. Johnson T, Holyoak T. Increasing the Conformational Entropy of the Omega-Loop Lid Domain in Phosphoenolpyruvate Carboxykinase Impairs Catalysis and Decreases Catalytic Fidelity. *Biochemistry*. 2010;49(25):5176-5187. doi:10.1021/bi100399e

39. Johnson TA, Holyoak T. The Ω-loop Lid Domain of Phosphoenolpyruvate Carboxykinase is Essential for Catalytic Function. *Biochemistry*. 2012;51(47):9547-9559. doi:10.1021/bi301278t

40. Johnson TA, Mcleod MJ, Holyoak T. Utilization of Substrate Intrinsic Binding Energy for Conformational Change and Catalytic Function in Phosphoenolpyruvate Carboxykinase. *Biochemistry*. 2016;55(3):575-587. doi:10.1021/acs.biochem.5b01215

41. Carlson GMGGM, Holyoak T. Structural insights into the mechanism of phosphoenolpyruvate carboxykinase catalysis. *J Biol Chem*. 2009;284(40):27037-27041. doi:10.1074/jbc.R109.040568

42. Xu D, Wang Z, Xia Y, et al. The gluconeogenic enzyme PCK1 phosphorylates INSIG1/2 for lipogenesis. *Nature*. 2020;580(7804):530-535. doi:10.1038/s41586-020-2183-2

43. Yang J, Kalhan SC, Hanson RW. What is the metabolic role of phosphoenolpyruvate carboxykinase? *J Biol Chem*. 2009;284(40):27025-27029. doi:10.1074/jbc.R109.040543

44. Burgess SC, He TT, Yan Z, et al. Cytosolic phosphoenolpyruvate carboxykinase does not solely control the rate of hepatic gluconeogenesis in the intact mouse liver. *Cell Metab*. 2007;5(4):313-320. doi:10.1016/j.cmet.2007.03.004

45. Lamers WH, Hanson RW, Meisner HM. cAMP stimulates transcription of the gene for cytosolic phosphoenolpyruvate carboxykinase in rat liver nuclei. *Proc Natl Acad Sci*. 1982;79(17):5137-5141. http://www.jstor.org/stable/12490.

46. Granner D, Andreone T, Sasaki K, Beale E. Inhibition of transcription of the phosphoenolpyruvate carboxykinase gene by insulin. *Nature*. 1983;305:549-551.

47. Hall RK, Wang XL, George L, Koch SR, Granner DK. Insulin represses phosphoenolpyruvate carboxykinase gene transcription by causing the rapid disruption of an active transcription complex: a potential epigenetic effect. *Mol Endocrinol*. 2007;21(2):550-563. doi:10.1210/me.2006-0307

48. Kaiser S, Curthoys NP. Effect of pH and bicarbonate on phosphoenolpyruvate carboxykinase and glutaminase mRNA levels in cultured renal epithelial cells. *J Biol Chem*. 1991;266(15):9397-9402.

49. Curthoys NP, Gstraunthaler G. Mechanism of increased renal gene expression during metabolic acidosis. *Am J Physiol Renal Physiol*. 2001;281(3):381-390. http://www.ncbi.nlm.nih.gov/pubmed/11502586.

50. Pollock AS, Long JA. The 5' region of the rat phosphoenolpyruvate carboxykinase gene confers pH sensitivity to chimeric genes expressed in renal and liver cell lines capable of expressing PEPCK. *Biochem Biophys Res Commun*. 1989;164(1):81-87.

51. Benvenisty N, Mencher D, Meyuhas O, Razin A, Reshef L. Sequential Changes in DNA Methylation Patterns of the Rat Phosphoenolpyruvate Carboxykinase Gene during Development. *Proc Natl Acad Sci*. 2016;82(2):267-271. http://www.jstor.org/stable/24610.

52. Jiang W, Wang S, Xiao M, et al. Acetylation regulates gluconeogenesis by promoting PEPCK1 degradation via recruiting the UBR5 ubiquitin ligase. *Mol Cell*. 2011;43(1):33-44. doi:10.1016/j.molcel.2011.04.028

53. Lin Y, Lu J, Zhang J, et al. Protein Acetylation Microarray Reveals NuA4 Controls Key Metabolic Target Regulating Gluconeogenesis. *Cell*. 2010;136(6):1073-1084. doi:10.1016/j.cell.2009.01.033.Protein

54. Knyphausen P, de Boor S, Kuhlmann N, et al. Insights into lysine-deacetylation of natively folded substrate proteins by sirtuins. *J Biol Chem*. 2016;291(28):14677-14694. doi:10.1074/jbc.M116.726307

55. Gao. XH, Ling L, Parisien M, et al. Discovery of a redox-thiol switch regulating cellular energy metabolism. *bioRxiv*. 2019;53(9):1689-1699. doi:10.1017/CBO9781107415324.004

56. Aich S, Delbaere LTJ. Phylogenetic study of the evolution of PEP-carboxykinase. *Evol Bioinforma*. 2007;3(306):333-340. doi:10.1177/117693430700300012

57. Mendler K, Chen H, Parks DH, Lobb B, Hug LA, Doxey AC. Annotree: Visualization and exploration of a functionally annotated microbial tree of life. *Nucleic Acids Res*. 2019;47(9):4442-4448. doi:10.1093/nar/gkz246

58. Chiba Y, Kamikawa R, Nakada-Tsukui K, Saito-Nakano Y, Nozaki T. Discovery of PPi-type phosphoenolpyruvate carboxykinase genes in Eukaryotes and Bacteria. *J Biol Chem*. 2015;290(39):23960-23970. doi:10.1074/jbc.M115.672907

59. Lolis E, Petsko GA. Crystallographic Analysis of the Complex between Triosephosphate Isomerase and 2-Phosphoglycolate at 2.5-Å Resolution: Implications for Catalysis. *Biochemistry*. 1990;29(28):6619-6625. doi:10.1021/bi00480a010

60. Joseph D, Petsko GA, Karplus M. Anatomy of a conformational change: Hinged “lid” motion of the triosephosphate isomerase loop. *Science (80-)*. 1990;249(4975):1425-1428. doi:10.1126/science.2402636

61. Davis JJ, Willard JM, Wood HG. Phosphoenolpyruvate carboxytransphosphorylase. III. Comparison of the fixation of carbon dioxide and the conversion of phosphoenolpyruvate and phosphate into pyruvate and pyrophosphate. *Biochemistry*. 1969;8(8):3127-3136. doi:10.1021/bi00836a001

62. Sullivan SM, Holyoak T. Enzymes with lid-gated active sites must operate by an induced fit mechanism instead of conformational selection. *Proc Natl Acad Sci U S A*. 2008;105(37):13829-13834.

63. Stiffin RM, Sullivan SM, Carlson GM, Holyoak T. Differential inhibition of cytosolic PEPCK by substrate analogues. Kinetic and structural characterization of inhibitor recognition. *Biochemistry*. 2008;47(7):2099-2109. doi:10.1021/bi7020662

64. Knorre D, Kudryashova N, Godovikova T. Chemical and Functional Aspects of Posttranslational Modification of Proteins. *Acta Naturae*. 2009;1(3):29-51. doi:10.32607/actanaturae.10755

65. Balan MD, McLeod MJ, Lotosky WR, Ghaly M, Holyoak T. Inhibition and Allosteric Regulation of Monomeric Phosphoenolpyruvate Carboxykinase by 3-Mercaptopicolinic Acid. *Biochemistry*. 2015;54(38):5878-5887. doi:10.1021/acs.biochem.5b00822

66. Verespy S, Mehta AY, Afosah D, Al-Horani RA, Desai UR. Allosteric Partial Inhibition of Monomeric Proteases. Sulfated Coumarins Induce Regulation, not just Inhibition, of Thrombin. *Sci Rep*. 2016;6:1-13. doi:10.1038/srep24043

67. Bellelli A. Hemoglobin and Cooperativity: Experiments and Theories. *Curr Protein Pept Sci*. 2010;11(1):2-36. doi:10.2174/138920310790274653

68. Frieden C. Slow Transitions in Enzymes. *Ann Rev Biochem*. 1979;48:471-489.

69. Cardenas ML, Rabajille E, Niemeyer H. Suppression of kinetic cooperativity of hexokinase D (glucokinase) by competitive inhibitors: A Slow transition model. *Eur J Biochem*. 1984;145(1):163-171. doi:10.1111/j.1432-1033.1984.tb08536.x

70. Ainslie Jr. GR, Shill, JP, Neet KE. Transients and Cooperativity: A Slow Transition Model for Relating Transients and Cooperative Kinetics of Enzymes. *J Biol Chem*. 1972;247(21):7088-7096.

71. Monod J, Wyman J, Changeux JP. On the nature of allosteric transitions: A plausible model. *J Mol Biol*. 1965;12(1):88-118. doi:10.1016/S0022-2836(65)80285-6

72. Porter CM, Miller BG. Cooperativity in monomeric enzymes with single ligand-binding sites. *Bioorg Chem*. 2012;43:44-50. doi:10.1016/j.bioorg.2011.11.001

73. Koshland DE, Nemethy JG, Filmer D. Comparison of Experimental Binding Data and Theoretical Models in Proteins Containing Subunits. *Biochemistry*. 1966;5(1):365-385. doi:10.1021/bi00865a047

74. Jiang Y, Li X, Morrow BR, et al. Single-Molecule Mechanistic Study of Enzyme Hysteresis. *ACS Cent Sci*. 2019;5(10):1691-1698. doi:10.1021/acscentsci.9b00718

75. Ricard J, Meunier JC, Buc J. Regulatory Behavior of Monomeric Enzymes: The Mnemonical Enzyme Concept. *Eur J Biochem*. 1974;49(1):195-208. doi:10.1111/j.1432-1033.1974.tb03825.x

76. Meunier JC, Buc J, Navarro A, Ricard J. Regulatory Behavior of Monomeric Enzymes: A Wheat-Germ Hexokinase as a Mnemonical Enzyme. *Eur J Biochem*. 1974;49(1):195-208. doi:10.1111/j.1432-1033.1974.tb03825.x

77. Hanozet G, Pircher HP, Vanni P, Oesch B, Semenza G. An example of enzyme hysteresis. The slow and tight interaction of some fully competitive inhibitors with small intestinal sucrase. *J Biol Chem*. 1981;256(8):3703-3711.

78. Studier FW. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif*. 2005;41(1):207-234.

79. Otwinowski Z, Minor W. Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzym*. 1997;276:307-326.

80. Vagin A, Teplyakov A. MOLREP: An Automated Program for Molecular Replacement. *J Appl Crystallogr*. 1997;30(6):1022-1025. doi:10.1107/S0021889897006766

81. 4 CCPN, Bailey S. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr Sect D Biol Crystallogr*. 1994;50(5):760-763. doi:10.1107/S0907444994003112

82. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Cryst*. 2004;D60:2126-2132.

83. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Cryst*. 1997;D53(3):240-255. doi:10.1107/S0907444996012255

84. Chen VB, Arendall WB, Headd JJ, et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Cryst*. 2010;66(Pt 1):12-21. doi:10.1107/S0907444909042073

85. Barwell S. Structural and kinetic characterization of mechanisms of regulation of phosphoenolpyruvate carboxykinase by anions. *UW Sp - Thesis*. 2019:1-94. doi:10.1017/CBO9781107415324.004

86. Fenimore PW, Frauenfelder H, McMahon BH, Parak FG. Slaving: Solvent fluctuations dominate protein dynamics and functions. *Proc Natl Acad Sci U S A*. 2002;99(25):16047-16051. doi:10.1073/pnas.212637899

87. Pucci F, Rooman M. Physical and molecular bases of protein thermal stability and cold adaptation. *Curr Opin Struct Biol*. 2017;42:117-128. doi:10.1016/j.sbi.2016.12.007

88. Arnold FH. Enzyme engineering reaches the boiling point. *Proc Natl Acad Sci U S A*. 1998;95(5):2035-2036. doi:10.1073/pnas.95.5.2035

89. Arnold FH, Wintrode PL, Miyazaki K, Gershenson A. How proteins adapt: Lessons from directed evolution. *Trends Biochem Sci*. 2009;26(2):100-106. doi:10.1101/sqb.2009.74.046

90. Somero GN. Temperature Adaptation of Enzymes: Biological Optimization Through Structure-Function Compromises. *Annu Rev Ecol Syst*. 1978;9(1):1-29. doi:10.1146/annurev.es.09.110178.000245

91. Závodszky P, Kardos J, Svingor Á, Petsko GA. Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc Natl Acad Sci U S A*. 1998;95(13):7406-7411. doi:10.1073/pnas.95.13.7406

92. Daniel RM, Danson MJ, Eisenthal R. The temperature optima of enzymes: A new perspective on an old phenomenon. *Trends Biochem Sci*. 2001;26(4):223-225. doi:10.1016/S0968-0004(01)01803-5

93. Daniel RM, Peterson ME, Danson MJ, et al. The molecular basis of the effect of temperature on enzyme activity. *Biochem J*. 2010;425(2):353-360. doi:10.1042/BJ20091254

94. Dias CL. Unifying microscopic mechanism for pressure and cold denaturations of proteins. *Phys Rev Lett*. 2012;109(4):1-6. doi:10.1103/PhysRevLett.109.048104

95. Dias CL, Ala-Nissila T, Karttunen M, Vattulainen I, Grant M. Microscopic mechanism for cold denaturation. *Phys Rev Lett*. 2008;100(11):1-4. doi:10.1103/PhysRevLett.100.118101

96. Dias CL, Ala-Nissila T, Wong-ekkabut J, Vattulainen I, Grant M, Karttunen M. The hydrophobic effect and its role in cold denaturation. *Cryobiology*. 2010;60(1):91-99. doi:10.1016/j.cryobiol.2009.07.005

97. Ben-Naim A. Theory of cold denaturation of proteins. *Adv Biol Chem*. 2013;03(01):29-39. doi:10.4236/abc.2013.31005

98. Feller G. Protein stability and enzyme activity at extreme biological temperatures. *J Phys Condens Matter*. 2010;22(32). doi:10.1088/0953-8984/22/32/323101

99. Low PS, Bada JL, Somero GN. Temperature adaptation of enzymes: roles of the free energy, the enthalpy, and the entropy of activation. *Proc Natl Acad Sci U S A*. 1973;70(2):430-432. doi:10.1073/pnas.70.2.430

100. Jeon CO, Park W, Ghiorse WC, Madsen EL. Polaromonas naphthalenivorans sp. nov., a naphthalene-degrading bacterium from naphthalene-contaminated sediment. *Int J Syst Evol Microbiol*. 2004;54(1):93-97. doi:10.1099/ijs.0.02636-0

101. Freier D, Mothershed CP, Wiegel J. Characterization of Clostridium thermocellum JW20. *Appl Environ Microbiol*. 1988;54(1):204-211. doi:10.1128/aem.54.1.204-211.1988

102. Segur JB, Oberstar HE. Viscosity of Glycerol and Its Aqueous Solutions. *Ind Eng Chem*. 1951;43(9):2117-2120. doi:10.1021/ie50501a040

103. Tsai CS. Spontaneous decarboxylation of oxalacetic acid. *Can J Chem*. 1967;46(4):873-890. doi:10.1139/v68-111

104. Vieille C, Zeikus GJ. Hyperthermophilic Enzymes: Sources, Uses, and Molecular Mechanisms for Thermostability. *Microbiol Mol Biol Rev*. 2001;65(1):1-43. doi:10.1128/MMBR.65.1.1-43.2001

105. Han MH. Non-linear Arrhenius plots in temperature-dependent kinetic studies of enzyme reactions. I. Single transition processes. *J Theor Biol*. 1972;35:543-568. doi:10.1016/0022-5193(72)90150-6

106. Beece D, Eisenstein L, Frauenfelder H, et al. Solvent Viscosity and Protein Dynamics. *Biochemistry*. 1980;19(23):5147-5157. doi:10.1021/bi00564a001

107. Fields PA, Somero GN. Hot Spots In Cold Adaptation: Localized Increasing in Conformational Flexibility in Lacatate Dehydrogenase A4 Orthologs of Antarctic Notothenioid Fishes. *PNAS*. 1998;95(19):11476-11481. https://www.jstor.org/stable/46333 Accessed:

108. Somero GN. Proteins and Temperature. *Annu Rev Biochem*. 1995;57:43-68. doi:10.1016/B978-0-12-374553-8.00192-1

109. Bjelic S, Brandsdal BO, Åqvist J. Cold adaptation of enzyme reaction rates. *Biochemistry*. 2008;47(38):10049-10057. doi:10.1021/bi801177k

110. Isaksen GV, Åqvist J, Brandsdal BO. Protein Surface Softness Is the Origin of Enzyme Cold-Adaptation of Trypsin. *PLoS Comput Biol*. 2014;10(8). doi:10.1371/journal.pcbi.1003813

111. Privalov PL, Gill SJ. Stability of Protein Structure and Hydrophobic Interaction. *Adv Protein Chem*. 1988;39:191-234. doi:10.1017/CBO9781107415324.004

112. Lonhienne T, Gerday C, Feller G. Psychrophilic enzymes: Revisiting the thermodynamic parameters of activation may explain local flexibility. *Biochim Biophys Acta*. 2000;1543(1):1-10. doi:10.1016/S0167-4838(00)00210-7

113. Miyazaki K, Wintrode PL, Grayling RA, Rubingh DN, Arnold FH. Directed evolution study of temperature adaptation in a psychrophilic enzyme. *J Mol Biol*. 2000;297(4):1015-1026. doi:10.1006/jmbi.2000.3612

114. Zhao H, Arnold FH. Directed evolution converts subtilisin E into a functional equivalent of thermitase. *Protein Eng*. 1999;12(1):47-53. doi:10.1093/protein/12.1.47

115. Gershenson A, Schauerte JA, Giver L, Arnold FH. Tryptophan phosphorescence study of enzyme flexibility and unfolding in laboratory-evolved thermostable esterases. *Biochemistry*. 2000;39(16):4658-4665. doi:10.1021/bi992473s

116. Cooper TG, Tchen TT, Wood HG, Benedict CR. The Carboxylation of Phosphoenolpyruvate and Pyruvate: The Active Species of CO2 Utilized by Phosphoenolpyruvate Carboxykinase, Carboxytransphosphorylase, and Pyruvate Carboxylase. *J Biol Chem*. 1968;243(14):3857-3863.

117. Lochmuller, H., Wood, HM., Davis J. Phosphoenolpyruvate Carboxytransphosphorylase: Crystallization and Properties. *J Biol Chem*. 1966;241(23):5678-5691.

118. Willard JM, Rose IA. Formation of Enolpyruvate in the Phosphoenolpyruvate Carboxytransphosphorylase Reaction. *Biochemistry*. 1973;12(26):5241-5246. doi:10.1021/bi00750a003

119. O’Brien WE, Singleton R, Wood HG. Phosphoenolpyruvate Carboxytransphosphorylase. An Investigation of the Mechanism with 18O. *Biochemistry*. 1973;12(26):5247-5253. doi:10.1021/bi00750a004

120. O’Brien WE, Wood HG. Carboxytransphosphorylase: VIII. Ligand-Mediated Interaction of Subunits as a Possible Control Mechanism in Propionibacteria. *J Biol Chem*. 1974;249(15):4917-4925.

121. Willard JM, Davis JJ, Wood HG. Phosphoenolpyruvate carboxytransphosphorylase. IV. Requirement for metal cations. *Biochemistry*. 1969;8(8):3137-3144. doi:10.1021/bi00836a002

122. Wood, H., Davis, J., Willard. J, Wood HG, Davis JJ, Willard JM, Wood, H., Davis, J., Willard. J. Phosphoenolpyruvate carboxytransphosphorylase. V. Mechanism of the reaction and the role of metal ions. *Biochemistry*. 1969;8(8):3145-3155. doi:10.1021/bi00836a003

123. Hunt M, Kohler P. Purification and characterization of phosphoenolpyruvate carboxykinase from Trypanosma brucei. *Biochim Biophys Acta*. 1994;(1249):15-22.

124. Tortora P, Hanozet GM, Guerritore A. Purification of phosphoenolpyruvate carboxykinase from Saccharomyces cerevisiae and its use for bicarbonate assay. *Anal Biochem*. 1985;144(1):179-185. doi:10.1016/0003-2697(85)90101-0

125. Burnell J. Purification and Properties of Phosphoenolpyruvate Carboxykinase from C4 Plants. *Funct Plant Biol*. 1986;13(5):577. doi:10.1071/pp9860577

126. Haberland ME, Willard JM, Wood HHG. Phosphoenolpyruvate Carboxytransphosphorylase. Study of the Catalytic and Physical Structures. *Biochemistry*. 1972;11(5):712-722. doi:10.1021/bi00755a007

127. Wood HG, Stjernholm RL, Leaver-Fay A. The Metabolism of Labeled Glucose by the Propionic Acid Bacteria. *J Bacteriol*. 1955;70(5):510-520.

128. Patrick M, Sitj L, Wood HG. Phosphoenolpyruvic Carboxytransphosphorylase, A CO2 Fixation Enzyme From Propionic Acid Bacteria. *J Biol Chem*. 1962;237(10):3044-3051.

129. Siu M, Wood HG, Stjernholm RL. Fixation of CO2 by Phosphoenolpyruvic Carboxytransphosphorylase. *J Biol Chem*. 1961;236(4):21-22.

130. Van Eunen K, Bouwman J, Daran-Lapujade P, et al. Measuring enzyme activities under standardized in vivo-like conditions for systems biology. *FEBS J*. 2010;277(3):749-760. doi:10.1111/j.1742-4658.2009.07524.x

131. Greenfield NJ, Hussain M, Lenard J. Effects of growth state and amines on cytoplasmic and vacuolar pH, phosphate and polyphosphate levels in Saccharomyces cerevisiae: A 31P-nuclear magnetic resonance study. *Biochim Biophys Acta*. 1987;926:205-214.

132. Ingerson-Mahar M, Briegel A, Werner J, Jensen G, Gitai Z. The metabolic enzyme CTP synthase froms cytoskeletal filaments. *Nat Cell Biol*. 2010;12(8):739-746. doi:10.1038/ncb2087.

133. den Hollander JA, Shulman RG, Ugurbil K, Brown TR. Phosphorus-31 Nuclear Magnetic Resonance Studies of the Effect of Oxygen upon Glycolysis in Yeast. *Biochemistry*. 1981;20(20):5871-5880. doi:10.1021/bi00523a034

134. Holyoak T, Nowak T, Carboxykinase P, Holyoak T, Nowak T. pH dependence of the reaction catalyzed by avian mitochondrial phosphoenolpyruvate carboxykinase. *Biochemistry*. 2004;43(22):7054-7065. doi:10.1021/bi049707e

135. Falentin H, Deutsch SM, Jan G, et al. The Complete Genome of Propionibacterium freudenreichii CIRM-BIA1T, a Hardy Actinobacterium with Food and probiotic applications. *PLoS One*. 2010;5(7):1-12. doi:10.1371/journal.pone.0011748

136. Case CL, Mukhopadhyay B. Kinetic characterization of recombinant human cytosolic phosphoenolpyruvate carboxykinase with and without a His10-tag. *Biochim Biophys Acta - Gen Subj*. 2007;1770(11):1576-1584. doi:10.1016/j.bbagen.2007.07.012

137. McCluskey GD, Bearne SL. “Pinching” the ammonia tunnel of CTP synthase unveils coordinated catalytic and allosteric-dependent control of ammonia passage. *Biochim Biophys Acta - Gen Subj*. 2018;1862(12):2714-2727. doi:10.1016/j.bbagen.2018.08.008

138. Levitzki A. Cytidine Triphosphate Synthetase. Covalent Intermediates and Mechanisms of Action. *Biochemistry*. 1971;10(18):3365-3371. doi:10.1021/bi00794a008

139. MacDonnell JE, Lunn FA, Bearne SL. Inhibition of E. coli CTP synthase by the “positive” allosteric effector GTP. *Biochim Biophys Acta - Proteins Proteomics*. 2004;1699(1-2):213-220. doi:10.1016/j.bbapap.2004.03.002

140. Levitzkif A, Koshland DE. Role of an Allosteric Effector. Guanosine Triphosphate Activation in Cytosme Triphosphate Synthetase. *Biochemistry*. 1972;11(2):241-246. doi:10.1021/bi00752a015

141. Levitzkit A, Koshland DE. Ligand-Induced Dimer-to-Tetramer Transformation in Cytosine Triphosphate Synthetase. *Biochemistry*. 1972;11(2):247-253. doi:10.1021/bi00752a016

142. Lynch EM, Hicks DR, Shepherd M, et al. Human CTP synthase filament structure reveals the active enzyme conformation. *Nat Struct Mol Biol*. 2017;24(6):507-514. doi:10.1038/nsmb.3407

143. Endrizzi JA, Kim H, Anderson PM, Baldwin EP. Mechanisms of product feedback regulation and drug resistance in cytidine triphosphate synthetases from the structure of a CTP-inhibited complex. *Biochemistry*. 2005;44(41):13491-13499. doi:10.1021/bi051282o

144. Endrizzi JA, Kim H, Anderson PM, Baldwin EP. Crystal structure of Escherichia coli cytidine triphosphate synthetase, a nucleotide-regulated glutamine amidotransferase/ATP-dependent amidoligase fusion protein and homologue of anticancer and antiparasitic drug targets. *Biochemistry*. 2004;43(21):6447-6463. doi:10.1021/bi0496945

145. Fijolek A, Hofer A, Thelander L. Expression, purification, characterization, and in vivo targeting of trypanosome CTP synthetase for treatment of African sleeping sickness. *J Biol Chem*. 2007;282(16):11858-11865. doi:10.1074/jbc.M611580200

146. Meza E, Becker J, Bolivar F, Gosset G, Wittmann C. Consequences of phosphoenolpyruvate:sugar phosphotranferase system and pyruvate kinase isozymes inactivation in central carbon metabolism flux distribution in Escherichia coli. *Microb Cell Fact*. 2012;11(1):127. doi:10.1186/1475-2859-11-127

147. Toschi L, Finocchiaro G, Bartolini S, Gioia V, Cappuzzo F. Role of gemcitabine in cancer therapy. *Futur Oncol*. 2005;1(1):7-17. doi:10.1517/14796694.1.1.7

148. Wang MC, Bloch A, McPartland RP, Weinfeld H. Cytidine 5’Synthetase as a Target for Inhibition by the Antitumor Agent 3-Deazauridine. *Cancer Res*. 1974;34(11):3107-3111.

149. Gharehbaghi K, Zhen W, Fritzer-Szekeres M, Szekeres T, Jayaram H. Studies on the Antitumor Activity and Biochemical Actions of Cyclopentenyl. *Life Sci*. 1999;64(2):103-112.

150. Schimmel K. *The Cytotoxic Drug Cyclo-Pentenyl Cytosine: From Manufacturing to Anti-Tumor Activity and Cardio Toxicity*.; 2007.

151. Zhang H, Cooney DA, Zhang MH, Ahluwalia G, Ford H, Johns DG. Resistance to Cyclopentenylcytosine in Murine Leukemia L1210 Cells. *Cancer Res*. 1993;53(23):5714-5720.

152. Blaney SM, Grem JL, Balis FM, Cole DE, Adamson PC, Poplack DG. Mechanism of resistance to cyclopentenyl cytosine (CPE-C) in Molt-4 lymphoblasts. *Biochem Pharmacol*. 1993;45(7):1493-1501. doi:10.1016/0006-2952(93)90050-7

153. Whelan J, Phear G, Yamauchi M, Meuth M. Clustered base substitutions in CTP synthetase conferring drug resistance in Chinese hamster ovary cells. *Nat Genet*. 1993;3(4):317-322. doi:10.1038/ng0493-317

154. McCluskey GD, Mohamady S, Taylor SD, Bearne SL. Exploring the Potent Inhibition of CTP Synthase by Gemcitabine-5′-Triphosphate. *ChemBioChem*. 2016;17(23):2240-2249. doi:10.1002/cbic.201600405

155. McCluskey GD, Bearne SL. Biophysical Analysis of Bacterial CTP Synthase Filaments Formed in the Presence of the Chemotherapeutic Metabolite Gemcitabine-5′-triphosphate. *J Mol Biol*. 2018;430(8):1201-1217. doi:10.1016/j.jmb.2018.02.019

156. Bearne SL, Hekmat O, MacDonnell JE. Inhibition of Escherichia coli CTP synthase by glutamate γ-semialdehyde and the role of the allosteric effector GTP in glutamine hydrolysis. *Biochem J*. 2001;356(1):223-232. doi:10.1042/0264-6021:3560223

157. Holyoak T, Fenn TD, Wilson MA, Moulin AG, Ringe D, Petsko GA. Malonate: A versatile cryoprotectant and stabilizing solution for salt-grown macromolecular crystals. *Acta Crystallogr - Sect D Biol Crystallogr*. 2003;59(12):2356-2358. doi:10.1107/S0907444903021784

158. Biffinger JC, Kim HW, DiMagno SG. The polar hydrophobicity of fluorinated compounds. *ChemBioChem*. 2004;5(5):622-627. doi:10.1002/cbic.200300910

159. Lynch EM, Kollman JM. Coupled structural transitions enable highly cooperative regulation of human CTPS2 filaments. *Nat Struct Mol Biol*. 2020;27(1):42-48. doi:10.1038/s41594-019-0352-5

160. Dalvi VH, Rossky PJ. Molecular origins of fluorocarbon hydrophobicity. *Proc Natl Acad Sci U S A*. 2010;107(31):13603-13607. doi:10.1073/pnas.0915169107

161. Wylie JL, Wang LL, Tipples G, McClarty G. A single point mutation in CTP synthetase of Chlamydia trachomatis confers resistance to cyclopentenyl cytosine. *J Biol Chem*. 1996;271(26):15393-15400. doi:10.1074/jbc.271.26.15393

Appendices

7.1 Materials

GDP was purchased from Combi-Blocks. TCEP and DTT were purchased from GoldBio. GTP, PEP, and PP_i were purchased from ChemImpex. PK, LDH, and MDH were purchased from Calzyme. OAA was purchased from Calbiochem. Nickel-NTA was purchased from Qiagen. P6-DG desalting resin was purchase from BioRad. All other chemicals were purchased from the highest grade available.

7.2 Supplemental Information for Thermal Adaptions of PEPCK

Cloned Sequence for *Polaramonas naphthalenivorans* PEPCK

ATGAACCAGCCGGCGATGCAAGGTCTGAACCTGAACACCCCGGCGTACGTTAAGAACGCG
AAACTGATCGCGTGGGTTGCGGACATGGTGGCGCTGTGCAAGCCGGATACCGTTTACTGGT
GCGACGGTAGCGATGCGGAGTATCAGCGTCTGTGCCAGCAACTGGTGGACGCGGGCACCT
TTAAGAAACTGAACGAAGCGAAGCGTCCGAACAGCTACCTGGCGTGCAGCGACCCGAGCG
ATGTGGCGCGTGTTGAGGATCGTACCTATATTTGCAGCGCGAAAGAGGAAAACGCGGGTCC
GACCAACAACTGGATGGAGCCGAGCGAAATGCGTGCGACCCTGCAACCGCTGTTTGACGG
CTGCATGCGTGGTCGTACCATGTACGTGGTTCCGTTTAGCATGGGTCCGCTGGGTAGCCCGA
TTGCGCACATCGGTATTGAACTGAGCGATAGCCCGTATGTGGCGGTTAACATGAAGATTATG
ACCCGTATGGGCAAAGCGGTGTTCGAGGTTCTGGGCACCGAGGGTGACTTTGTGCCGTGCG
TTCACACCGTGGGTGCGCCGCTGGCGGAGGGTGAACAGGATGTTACCTGGCCGTGCAACA
AGACCAAATACATCGTGCACTATCCGGAAACCCGTGAGATTTGGAGCTACGGTAGCGGCTAT
GGTGGCAACGCGCTGCTGGGTAAGAAATGCTTCGCGCTGCGTATCGCGAGCAACATGGGCC
GTGACGAGGGTTGGCTGGCGGAACACATGCTGATTCTGGGCGTTACCAACCCGGAGGGTA
AGAAATACCATGTGGCGGCGGCGTTCCCGAGCGCGTGCGGCAAGACCAACTTTAGCATGCT
GGTTCCGCCGGCGGGTTTTGATGGTTGGAGCGTGACCACCATCGGCGACGATATCGCGTGG
ATTAAGCCGGCGGCGGATGGTAAACTGTACGCGATTAACCCGGAAGCGGGCTATTTTGGTG
TTGCGCCGGGTACCAACATGCTGACCAACCCGAACTGCATGCTGAGCGTGCGTGGCAACAC
CATCTTCACCAACGTTGCGCTGACCGACGATGGCGACGTGTGGTGGGAGGGTATGGAACAA
GATGCGCCGGGCAAGGCGCTGCCGGACCACCTGATCGATTGGCAGGGTAAAGACTGGACC
CCGCAAATTGCGAAGGAAACCGGTGCGAAAGCGGCGCACCCGAACAGCCGTTTTACCGTT
GCGGCGACCAACAACCCGGCGCTGGACAGCGCGTGGGATGATCCGAAAGGTGTGGCGATC
GATGCGTTCATTTTTGGTGGCCGTCGTAGCACCACCGTTCCGCTGGTGACCGAGGCGCGTA
GCTGGGTTGAAGGCGTGTACATGGCGGCGACCATGGGTAGCGAAACCACCGCGGCGGCGA
CCGGTCAGGCGGGTGTGGTTCGTCGTGACCCGTTCGCGATGCTGCCGTTTATGGGTTACAA
CATGAGCGACTATTTCCAACACTGGCTGGATCTGGGCAAGAAACTGGCGGCGAGCGGTGC
GACCCTGCCGAAGATCTACACCACCAACTGGTTCCGTAAGGGCACCGATGGTAAATTTGTT
TGGCCGGGTTATGGCGAGAACATGCGTGTGCTGGCGTGGATGATCGACCGTATTGAGGGTA
AAGGCCAGGGTGAAGATCACGTTACCGGTATCAGCCCGAGCTATGAAGACCTGAACTGGAC
CGGCCTGGACTTCAGCGCGGATCAGTTTAACACCGTGACCAGCATTGATAAGGCGGCGTGG
CTGAAAGAGCTGGAACTGCACGCGGAGCTGTTCAAACAACTGGAACACCACCTGCCGAAG
GAGCTGGGTGAAACCAAAGCGGAGATTGAACGTCGTCTGGCGGCGTAA

Cloned Sequence for *Hungateiclostridium thermocellum* PEPCK

ATGACCAGCACCAACATGACCAAGAACAAGAAACTGCTGGACTGGGTGAAAGAAATGGCG
GAGATGTGCCAGCCGGACGAGATCTACTGGTGCGATGGTAGCGAGGAAGAGAACGAACGT
CTGATTAAGCTGATGGTTGATAGCGGCCTGGCGACCCCGCTGAACCCGGAGAAACGTCCGG
GTTGCTACCTGTTCCGTAGCGACCCGAGCGATGTTGCGCGTGTGGAAGATCGTACCTTTATC
GCGAGCAAGACCAAAGAGGACGCGGGTCCGACCAACAACTGGATTGATCCGGTGGAACTG
AAGGCGACCATGAAAGAGCTGTACAAGGGTTGCATGAAAGGCCGTACCATGTATGTGATCC
CGTTCAGCATGGGTCCGATTGGCAGCCCGATCAGCAAGATTGGTGTTGAGCTGACCGACAG
CCCGTACGTGGTTGTGAACATGCGTATCATGACCCGTATTGGCAAAGCGGTTCTGGATCAAC
TGGGTGAAGACGGCGATTTTGTTCCGTGCCTGCACAGCGTGGGTGCGCCGCTGAAGGAAG
GCGAGAAGGACAAAGGTTGGCCGTGCGCGCCGATCGAGAAGAAATATATTAGCCACTTCCC
GGAAGAGCGTACCATCTGGAGCTACGGTAGCGGCTATGGTGGCAACGCGCTGCTGGGTAAG
AAATGCTTTGCGCTGCGTATTGCGAGCGTGATGGCGCGTGACGAAGGTTGGCTGGCGGAGC
ACATGCTGATCCTGCGTATTACCGATCCGGAAGGCAACAAGACCTATGTGACCGGTGCGTTC
CCGAGCGCGTGCGGCAAAACCAACCTGGCGATGCTGATCCCGACCATTCCGGGTTGGAAG
GTTGAAACCATCGGCGACGATATTGCGTGGATGCGTTTTGGTAAAGACGGCCGTCTGTATGC
GATCAACCCGGAAGCGGGTTTCTTTGGCGTTGCGCCGGGTACCAGCATGGATAGCAACCCG
AACGCGATGCACACCATCAAGAAAAACACCATTTTCACCAACGTTGCGCTGACCGACGATG
GCGACGTGTGGTGGGAAGGTATTGGTACCGAGCCGCCGGCGCACCTGATTGATTGGCAGGG
TAAAGACTGGACCCCGGATAGCGGTACCCTGGCGGCGCACCCGAACGGTCGTTTTACCGCG
CCGGCGAGCCAATGCCCGGTTATTGCGCCGGAATGGGAGGACCCGGAGGGTGTTCCGATTA
GCGCGATCCTGATTGGTGGCCGTCGTCCGAACACCATCCCGCTGGTTCACGAGAGCTTCGA
TTGGAACCACGGTGTGTTTATGGGCAGCATCATGGGTAGCGAAATTACCGCGGCGGCGATC
AGCAACAAGATTGGCCAAGTGCGTCGTGACCCGTTCGCGATGCTGCCGTTTATCGGTTACA
ACGTTAACGACTATCTGCAACACTGGCTGAACATGGGCACCAAAACCGATCCGAGCAAGCT
GCCGAAAATTTTCTACGTGAACTGGTTTCGTAAGGATAGCAACGGTAAATGGCTGTGGCCG
GGTTATGGCGAAAACAGCCGTGTTCTGAAGTGGATCGTTGAGCGTGTGAACGGCAAGGGC
AAAGCGGTTAAAACCCCGATCGGCTACATGCCGACCGTGGACGCGATTGATACCACCGGTC
TGGACGTTAGCAAGGAAGATATGGAAGAGCTGCTGAGCGTTAACAAAGAGCAGTGGCTGC
AAGAAGTGGAGAGCATCAAGGAACACTACAAAAGCTATGGTGAAAAGCTGCCGAAAGAG
CTGTGGGCGCAGCTGGAAGCGCTGGAGCAACGTCTGAAGGAATACAACGGCTAA

7.3 Supplemental data for Biochemical, Structural, and Kinetic Characterization of PP$_i$-PEPCK from *Propionibacterium freudenreichii*

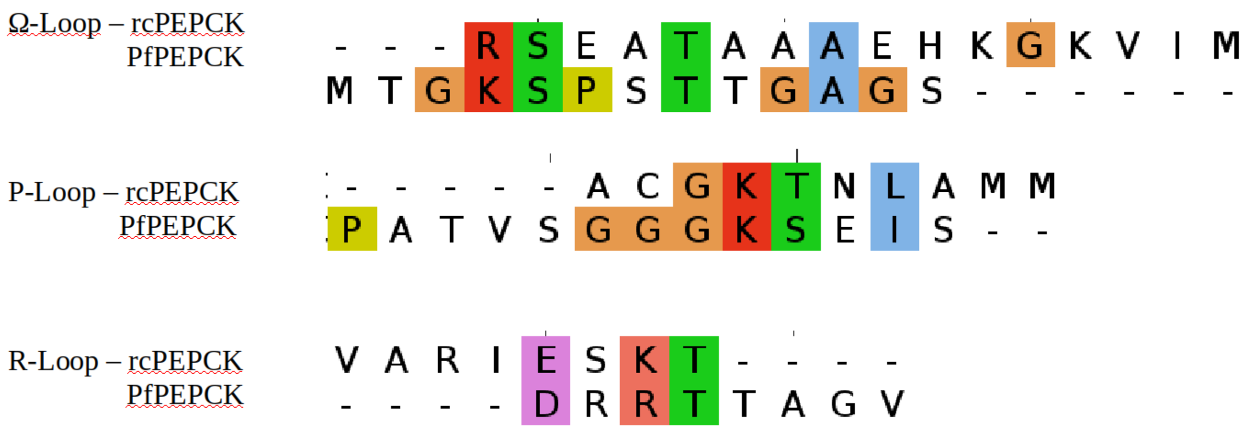
Cloned Sequence for *Propionibacterium freudenreichii* PEPCK

ATGAGCGTGGTTGAGCGTCGTCAGATCAACGCGGCGATTAACCTGCGTCTGAGCCTGCTGG
GTCTGCCGCACCCGGACAGCAACGCGGAAAGCCCGGATGCGATCCTGGTTGAGCCGCTGC
TGGCGCGTCAACGTGAACTGAGCCGTCGTCTGAAGGATCGTCTGAGCGCGCCGGACCTGC
GTATTCAGCGTTTCCTGGACGATTACCTGGCGGACTGCGATGAGCACCCGCAACTGCCGCG
TACCACCCTGGTGCTGGATGAACCGGGTCTGGCGCGTGGCCTGAGCCTGCCGGTTGACGGT
GATGAATTTCACAGCGACATCGTGGCGAGCTATCGTCTGGTTAACGGCGTGCTGCACAACC
CGAAACACGATCGTCGTACCACCGCGGGTGTTTTCCACATCAGCACCGGTGGCCTGCCGAT
TCCGCAGGATAAGGTTGAGGTGGACAAAAACGTGTACGCGCGTATTCTGGCGCGTGCGTTT
CAAGCGCCGGATGAGGAACTGGCGCTGCCGTATACCGCGAACCTGCCGGAACAAGCGCAC
TGCTGGGCGAGCCTGCTGATGCGTCCGACCGTTCTGCCGGCGGTGCCGGGTCGTACCACCG
AAAAGAGCTACGAGGTTCACTTCATCGTGCCGGGTGGCCTGATGTGCAACCTGGACTTTGT
TGAGGGCATTTTCGGTAACGCGGGCGACCCGTATCTGCCGGAAAACGATGCGAGCCTGGAC
CCGGATAGCTGGACCGGTCATACCGGTTGCGTGATCCTGGCGCCGCACCTGACCACCATGA
CCAAGAAAAGCCTGGGCATGCCGCACTACGACGATGCGACCGAGCGTCAGCGTCGTGACG
GTCAATGCTGGCGTCACGAAGACGATCTGTATAACGACGGCAAGGCGTTTAAAGTTTGCGC
GCGTGATGAGCGTGGCGTGATCGTTACCGTGATTGCGGATAACTACTTCGGTTATTGCAAGA
AAGAAGTGAAGACCCAGATCAGCTATAGCGCGAACCTGCTGGGTGGCGCGGAGGAAGAGC
ACAGCGGTGGTGCGGAAGTGTATCCGGCGTGGAACCTGAACCAAGATTTTACCGACCGTAC
CCCGGACGATTTCACCCTGGCGGACGTGATTAGCACCAACCGTGAACTGCTGGATGTTCGT
CCGGAGGGCTACGCGGTGTATAAACCGGAGCCGAACATCGTTTTCATTCCGGAACACAGCC
ACTACAGCATGCGTACCCAGACCATTAGCTGGACCGCGCATGGTGCGGAGCAAACCATTAA
GCTGCTGGCGGGTAAACACTACCTGAGCCCGGACGGCTATCGTATCCACGCGAAGCACCGT
GAAATGGATGCGACCCAGTGGCACCTGATTGGCACCAGCAGCCGTGCGGTTACCTGC
CATAAACCGGCGACCGTGAGCGGTGGCGGTAAAAGCGAGATCAGCAAAAGCATTAGCGAC
GCGTTCGTTTTTGGTAACGCGTTTAGCCACGACATCGATAGCGCGATGGATCAGGTGCAAGC
GCTGTTCGACACCGATTTTACCAACCGTTTCGCGGACGCGAGCCGTAACGGCACCGATCAC
CGTCCGGTTCTGAGCATCGACCGTAGCCTGGGTAGCGTGATCAAGCTGCTGACCCCGAGCA
TTCAGTACAACGACGAGTATAACGCGTTTCTGGAAGGCATTGAGCCGGATGTTAAGGAACT
GGCGTTCACCGTGAAACGTTACTATCTGCCGGAATGGGGTGAGGACTGGCGTAGCCACTTC
ACCGTTGGCATCATGAACGGTCGTCACGGCAACATGGTGCGTCTGGATGGTAAGAAAATCA
TTACCAACATGCTGCGTGTTGGTTTTCGTGAGGACGGCAGCTGGCGTCTGTTCACCCTGCG
TCCGGATTACAGCCCGGCGGTTAAAGTGCAAACCGAAGACGATATTACCGCGAGCACCGTT
ACCCCGCCGTGGGAAGATGCGGAGGGTCTGCCGCGTAAGTACGTGACCAACTGCGAGCAC
CTGCTGTTTCAGCGTCCGGACGATGCGATCCACCGTGGCTATGATAAACAAGCGGAATTCG
ACCTGGCGAGCGGTACCGATACCTTTATTAGCAACTTCGAACCGCTGACCCACGAGCAGGC
GCGTGATCTGCTGACCGACGTGCAAGCGTACAGCGAGTTTACCAAGCCGGTTCGTAAACTG
ATCGAACGTGTTGCGGCGATGCCGGATGATCAGAGCCCGGAGTTCTGGGTTTGCAGCGATG
ATCCGCGTCACCTGCCGGATGGCGGTCGTAGCAAGAACCCGCGTTATCTGCAAGTGCGTCC
GACCGACAGCAACCCGGAACTGACCACCGTTGCGGATGTGGCGGGCAAGCTGGCGCGTAA

ACTGCCGCTGGCGGGTCATGCGCCGCAGCCGATCGACGTGGTTGCGGCGGGTCGTCGTAAC
AACCCGCCGGAGGATAAAGTTCCGGCGCTGTGCGCGTACAACCCGCTGCACTATATGGAAC
TGCCGGAGCTGTTTATGGAGTACATTAGCAGCATGACCGGTAAAAGCCCGAGCACCACCGG
TGCGGGTAGCGAAGGTGCGCTGACCAAGGGTCCGTTCAACGCGCTGCCGGCGGTTTACGA
CCTGAACGCGGCGGTGCTGAGCTATGCGCTGACCGACTATGATGGTTGGCTGAGCAGCGCG
GGTTATATCGGCCCGAACGCGCGTGTGGACCACGATATCAGCATGCTGATTCCGGAACTGTT
CAGCCACATGGGCCCGAACGACCGTAACACCAAGCGTCTGATTAGCGAAGGTTACCTGGA
GAAAATGCAGGACTTCGATTTTGACGGTCACCGTGTTCTGGCGAGCCGTCTGGGCTACCGT
ATCAACGATCGTTTTGTGACCCACTATTTTGGCCGTATTTTCCTGCACCCGGACGTGGTTTTT
AGCGAAGAGATGCTGCGTCCGGAACTGCAAGATGAGAAGATCTTCGCGGATAGCATCGAC
GTTATTGTGAAAACCCACCAGCGTGTTGCGCAAATGTACTTCGATGATGGTACCGTGAGCCT
GGCGTGCCCGCCGATCCGTGCGCTGCTGGAGATTATGGCGCATGGTGCGAGCGCGGAGGGT
TGGACCCTGGACAGCCCGGAATTTCGTAAGCTGTTCGAACGTGAGAGCGTTCTGGCGAGC
GATTGGTATGCGGCGCGTCTGGATGCGAAGCAGGCGGAGGATGTTAAACAAACCGAAGAG
GGCGTGGAACGTCTGAAAGAATATATCGAGCGTCCGGACAGCGGTAGCGTGAGCGCGCGT
CTGCACCTGGCGGATCGTCTGCGTGAACTGGAGGCGCAGCTGACCTACGAACGTAGCCCG
GAGTATCGTCGTAGCCTGGTTGGTACCCTGGGCCGTCAACCGCGTTTTGTGTAA

Table 7.3.1: Comparison of the various dimer sets from all structures of PP_i PEPCK. The interface surface area was calculated from the PISA server, and the RMSD differences are from a global alignment through CCP4MG.

Malate Dimer (Chain ID)	**Interface Surface Area (Å2)**	**RMSD (Å)**
A-F	1692	
B-H	1642	1.04
C-D	1717	0.99
E-G	1735	1.25
Holo Dimer		
A-C	1711	1.21
B-D’	1693	1.22
6K31		
A-A’	1601	2.63
B-B’	1599	2.37


Ω-Loop – rcPEPCK - - - R S E A T A A A E H K G K V I M
PfPEPCK M T G K S P S T T G A G S - - - - - -

P-Loop – rcPEPCK - - - - - A C G K T N L A M M
PfPEPCK P A T V S G G G K S E I S - -

R-Loop – rcPEPCK V A R I E S K T - - - -
PfPEPCK - - - - D R R T T A G V

Figure 7.3.1: Sequence alignment of the active site loops of rcPEPCK and PfPEPCK. Aligned with ClustalO and coloured based on ClustalO scheme.

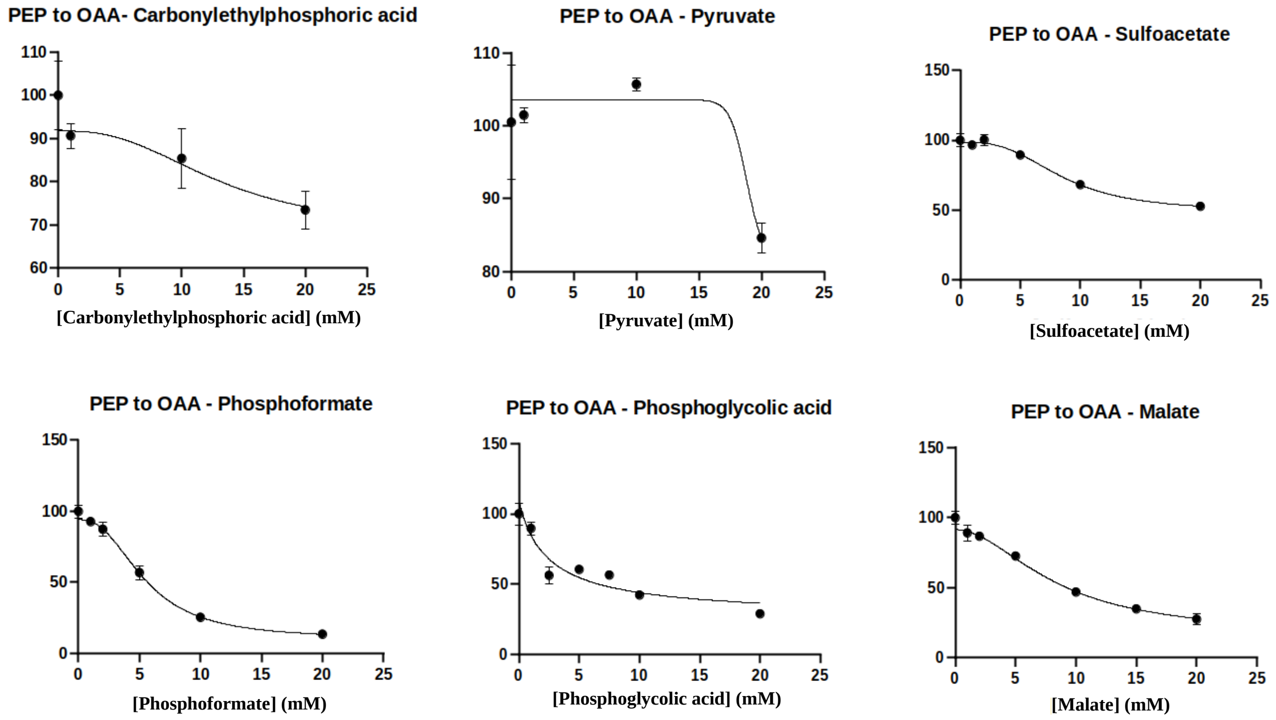


Figure 7.3.2: IC_{50} curves for inhibitors of PEPCK against the PEP carboxylation reaction. Y-axis is percent activity remaining. Errors bars may be hidden within the data point.

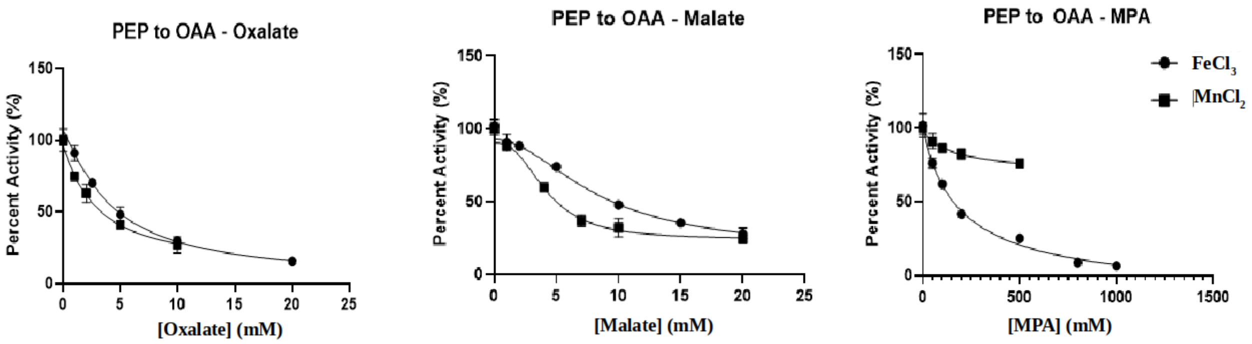


Figure 7.3.3: IC$_{50}$ of three inhibitors of PEPCK and there response to changing the M1 metal. Errors bars may be hidden within the data point.

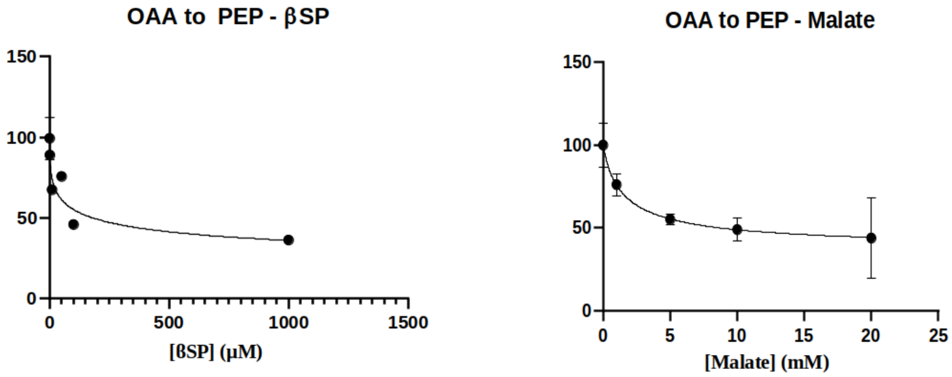


Figure 7.3.4: IC$_{50}$ curves for the previously described inhibitors of PEPCK against the OAA decarboxylation reaction. Errors bars may be hidden within the data point

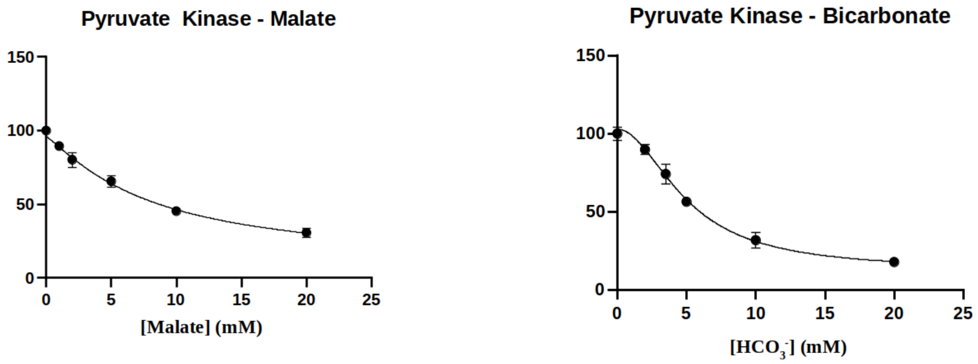
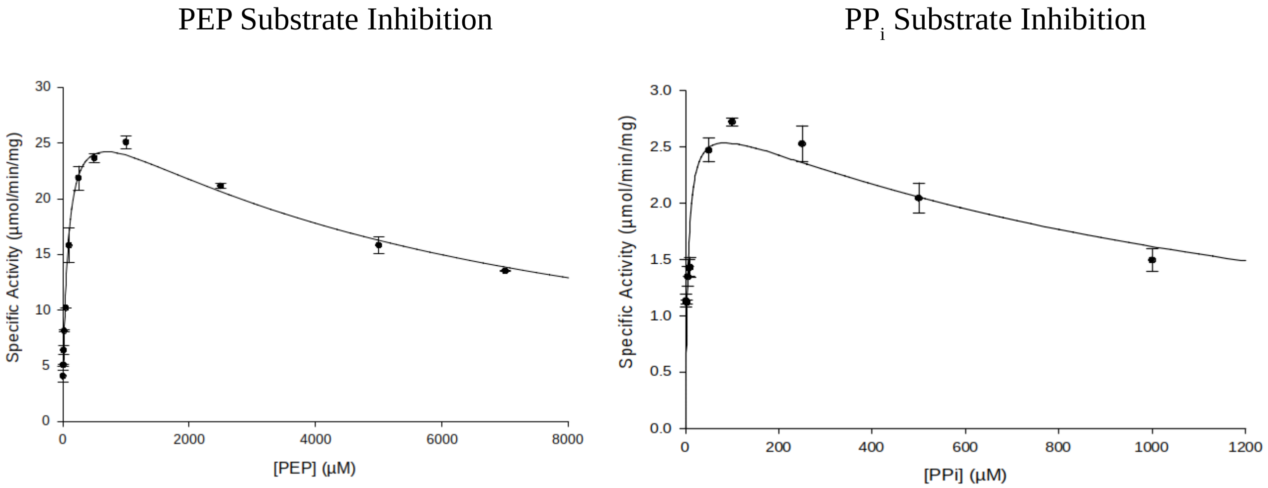


Figure 7.3.5: IC_{50} curves for the previously described inhibitors of PEPCK against the pyruvate kinase reaction. Errors bars may be hidden within the data point.



V_{max} = 29.5
K_m = 73.9
K_i = 6280.6

V_{max} = 2.9
K_m = 5.5
K_i = 1302.5

Figure 7.3.6: Substrate inhibition Michaelis-Menten plots for both PEP and PP_i against PEP carboxylation and OAA decarboxylation reactions respectively.

8.0 Structural Basis for the tight binding inhibition of *E coli.* CTP Synthase - Inhibition by Gemcitabine and its analogues

***Matthew J. Mcleod*[‡], *Stephen Bearne*[†], *Todd Holyoak*[‡]**

[†]*Department of Chemistry, Dalhousie University, Halifax, NS, B3H4R2, Canada*
[‡]*Department of Biology, University of Waterloo, Waterloo, ON, N2L3G1, Canada*

8.1 Introduction:

Cytosine 5’-triphosphate (CTP) synthase (CTPS) is a metabolic enzyme which regulates and is responsible for the essential *de novo* production of CTP. CTPS utilizes uridine 5’-triphosphate (UTP), ATP and ammonia in the production of CTP. While *in vitro*, the ammonia can be provided in the form of free ammonia, *in vivo* glutamine serves as the ammonia donor through the activity of the glutamine amidotransferase (GAT) domain of CTPS. Evidence suggests that the ammonia produced during the GAT catalyzed conversion of glutamine is transferred from the GAT domain of CTPS to the amidoligase domain where CTP synthesis is completed via a previously described substrate tunnel.[137] The chemical mechanism for CTP synthesis, as understood to this point, entails phosphoryl transfer of ATP’s γ-phosphate to UTP producing a 4-phospho-UTP intermediate which undergoes nucleophilic attack by ammonia resulting in the production CTP and inorganic phosphate.[138]

Due to its essential cellular role, the observation that CTPS is highly regulated based on available nucleotides is not surprising. GTP seems to be an activator at low concentration and an inhibitor at high concentrations.[139-140] This allosteric regulation has been studied but no affirmative structural data exists to shed light on the details of this mechanism or the GTP binding location. The *E. coli* variant of CTPS is a dimer in solution forming an active tetramer upon substrate addition which polymerizes to form inhibitory filaments when under particular controls, such as in the presence of high concentrations of the product CTP.[141] In contrast, the human variants form enzymatically activate filaments upon the addition of the substrate UTP.[142] Recent, cryo-EM studies have been insightful in demonstrating the global conformation changes that must occur for these two isozymes of CTPS to transition between the tetrameric to filamentous states. These global changes can be simply described as an increasing in packing of the tetramer interfaces to allow for stacking.[142,143]

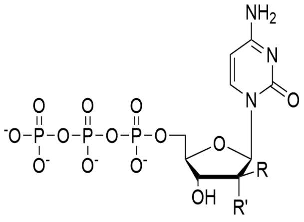
The crystal structure of *E coli.* CTPS was first solved in 2004 demonstrating the overall architecture of CTPS in its active tetrameric form.[144] Later in 2005, in an investigation by the same group, they characterized the product state with both CTP and ADP bound.[143] Taken together, these structural data in conjunction with the body of *in vitro* biochemical data suggests that the CTP and UTP sites only partially overlap through a common triphosphate binding pocket, while the pyrimidine rings of the respective nucleotides bind to unique sites. Both active sites are wedged between A-B-A' interface of the tetramer. This initial hypothesis on partially overlapping nature of the two nucleotide binding sites was recently confirmed in the determination of the human CTPS-UTP complex by cryoEM.[142]

Due to its essential role in nucleotide synthesis, CTPS has been shown to be an effective therapeutic target against protozoan parasites,[145,146] as well as several cancers.[147,148] Many CTP-like small molecules have been tested for their selective inhibition of CTPS. Historically, the most well characterized of these compounds are 3-deazauridine-5'-triphosphate (IC_{50} ~ 18μM) and cyclopentenyl cytosine 5'-triphosphate (CPEC) (IC_{50} ~ 6μM).[148,149] Although *in vivo* cellular work has generally shown them to be effective, CPEC has detrimental cardiotoxic effects which unfortunately does not make it an appropriate therapeutic candidate.[150] In addition, when used as antiproliferative agents, it is found that many cancers can alleviate the inhibition demonstrated by these two compounds through mutations that also removes normal CTPS product inhibition by CTP. It has been found that this mutational escape is achieved through modifications of the ribose and cytosine binding pocket of CTP and not the shared CTP/UTP triphosphate site.[151–153] Due to the this phenomenon, these particular mutations allow for UTP utilization with none of the inhibition generated by CTP binding.

In contrast to the compounds described above, gemcitabine triphosphate (dF-dCTP, Table 1) has been utilized as a therapeutic towards solid tumours and seems to be well tolerated.[147] In light of these positive aspects, a series of fluorinated CTP derivatives including dF-dCTP were tested for their ability to inhibit CTPS (Table 8.1).[154] dF-dCTP was the most inhibitory compound, being approximately 40 times more potent than CTP, and 30 times more potent that its tightest binding monofluorinated counterpart and can form filaments at a lower concentration than CTP.[155] This indicates that dF-dCTP may be a suitable starting point for further improvements then the previous generation of selective inhibitors. Based upon this prior work, we have undertaken the structural characterizations of these fluorinated analogues of CTP in efforts to determine the molecular basis for the tight binding of dF-dCTP when compared to the monofluorinated derivatives. It is hoped that these molecular details can further medicinal chemistry efforts to create more selective inhibitors of CTPS. Based upon this new structural data in combination with previous kinetic work, we demonstrate the dF-dCTP potency arises from binding to the closed, filament forming state of CTPS. This filament state allows for structural rearrangements of both inhibitor and active site to allow for the association of a cation to the triphosphate tail. This presents novel opportunities to exploit enzyme interactions with this new molecular scaffold for the next generation of CTPS inhibitors.

Table 8.1: Inhibitor Structure and Potencies. Adaptation from [154]

Inhibitor	IC_{50} Potency (μM)	R	R’
CTP	97 ±10	H	OH
F-dCTP	110 ± 7	H	F
F-araCTP	35 ± 5	F	H
dF-dCTP (Gemcitabine)	1.2 ± 0.1	F	F



8.2 Materials and Methods

8.2.1 Materials

Sodium malonate was purchased from Hampton Research. CTP and ATP were purchased from BioBasics. ADP was purchase from ChemImpex. dF-dCTP was purchased from JenaBiosciences. F-dCTP and F-araCTP was purchased from TriLink Biotechnologies. All other chemicals were purchased from the highest grade available.

8.2.2 Construct, Expression, and Purification

The construct used for this study has been outlined previously.[156] The protein was expressed in 2L of LB media and was induced with 1mM IPTG at an OD_{600} reading between 0.6 and 0.8 for 3 hours. Cells were harvested by centrifugation and resuspended in Buffer 1 (B1 - 25mM HEPES pH 7.9, 1M NaCl, 20mM Imidazole). Cells were sonicated in 5 pulses lasting 20s each while on ice. Lysate was spun down at 12 000G for 1 hour to pellet the insoluble portion. The supernatant was removed and added to pre-equilibrated nickel-NTA for 1 hour. The nickel was then washed with B1 until the flow-through had an A_{280} absorbance reading below 0.1. The bound protein was eluted with Buffer 2 (B2 – 25mM HEPES pH 7.9, 1M NaCl, 250mM imidazole) in 10mL fractions. These fractions were concentrated with an Amicon nitrogen gas concentrator using a 10kDa Millipore Ultracel© cellulose membranes.

The concentrated protein was then buffer exchanged into Buffer 3 (B3 – 25mM HEPES pH 7.9) by passage through a pre-equilibrated P6-DG column from BioRad. This protein was digested with 0.5U of thrombin to cleave the 6xHis tag over-night for 9 hours – any longer and the CTPS would become proteolysed. Approximately 75% of CTPS precipitated. To inactivate thrombin, 1mM PMSF was added. The remaining digested protein was then added to cleaned nickel-NTA resin equilibrated in B3, for 1 hour. The nickel resin was washed with B3 and the protein, now in the flow-through, was concentrated to 15mgmL$^{-1.}$. The final protein concentration was determined by nanodrop using the $E_{280\ 1\%}$ of 6.7mLmg^{-1}. The protein was frozen by direct submersion in liquid nitrogen in 30μL drops and subsequently stored at -80°C.

8.2.3 Crystallization and Data Collection

Crystals were grown as previously described.[144] In brief, crystals were setup at room temperature but grown by hanging drop vapour diffusion at 4°C. Mother liquor conditions consisted of 0.1M TRIS-Cl pH 8.0 (at room temperature), 5mM $MgCl_2$ and varied concentrations of $(NH_4)_2SO_4$ between 1.15M and 1.4M at a ratio of 2μL of protein at 15mgmL^{-1} to 2μL of mother liquor. Crystals were harvested and transferred into increasing concentrations of sodium malonate (pH 7.0) cryoprotectant supplemented with 5mM $MgCl_2$.[157] Crystals were initially placed in 1.5M sodium malonate and subsequently transferred to a 1.9M solution. Finally crystals were transferred to 2.4M sodium malonate with the ligands present at 1mM where they soaked at 4°C for 1 hour. Many crystals appeared to crack and ones that were stable were then cryopreserved by submersion into liquid N_2. Data were collected at the Canadian Light Source in Saskatoon, Saskatchewan on the 08B1-1 beamline. Data were indexed and scaled in HKL-2000[79] with final statistics presented in Table 8.2.

8.2.4 Structure Determination

All structures were solved through molecular replacement using the previously determined C268A mutant *E coli.* CTPS structure (PDB ID# 5TKV) and MolRep in the CCP4 package[80,142]. Iterative rounds of real space refinement and the addition of water molecules and heteroatoms were carried out with COOT, followed by refinement with REFMAC5[81,82]. Final MolProbity statistics were collected with the online server at http://molprobity.biochem.duke.edu/index.php and final model statistics are present in Table 8.2 [84].

Table 8.2: Collection and Refinement Statistics for CTPS Complexes

	dF-dCTP	**dF-dCTP + ATP**	**dF-dCTP + ADP**
Wavelength (Å)	1	1	1
Resolution range (Å)	54 - 2.2 (2.3 - 2.2)	50 - 2.4 (2.5 - 2.4)	57 – 2.0 (2.1 – 2.0)
Space group	P 21 21 2	P 21 21 2	P 21 21 2
Unit cell (Å)	158.415 108.962 129.029 90 90 90	157.991 109.549 129.105 90 90 90	157.748 108.763 128.86 90 90 90
Total reflections	3667046	3450292	6031925
Unique reflections	113698 (11244)	87367 (8468)	150135 (14297)
Multiplicity	15.0 (13.8)	14.3 (13.4)	13.8 (10.3)
Completeness (%)	100 (100.00)	99.7 (97.6)	99.2 (95.7)
Mean I/sigma(I)	18.8 (1.9)	15.8 (2.1)	16.7 (1.2)
Wilson B-factor (Å^2)	34.7	35.2	37.3
R-merge	0.149 (1.286)	0.180 (1.25)	0.127 (1.9)
R-meas	0.154 (1.334)	0.187 (1.3)	0.132 (2.067)
R-pim	0.04 (0.357)	0.049 (0.346)	0.035 (0.630)
CC1/2	0.986 (0.930)	0.996 (0.754)	0.973 (0.754)
Reflections used in refinement	113688 (11244)	87362 (8468)	150129 (14297)
Reflections used for R-free	5636 (566)	4366 (415)	7506 (728)
R-work	0.197 (0.281)	0.194 (0.266)	0.195 (0.308)
R-free	0.213 (0.301)	0.222 (0.295)	0.219 (0.313)
Number of non-hydrogen atoms	8985	8996	8968
macromolecules	8410	8402	8408
ligands	76	124	116
solvent	499	470	444
Protein residues	1070	1073	1069
RMS(bonds)	0.013	0.013	0.013
RMS(angles)	1.7	1.7	1.7
Ramachandran favored (%)	97.9	97.2	97.6
Ramachandran allowed (%)	2.1	2.7	2.2
Ramachandran outliers (%)	0	0.1	0.2
Rotamer outliers (%)	1.9	2.5	1
Clashscore	2.1	2.4	2.5
Average B-factor	45.6	45.5	46.7
macromolecules	45.9	45.9	46.6
ligands	28.9	41.9	44.6
solvent	42.3	38.4	47.8

	CTP	**F-dCTP**	**F-araCTP**
Wavelength (Å)	1	1	1
Resolution range (Å)	44 - 2.0 (2.1 - 2.0)	44 - 2.0 (2.1 - 2.0)	64 – 2.3 (2.4 - 2.3)
Space group	P 21 21 2	P 21 21 2	P 21 21 2
Unit cell (Å)	164.707 103.907 131.835 90 90 90	164.419 103.506 131.562 90 90 90	158.997 108.58 129.014 90 90 90
Total reflections	5344399	6118499	4389151
Unique reflections	142128 (10570)	152954 (14592)	103170 (9913)
Multiplicity	14.0 (12.1)	14.4 (6.9)	12.9 (12.1)
Completeness (%)	97.3 (73.3)	99.6 (96.2)	99.6 (96.7)
Mean I/sigma(I)	12.7 (3.8)	16.8 (2.0)	13.8 (3.0)
Wilson B-factor (Å^2)	33.4	33.6	48.5
R-merge	0.166 (0.851)	0.124 (0.878)	0.141 (1.741)
R-meas	0.172 (0.887)	0.101 (0.950)	0.147 (1.820)
R-pim	0.045 (0.248)	0.033 (0.346)	0.041 (0.524)
CC1/2	0.992 (0.971)	0.989 (0.887)	0.989 (0.920)
Reflections used in refinement	142115 (10570)	152938 (14592)	103168 (9914)
Reflections used for R-free	7162 (518)	7766 (726)	5051 (516)
R-work	0.186 (0.269)	0.183 (0.253)	0.203 (0.277)
R-free	0.206 (0.293)	0.198 (0.265)	0.231 (0.324)
Number of non-hydrogen atoms	9432	9462	8678
macromolecules	8497	8489	8394
ligands	86	86	65
solvent	849	887	219
Protein residues	1068	1066	1068
RMS(bonds)	0.013	0.013	0.013
RMS(angles)	1.7	1.6	1.7
Ramachandran favored (%)	98.2	98.8	97.7
Ramachandran allowed (%)	1.6	1.2	2.2
Ramachandran outliers (%)	0.2	0	0.1
Rotamer outliers (%)	0.6	0.4	2.5
Clashscore	2	2.1	2.7
Average B-factor	39.3	39.4	56.9
macromolecules	38.9	38.7	57.1

8.3 Results

The solved structures are all soaked crystals of the various different ligand compositions at high enough concentrations to get an occupancy of 1.0 at the binding sites.

8.3.1 dF-dCTP + ATP/ADP Complexes

To determine if there were any conformational changes that were induced by binding of ATP or ADP with CTP analogues, dF-dCTP complexes with these nucleotides were solved. Unsurprisingly, there were no observed changes between the dF-dCTP structure, and the complexes also with ATP or ADP. The global conformation of the crystallographic dimers are identical with a global RMSD change of 0.22Å and 0.27Å for the ADP and ATP complexes respectively (Figure 8.1). The B-factors of the residues interacting with the adenosine nucleotides are essentially comparable across all three structures, likely because even the dF-dCTP structure has malonate from the cryoprotectant bound in the ATP/ADP polyphosphate binding pocket. No conformational changes were observed between the three structures for any side chain rotamers as well. The equivalency of the dF-dCTP structures simplifies the discussion so all comparisons discussed are going to be made in reference to the dF-dCTP structure, but can be extrapolated to the adenosine nucleotide complexes.

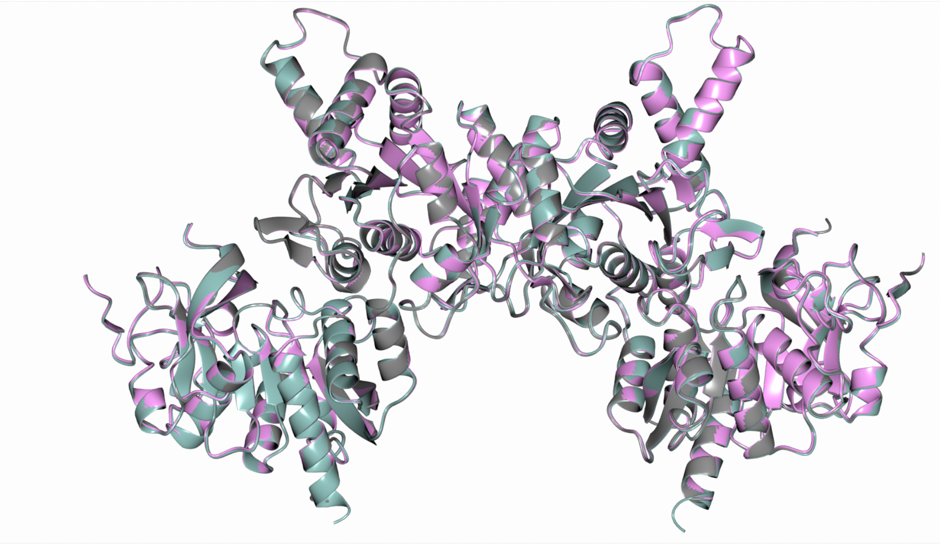


Figure 8.1: Global alignment of the dF-dCTP complexes. Grey is dF-dCTP, pink is dF-dCTP+ADP, and teal is dF-dCTP+ATP.

8.3.2 Global Conformation Changes

Global conformational comparison between CTP and the 3 fluorinated inhibitors was completed. The RMSD values are from a global superposition of the inhibitors against the CTP bound structure (Figure 8.2). As shown from the RMSD values, both CTP and F-dCTP complexes are globally identical. Unfortunately, the global alignment and its RMSD values of F-araCTP and dF-dCTP do not accurately show the actual conformational rearrangements that are going on between the three states. To further show the differences between the CTP (and F-dCTP) in comparison to F-araCTP and dF-dCTP a superimpostion was completed while anchoring the structures with one of the sub-units (Figure 8.2). The superimposition highlights the observed three global conformations. Even though F-araCTP and dF-dCTP are not the same conformation, they are more similar to one another than either to the mutant CTP structure (PDB ID# 5TKV) or dF-dCTP. Interestingly, the dF-dCTP structure is most similar to the previously solved C268A structure (PDB ID# 5TKV) with a RSMD difference of 0.4Å (as opposed to 0.6Å when compared to F-araCTP).

Complex	RMSD (Å)
F-dCTP	0.18
F-araCTP	1.39
dF-dCTP	1.46
dF-dCTP + ADP	1.49
dF-dCTP + ATP	1.53

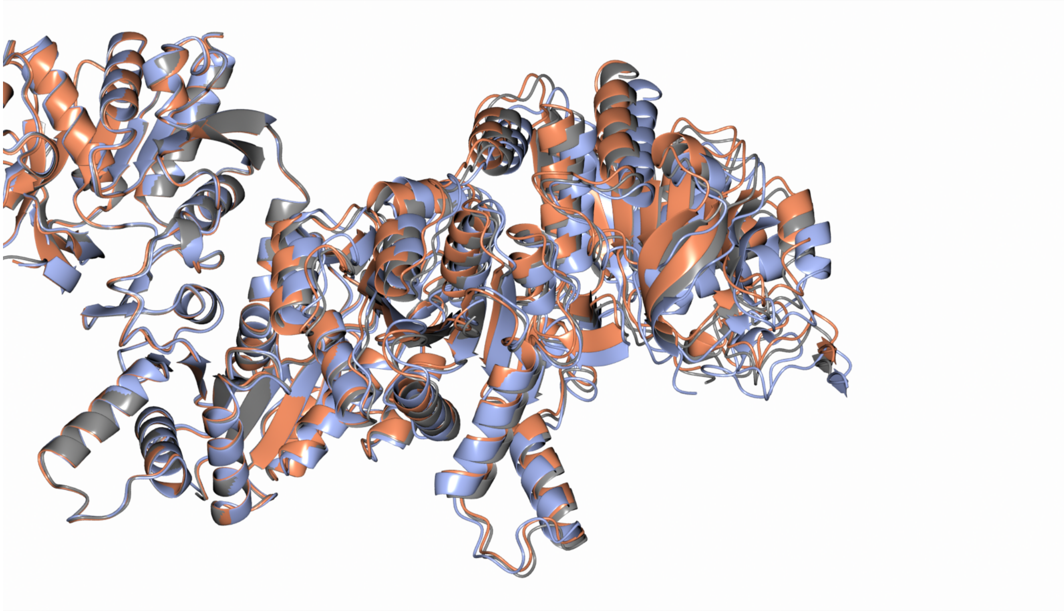


Figure 8.2: CTPS complexes anchored alignment and RMSD variability. The table is a comparison of the RMSD values of the complexes vs the CTP bound structure after a global superimposition. The right panel is a superimposition when aligned through one of the sub-units. The CTP structure is in ice blue, F-araCTP in coral, and dF-dFCTP is in grey. As seen, the left sub-unit is well aligned with all of the molecules but in the right sub-unit each of these three structures have different global conformations.

As previously mentioned, the CTP binding site is composed of 3 different sub-units of the crystallographic dimer, A-B-A'. Because of this, the tetramer is necessary to understand all binding interactions of this pocket. The crystallographic dimer and its symmetry mate fortunately recapitulate the tetrameric form of the enzyme. All of the further analysis and observations are completed with the tetrameric form of the enzyme. Furthermore, the various models were superimposed while anchoring chain A, and this pocket was used for the visualization so the local conformation changes can be most accurately described.

8.3.3 CTP Binding Site and Inhibitor Conformation

Along the same trend of the global conformational modes, there are three distinct local binding conformations for CTP and its analogues. CTP and F-dCTP are found in the same conformation as represented by the CTP molecule, while F-araCTP and dF-dCTP are found in two different conformations, although they are quite similar (Figure 8.3). The consequences of the global rearrangement of the tetramer interface can be seen in the three helices denoted.

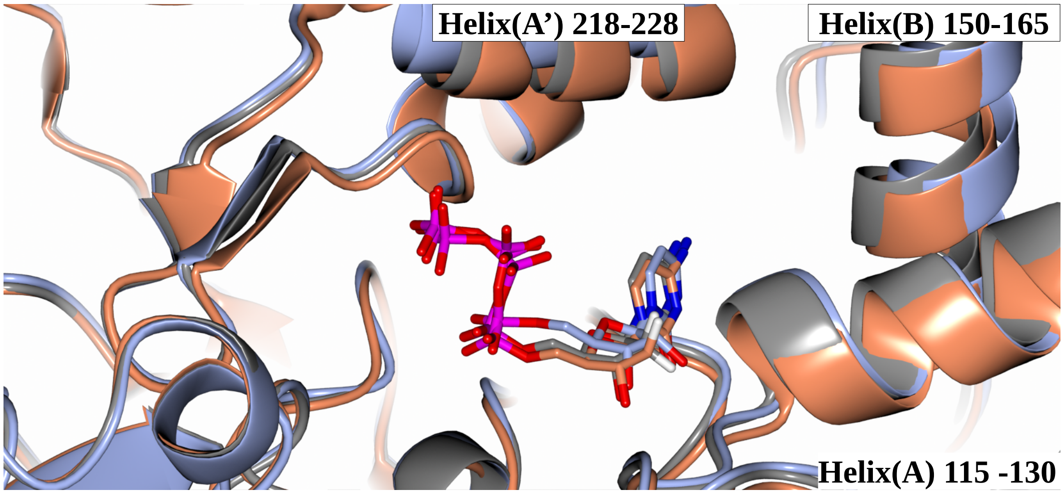


Figure 8.3: Ligand conformations in CTP binding pocket. Three ligand binding modes in the CTP pocket of the superimposed A sub-unit. Three helices making up the CTP binding pocket are denoted with chain identifier and residues numbers. The background protein structures are CTP bound in ice blue, and dF-dCTP in grey, F-araCTP in coral. All other atom colourings are as followed, green carbon, red oxygen, blue nitrogen, pink phosphorous, white fluorine.

8.3.4 CTP Binding Site Interactions and Metal Binding

Many binding interactions (and subsequent bond lengths) were automatically generated through LigPlot+ but manual curation was needed as some interactions were missed. The manual inspection of the potential interactions and bond lengths were completed in COOT. In general, the binding interactions are mostly conserved (Figure 8.4). This is unsurprising because although there are slight changes to the binding of the ligand and the binding pockets accommodation, the same residues are being used. The circled atoms/residues show specific additions to the binding interactions between the CTP reference and that of the inhibitor structures (Figure 8.4). Of note, there is a new interaction between F227 and the F-araCTP and dF-dCTP inhibitors. Also, as discussed later, there is evidence of a cation binding between the α-, β-, and γ-phosphates in the dF-dCTP structure which is absent in all other complexes (Figure 8.5).

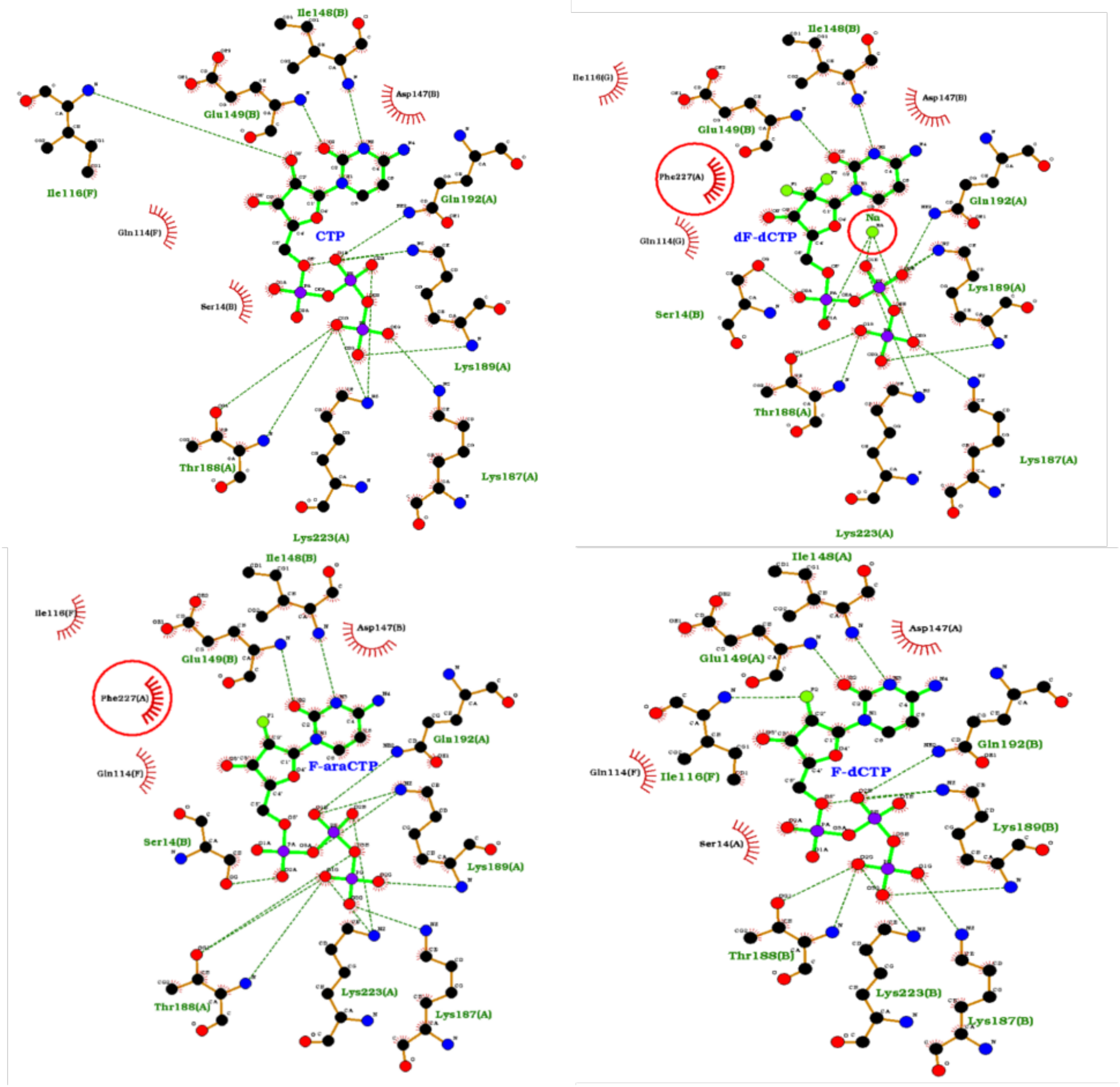


Figure 8.4: LigPlot+ schematics for the solved complexes. Atoms are coloured based on type; black is carbon, red is oxygen, blue is nitrogen, purple is phosphorous, lime green are other atoms (fluorine or sodium). Bonds for contacting residues are in brown, whereas ligand bonds are in green. Potential interactions are in dashed green and suspected hydrophobic interactions are shown as ray diagrams. The red circles indicate new interactions that are not found in the CTP reference structures. Residues are labelled with their chain notation as well.

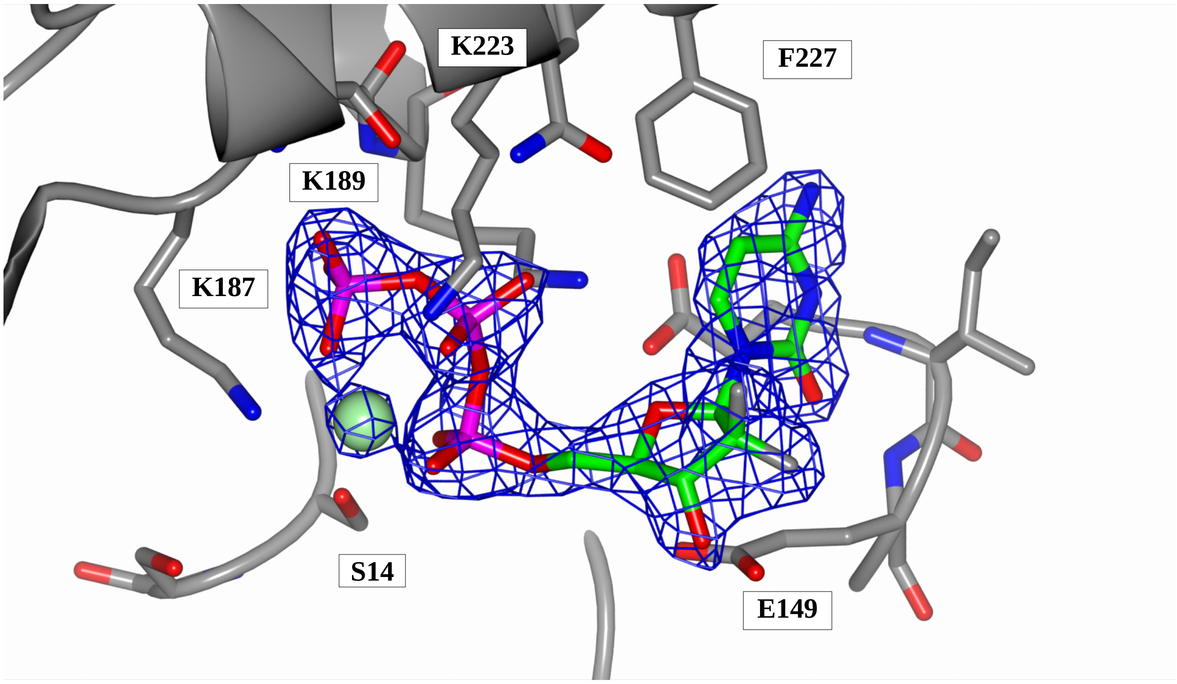


Figure 8.5: Gemcitabine binding. dF-dCTP structure with the blue mesh indicating the electron density from the $2F_o$-F_c map (rendered at 3σ). There is clear density for all atoms in the inhibitor as well as for the sodium cation. Atoms are coloured by type; grey is carbon, red is oxygen, blue is nitrogen, purple is phosphorous, white is fluorine, light green is sodium.

8.3.5 Bond Length Analysis

CTP vs F-dCTP

The first comparison of bond lengths was between the CTP and F-dCTP structures. Unsurprisingly, with the lack of change in the global conformation, inhibitor conformation, and reported IC_{50}, there are no significant changes in any residue – ligand bond lengths, besides E149 which has an increase (0.5Å) in bond length with the 2'-F-*ribo* of F-dCTP.

F-araCTP vs F-dCTP

It is apparent that the changes between F-araCTP and F-dCTP inhibitors are the same as those for dF-dCTP and F-dCTP. Because of this redundancy, these changes will be described below in the context of the dF-dCTP inhibitor.

dF-dCTP vs F-dCTP

Unsurprisingly, the majority of the differences between dF-dCTP and F-dCTP are a consequences of dF-dCTP having the additional fluorine at the R' (2'-*arabino*) position, and the corresponding metal ligand contact that are only found in the dF-dCTP structure . Four other interaction changes that are not because of the additional 2' fluorine are further described (Figure 8.6). First, K223 has changed positions, where in the F-dCTP complex it has a closer interaction with the γ-phosphate where in contrast, K223 is closer to the β-phosphate in the dF-dCTP complex. The net effect is shorter interactions from the K223 in the F-dCTP structure. Second, K189 has also changed it relative position between these two structures to interact with different oxygens of the triphosphate tail. In the dF-dCTP complex, K189 interacts with the bridging oxygen between the α and β phosphates while in the F-dCTP structure this lysine coordinates with the bridging oxygen between the ribose moiety and the α phosphate. When looking at the specific distance differences, the changes of the two modes likely at least partially negate one another as the magnitude of the changes are equivalent. Third, S14 of the dF-dCTP structure has moved positions to have a 1.1Å shorter H-bond between its hydroxyl and an oxygen of the α phosphate. Finally, F227 has rotated towards to the cytosine ring for a potential 3.7Å π-stack in the dF-dCTP structure, as well as 4.4Å interaction with the 2'-F-*arabino*. In the CTP and F-dCTP, F227 is disordered based on the lack of electron density to be properly fit, preventing this stabilizing interaction.

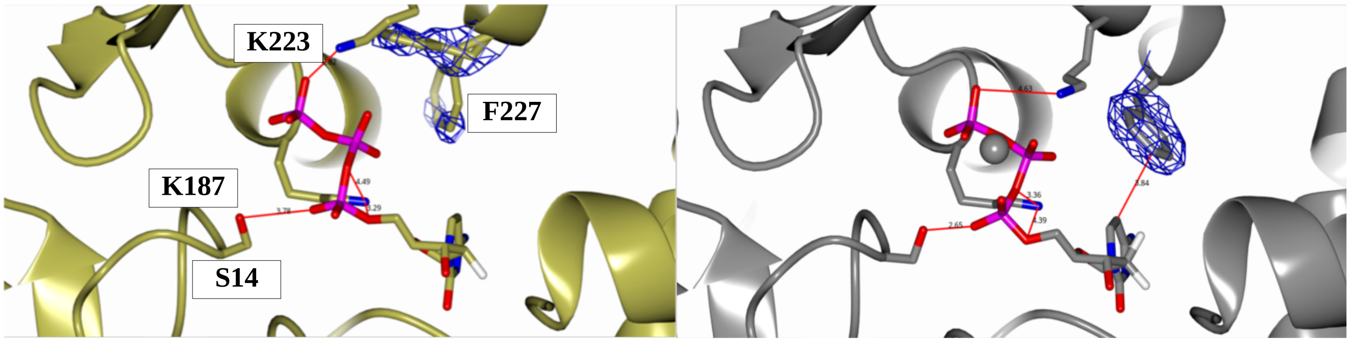


Figure 8.6: dF-dCTP and F-dCTP binding modes. Changing binding interactions between F-dCTP (left, gold) and dF-dCTP (right, grey). Blue mesh indicates electron density from the $2F_o$-F_c clipped to F227 (rendered at 1.7σ).The additional bonds that dF-dCTP has due to the additional R' fluorine are not shown.

dF-dCTP vs F-araCTP

The comparison between dF-dCTP and F-araCTP are to show the consequences that dF-dCTP has the additional fluorine at the R (2'-*ribo*) position. Most of the new bonds are due to increase contacts with this R fluorine, and the corresponding metal ligand contact that are only found in the dF-dCTP structure. As for interactions not coordinating with the second fluorine, K223 appears to have a shortened contact in the F-araCTP structure with a terminal oxygen on the γ phosphate but when inspecting the electron density maps of F-araCTP, there is no clear density to model the Cδ, Cε, Nζ atoms suggesting it is unstructured (Figure 8.7). K187 on the other hand is structure and does have a slightly shorter (0.7Å) interaction with terminal oxygen of the γ-phopshate in the dF-dCTP structure.

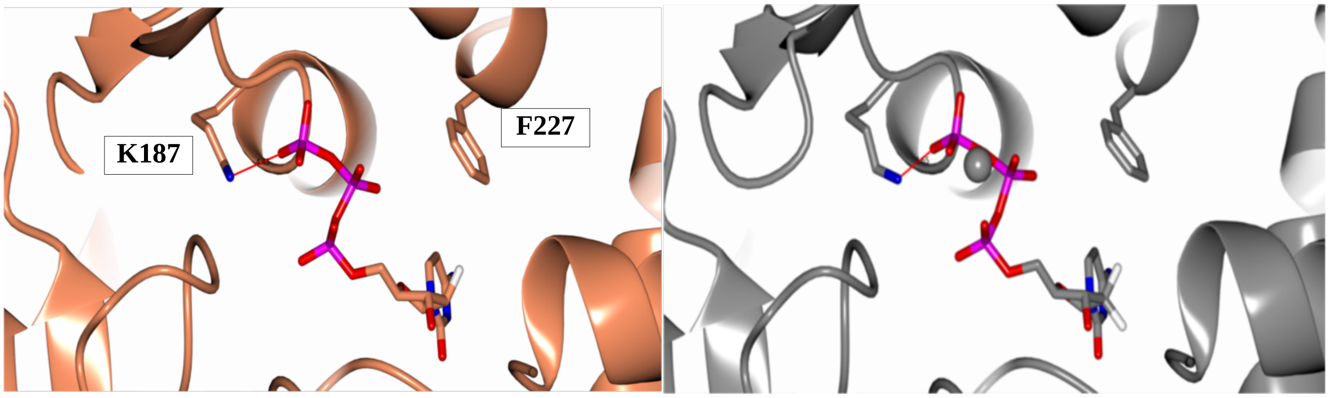


Figure 8.7: dF-dCTP and F-araCTP binding modes. Changing binding interactions between F-araCTP (left, coral) and dF-dCTP (right, grey). The additional bonds that dF-dCTP has due to the additional R fluorine are not shown.

dF-dCTP vs C268A + CTP (5TKV)

The final comparison is between the globally similar dF-dCTP complex and the previously solved structure of C268A *E. coli* CTPS (PDB ID# 5TKV). The global conformational similarities discussed earlier, and the presence of the cation binding between the triphosphate tails of both dF-dCTP and C268A+CTP structure led to the comparison of interacting ligand distances. Unsurprisingly, there are no significant differences between the bond lengths associated with the inhibitors. F227 is also rotated in the 5TKV structure to form the π-stack with cytosine moiety of CTP reminiscent of the dF-dCTP and F-araCTP structures.

8.4 Discussion

8.4.1 2’-*ribo* substitution - Similar Inhibition by CTP and F-dCTP

The previous investigation determined that CTP and F-dCTP inhibit CTPS with essentially the same potency, with IC_{50} values of 97µM and 109µM respectively.[154] The global conformational comparison indicates that the substitution for the 2’-*ribo* hydroxyl for fluorine does not change the global conformation of the enzyme(Figure 8.2). The observed local structure highlight only one difference between any of the participating binding residues, that is between E149 and the 2’-*ribo* group. In the case of CTP binding, the E149 terminal oxygen is 2.95Å from the 2’-OH-*ribo*, whereas in the F-dCTP (2’-F-*ribo*) structure this bond is 3.40Å. This increase in distance is accommodated by a slight movement of the ribose ring, whereas the triphosphate tail and cytosine moiety are in essentially identical positions. In previous discussions, questions were raised as to whether E149 was protonated.[154] Mutagenic studies were also completed to better understand the protonation state of E149. E149D was not able to be inhibited by either CTP or dF-dCTP, suggesting that the shorter aspartic acid residue (by ~1.5Å) removed this very important bonding interaction. Finally, E149Q, which removed the negative charge of glutamic acid and its H-bond accepting potential, had a relatively unperturbed effect on the inhibition of CTP and dF-dCTP. This finding lead to conclusions that E149 is protonated, as E149Q amino group could donate H-bonds (and thus inhibition was unchanged). E149 protonation seems likely from the previous biochemical data, and the subtle increase of bond length upon fluorine substitution. The C-F bonds are unique as they are H-bond acceptors, polar, non-polarizable, slightly hydrophobic, and partially negative.[158] If E149 is deprotonated, this partial negative charge of fluorine would repel with the formal negative charge on the oxygen of E149, likely leading to a large increase in distance between these two atoms, and relatively weaker binding which is not observed through the IC_{50} values. If protonated, potentially the "fatness", and electronegative potential of the fluorine may lead to ever so slight alterations in the H-bond distance as seen here but results in similar free-energy stabilization as observed from the biochemical inhibition and structural data. This structural information only adds to the understanding of E149 protonation state and is by no means conclusive evidence of either state. In brief, the structural evidence suggests lack of significant changes in active site interactions with the two compared ligands besides a 0.5Å increase between E149 and F-dCTP, leading to a lack of global conformational change and results in these two compounds having essentially equivalent potency.

8.4.2 2'-*arabino* Substitution - 3x Increased Inhibition of CTP/F-dCTP vs. F-araCTP

Although we compared structures of CTP or F-dCTP, which have the 2'-*ribo* position utilized, the 2'-OH-*arabino* inhibitor, araCTP, was previously tested for its potency.[154] It is an extremely poor inhibitor (IC_{50} ~>1000μM) clearly indicating that fluorine, rather than a hydroxyl, is important in the potency of this molecule at this position. As discussed later, this may be due to the proximal residues creating a hydrophobic pocket. The interpretation of this relatively small increase in binding efficiency by the 2'-F-*arabino* substitution in contrast to the 2'-F-*ribo* is quite difficult to deconvolute. Specifically, the binding interactions are different as the orientation of each of the fluorines allows for many different contacts. The sum effect of these different interactions is very difficult to understand and will therefore not be of focus. Changes identified in bond lengths between residues and the non fluorinated atoms may indicate the relatively small increase in binding efficiency for F-araCTP. As mentioned, S14 was observed to have a slightly shortened (~1Å) contact with the α-phosphate terminal oxygen. This decreased bond length may stabilize the inhibitor giving a modest increase in potency.

8.4.3 F227 Conformational Change and Filament Formation

The most significant change in the binding interactions between the F-dCTP and F-araCTP structure is the addition of F227 contact with 2'-F-*arabino.* F227 in the CTP and F-dCTP structures, as well as previously published structure of CTP-ADP complex (PDB ID# 2AD5) have a lack of electron density to model F227 in one position. In contrast, the F-araCTP, dF-dCTP, and mutagenically "closed" C268A mutant (PDB ID# 5TKV) structure indicate F227 is in one position, π-stacking with the cytosine ring (3.7Å), and contributing hydrophobic interactions with the 2'-F-*arabino* moiety (4.4Å) (Figure 8.6).[142] This new interaction that F-araCTP has is likely a significant cause for observed increase inhibition of this molecule for CTPS. F227 is important in polymerization as it shifts its helix (as seen in Figure 8.3) rearranging the CTP pocket interface.[142–144,159] This rearrangement leads to a global tightening and subsequent polymerization. This interaction has been previously studied to be important in filament formation where McCluskey et al employed two mutant forms of CTPS, F227L and F227A.[155] These mutants were used to study the types of interactions that made dF-dCTP a strong filament-former. F227A would expand the pocket, prevent any π-interaction and will slightly weaken potential hydrophobic interactions. F227L prevented π-interactions while mostly maintaining hydrophobicity. The IC_{50} data collected showed that CTP inhibition was unaffected by F227A (IC_{50} = 30μM) while the leucine mutant weakened the inhibition (IC_{50}=340μM) when compared the WT inhibition (IC_{50} = 81μM). This suggests that CTP inhibition is mostly through other auxiliary stabilizing forces (potentially E149 H-bonding), that the π-interaction is not vital, and hydrophobic interactions may be detrimental. In context with this new structural work, the mobility of F227 in the CTP structure (and F-dCTP), show that this is not a necessary interaction for inhibition of these two molecules. The two F227 mutants were also tested against dF-dCTP where the inhibition was relatively unaffected (F227 IC_{50} = 1.2μM, F227A IC_{50} = 1.3μM, F227L IC_{50} = 1.9μM) suggesting that the π-interaction is inconsequential whereas hydrophobicity nature of this pocket seems important for dF-dCTP inhibition.[155] Although our data suggests that due to the proximity of the F227 and the cytosine ring there is likely a π-interaction, based on the relative orientation of the two ring structures this may not be the main source of strong inhibition. The only other previously deposited structure of *E. coli* CTPS where F227 is edge-on with CTP is in C268A. Although no rationale was offered as to why this mutant would preferentially adopt the "closed" state, it does shift the 216-228 helix. Although the cryo-EM structures of the filament suggest a more drastic "closing", the C268A mutant has been used a crystallographic proxy for the filamentous state.[142] In summary, the conformational change of F227, coupled with the subsequent global closure of the enzyme suggests that both F-araCTP and dF-dCTP are a strong polymerizing agents and inhibitors. In contrast CTP and F-dCTP are weaker inhibitors and, at least for CTP, we know will cause tetramer formation, and in some circumstances filaments. Finally, the specific molecule recognition of this tight binding of the 2'-F-*arabino* inhibitors may be more of a hydrophobic interaction than a π-sharing interaction based on previously collected kinetic data coupled and the new structural information.

8.4.4 2'-*arabino* Hydrophobic Pocket

To further investigate the potential hydrophobic contributions inducing tight binding in particular inhibitors, we evaluated the surface of the CTP binding pocket, in which there appears to be a clear distinction between a hydrophilic and hydrophobic portion (Figure 8.8). Unsurprisingly the binding site for the triphosphate tail is hydrophilic while the cleft corresponding with the 2' site and the cytosine ring are hydrophobic. As previously mentioned, fluorine in slightly hydrophobic in which desolvation of the active site upon binding may impart increase potency of fluorinated compounds over non-fluorinated substitutes.[158,160] This may be important in the vastly tighter binding of dF-dCTP and the monofluorinated derivatives ability to inhibit CTPS. The residues lining this hydrophobic pocket and its potential interaction with the fluorinated inhibitors was described earlier[155], and the crystallographic data supports this previous hypothesis. With F227 now in a stabilizing conformation, both V115 and I116 side chains are 3.9Å and 3.5Å away from the 2'-F-*arabino* moiety suggesting hydrophobic interactions. This further supports the observation why the hydrophilic 2'-OH-*arabino* substitution was so ineffective. In summary, the likely increase in potency of of F-araCTP is through a new stabilizing interaction with F227, as well as the exploitation of a hydrophobic pocket built by V115 and I116.

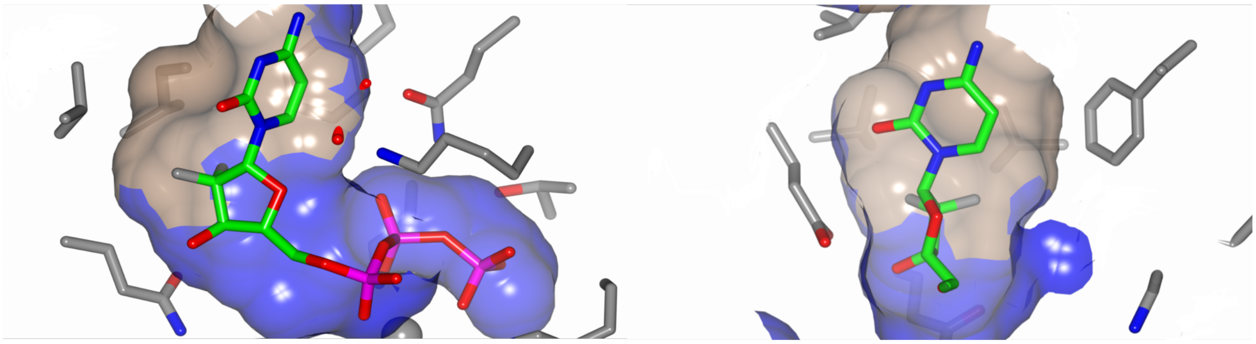


Figure 8.8: CTP hydrophobic binding site. Two different orientations (offest by 90°) of the hydrophobic surface of the CTP binding site. Blue surfaces indicate hydrophilicity and brown indicate hydrophobicity.

8.4.5 30x Increased Inhibition of dF-dCTP vs F-araCTP

The most important question that this study aims to shed light on is how dF-dCTP is so much more potent than either monofluorinated inhibitor. The most obvious suggestion for this increase in potency is the additive effect of being able to exploit both the 2'-F-*ribo* and -*arabino* groups. This is without a doubt very important and has even been suggested that doubly fluorinating the 2 position will actually lead to greater polarization of the C-F bonds, making the effects even stronger.[154] An important observation is the similarity of the global and local conformations of CTPS between the dF-dCTP complex the previously solved C268A mutant (PDB ID# 5TKV). Our dF-dCTP structure was able to induce this "closed" conformation that is closest towards the filamentous quaternary state, much like the C268A mutant. Interestingly, both the dF-dCTP and C268A structures are the only reported *E coli.* CTPS structures to have a cation bound to the triphosphate tail of their respective ligands. Both our CTP structure, and the previously deposited CTP-ADP complex (PDB ID# 2AD5), both are lacking this cation. This suggest that the binding of the metal is not due to the actual chemical species, but rather due to the global conformation state that is sampled. Both K187 and K223 are observed to be well modelled and interacting with dF-dCTP (or CTP if 5TKV). K187 are always found in the same position between all inhibitor structures, whereas K223 changes conformation for the C268A and dF-dCTP when compared to the CTP structure (Figure 8.9). Upon further investigation, K223 interacts with the terminal oxygen of the γ-phosphate (2.7Å) where it is unmodelled in the F-araCTP structure. K187 interacts with the terminal oxygen of the γ-phosphate at a distance of 2.75Å, whereas this length is increased to 3.3Å in the F-araCTP structure due to change in the inhibitor location. It is likely that these lysine-phosphate interactions are possible because of the global changes that occur as well as the local conformation of the inhibitor itself. The consequence of these lysine-phosphate interactions is the stabilization of the polyphosphate tail which allows for cation binding (Figure 8.9). The cation has an octahedral coordination with each terminal oxygen of the three phosphates and 3 water molecules. This cation binding will ultimately further reduce the energetic cost of stabilizing (and therefore induce tighter binding) CTP in the case of C268A, and dF-dCTP in our structure allowing it a more potent inhibitor. Based upon this hypothesis, it is likely that CTP potency is significantly increased against the C268A mutant as this inhibited state is adopted.

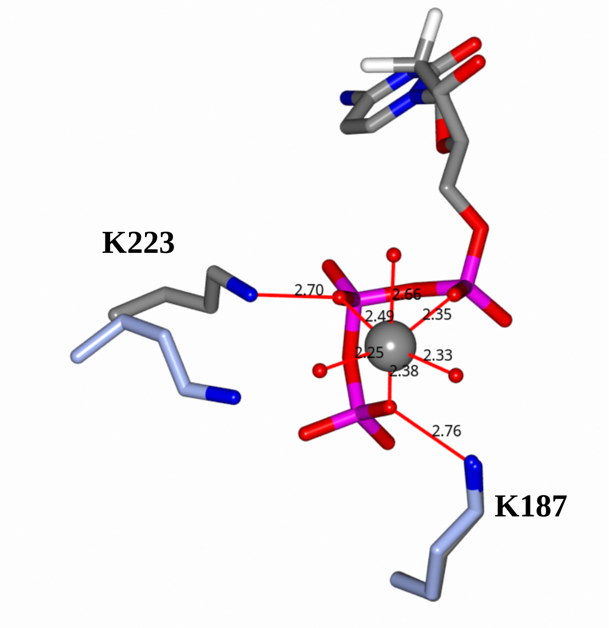


Figure 8.9: dF-dCTP metal coordination. Representative diagram of the metal binding ligands Lys187 and 223 under bound (dF-dCTP - grey) conditions, and non-bound (CTP – steel blue). Atoms are colored by type with oxygen in red, nitrogen in blue, phosphorous in pink, and flourine in white.

8.5 Conclusion

Our data presented above suggests new ligand binding interactions that may be exploited to create selective inhibitors of CTPS. As previously discussed, there are already two known compounds, 3-deazauridine-5’-triphosphate and cyclopentenyl cytosine 5’-triphosphate that are selective inhibitors of CTPS, although they have some unwanted side-effects and have been routinely selected against through mutations in cancer cell lines.[151–153,161] To date, dF-dCTP is the most potent inhibitor tested against CTPS and here we have shown, coupled with kinetic data, the molecular details to its potent inhibition. First, the 2’-F-*ribo* substitution over the OH functional group is relatively innocuous, as from the IC_{50} data, and over lack of global and local rearrangements suggest. Second, the 2’-F-*arabino* adds a new interaction with F227, which is likely hydrophobic in nature. The differences in the type of forces contributing to binding between CTP (electrostatic/H-bonding) and dF-dCTP (hydrophobic) may help prevent the aforementioned mutational escape of cancers for previous inhibitors of CTPS. dF-dCTP may utilize different types of interactions that potentially, these mutations do not diminish. Third, we still cannot rule out additive inductive effects by doubly fluorinating the 2’ position of the ribose ring further polarizing the C-F bonds and increasing the molecules hydrophobicity. Fourth, binding of dF-dCTP leads to a local rearrangement of the tetramer interface, which propagate to an overall conformational change through the entire enzyme. This induces a more closed filament-like state. Finally, both global and local conformational changes allow for a positional change in K223, which can stabilize the triphosphate tail of the dF-dCTP to allow for octahedral binding of cations, further reducing the energetic cost of binding. All of these aforementioned changes between our CTP structure, and dF-dCTP compex likely summate to account for the increased potency of this inhibitor. This study hopefully lends insight into the molecular mechanism of CTP binding as well as to offer an avenue towards further refinement of therapeutic targets.