Graduate School Form 30
Updated 1/15/2015

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Megan P. Rogers

Entitled
INVESTIGATION OF THE EVOLUTIONARY ASPECTS OF THIAMIN DIPHOSPHATE- DEPENDENT DECARBOXYLASES

For the degree of Master of Science

Is approved by the final examining committee:

Michael J. McLeish
Chair

Eric Long

Lei Li

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Michael J. McLeish

Approved by: Eric Long 4/6/2015
Head of the Departmental Graduate Program Date

INVESTIGATION OF THE EVOLUTIONARY ASPECTS OF THIAMIN DIPHOSPHATE-DEPENDENT DECARBOXYLASES

A Thesis

Submitted to the Faculty

of

Purdue University

by

Megan P. Rogers

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

May 2015

Purdue University

Indianapolis, Indiana

For my family

ACKNOWLEDGEMENTS

During my time at IUPUI, there are many people who deserve thanks for helping and guiding me along the way. I would first like to thank my advisor, Dr. Michael J. McLeish for allowing me the opportunity to work under his guidance and for providing support throughout the process. I would also like to thank Dr. Kathleen Marrs for giving me the opportunity to participate in the GK-12 fellowship program and for the ability to present science to students in a unique way. I would like to thank the members of the McLeish group and my fellow peers for their friendship and providing perspective at the most necessary of times. Lastly, I would like to thank my family, especially Will, for the support, patience, and understanding shown to me during this journey.

TABLE OF CONTENTS

Page

LIST OF TABLES viii
LIST OF FIGURES ix
LIST OF ABBREVIATIONS xii
ABSTRACT xiv
CHAPTER 1. INTRODUCTION 1
1.1 Thiamin Diphosphate 1
1.2 Pyruvate decarboxylase, a modern ThDP-dependent enzyme 5
1.2.1 The ThDP binding site 7
1.2.2 The active site 9
1.2.3 The regulatory site 10
1.3 Evolutionary aspects of ThDP-dependent enzymes 11
1.4 Research covered in this work 15
CHAPTER 2. MATERIALS AND METHODS 17
2.1 Materials 17
2.2 Transformations 18
2.2.1 *Escherichia coli* transformation 18
2.2.2 *Saccharomyces cerevisiae* transformation 19
2.3 Construction of expression vectors 20
2.3.1 Cloning and construction of expression vectors pET17PDC5-His, pET17PDC6-His, and pET17THI3-His 20
2.3.2 Construction of a zeocin yeast expression vectors 20

Page

2.3.3 Insertion of genes into zeocin yeast expression vector....22
2.4 Protein expression and purification of His-tagged enzymes....23
2.4.1 Expression of His-tagged enzymes....23
2.4.2 General purification of His-tagged proteins....23
2.4.3 Purification of phosphonopyruvate decarboxylase and phosphonoacetaldehyde hydrolase....25
2.5 Protein crystallization....26
2.5.1 Protein expression and purification for crystallization....26
2.5.2 Crystallization of PDC5....26
2.6 Characterization studies....27
2.6.1 Kinetic characterization of PDC1 and PDC5....27
2.6.2 Kinetic characterization of the benzoylformate decarboxylase T377L_A460Y variant....27
2.6.3 Steady state analysis of phosphonopyruvate decarboxylase....28
2.6.4 IC_{50} and K_i determination for difluorophosphonopyruvate....29
2.6.5 Reaction of difluorophosphonopyruvate with PhPyrDC monitored by circular dichroism (CD) spectroscopy and UV-VIS spectroscopy....29
2.6.6 Fluoride ion analysis of the reaction of difluorophosphonopyruvate with PhPyrDC....30
CHAPTER 3. INACTIVATION OF PHOSPHONOPYRUVATE DECARBOXYLASE FROM *BACTEROIDES FRAGILIS* BY DIFLUOROPHOSPHONOPYRUVATE....31
3.1 Background....31
3.2 Results and Discussion....34
3.2.1 Purification and kinetic characterization of PhPyrDC....34
3.2.2 Design and synthesis of difluorophosphonopyruvate....36
3.2.3 IC_{50} and K_i determination for DFPP....37
3.2.4 Dilution studies of PhPyrDC with DFPP....39

Page

3.2.5 Reaction of DFPP with PhPyrDC monitored by circular dichroism (CD) spectroscopy and UV-Vis spectroscopy 39
3.2.6 Fluoride ion analysis of the reaction of DFPP with PhPyrDC 42
3.2.7 Crystallization studies of PhPyrDC 43
3.3 Conclusions and Future Direction 44
CHAPTER 4. INVESTIGATION AND CHARACTERIZATION OF THE MINOR ISOZYMES OF PYRUVATE DECARBOXYLASE ISOLATED FROM BAKER'S YEAST 45
4.1 Background 45
4.1.1 PDC5 46
4.1.2 PDC6 48
4.1.3 THI3 50
4.2 Results and Discussion 51
4.2.1 Analysis of PDC5 51
4.2.1.1 Sequence alignments and homology model comparison of PDC5 and PDC1 51
4.2.1.2 Expression and kinetic characterization of PDC5 54
4.2.1.3 Crystallization, X-ray structure and structure analysis for PDC5 59
4.2.2 Analysis of PDC6 62
4.2.2.1 Sequence alignment and homology model comparison of PDC6 and PDC1 62
4.2.2.2 Expression conditions for PDC6 65
4.2.3 Analysis of THI3 66
4.2.3.1 Sequence alignment and homology model comparison of THI3 and PDC1 66
4.2.3.2 Attempted expression of THI3 70
4.3 Conclusions and Future Directions 70

Page

CHAPTER 5. INCREASING THE EFFICIENCY OF THE SCREENING AND SELECTION REQUIRED FOR SITE-SATURATION MUTAGENESIS OF PDC VARIANTS 72
5.1 Background 72
5.1.1 Site-saturation mutagenesis 72
5.1.2 Selection vs. screening 76
5.1.3 Evolution of benzoylformate decarboxylase to a pyruvate decarboxylase 77
5.2 Results and Discussion 80
5.2.1 The substrate spectrum of BFDC T377L-A460Y 80
5.2.2 Construction of yeast expression vectors 83
5.2.3 Yeast growth experiments 88
5.3 Conclusions and Future Directions 92
REFERENCES 94

LIST OF TABLES

Table Page

Table 2.1 *S. cerevisiae* strains used in this study 19

Table 2.2 Forward primers used to introduce new restriction sites..... 21

Table 2.3 Site-directed mutagenesis reaction mixture 21

Table 2.4 Site-directed mutagenesis temperature cycling parameters..... 21

Table 2.5 The zeocin yeast expression vector with gene inserts and the restriction enzymes used..... 22

Table 2.6 Molar extinction coefficients used in calculating protein concentration 25

Table 4.1 Differences surrounding binding sites between PDC1 and PDC5..... 53

Table 4.2 Kinetic Characterization of PDC5 carried out in a phosphate buffer..... 55

Table 4.3 Comparison of PDC1 and PDC5 in various buffers utilizing pyruvate as a substrate..... 56

Table 4.4 Kinetic Characterization of PDC1 utilizing a MES buffer..... 57

Table 4.5 Kinetic Characterization of PDC5 utilizing a MES buffer..... 58

Table 4.6 X-ray diffraction data collection and refinement statistics 60

Table 4.7 Differences surrounding binding sites between PDC1 and PDC6..... 64

Table 5.1 Steady-state Kinetic Characterization of BFDC T377L-A460Y..... 81

Table 5.2 Kinetic values for wild-type BFDC and BFDC T377L-A460Y 82

LIST OF FIGURES

Figure Page

Figure 1.1 Structure of ThDP in the 4'-aminopyrimidine (AP) form 2

Figure 1.2 Formation of the ThDP ylide 3

Figure 1.3 The archetypal catalytic mechanism of ThDP-dependent enzymes 4

Figure 1.4 Monomer of PDC illustrating the layout of the individual domains 5

Figure 1.5 Illustration of the PDC tetramer highlighting the individual monomers and the dimer of dimers 6

Figure 1.6 ThDP bound at the interface between the monomers 7

Figure 1.7 Diagram of the residues lining the ThDP binding site as well as residues involved in catalysis 8

Figure 1.8 Active site of PDC highlighting the catalytic acid/base residues 9

Figure 1.9 Regulatory binding site of PDC showing residues within 5 Å of pyruvate with Cys221 directly interacting with the substrate forming the thiohemiketal 10

Figure 1.10 The many types of reactions that can be catalyzed utilizing ThDP as a cofactor 12

Figure 1.11 The domain arrangements of the six main families of ThDP-dependent enzymes 13

Figure 1.12 Topology cartoon of the three domains of benzoylformate decarboxylase (BFDC) highlighting the structural similarities 14

Figure 2.1 The p426GPD-Zeo plasmid map indicating the location of the restriction enzyme sites 22

Figure 3.1 Reaction scheme of the conversion of phosphonopyruvate to phosphonoacetaldehyde and carbon dioxide catalyzed by phosphonopyruvate decarboxylase (PhPyrDC) 31

Figure Page

Figure 3.2 Scheme of the coupled assay used to monitor the conversion of phosphonopyruvate to ethanol....34
Figure 3.3 Michaelis-Menten plot for the reaction of PhPyrDC with phosphonopyruvate....35
Figure 3.4 The proposed mechanism of action of difluorophosphonopyruvate....36
Figure 3.5 Scheme of difluorophosphonopyruvate synthesis....37
Figure 3.6 The IC_{50} determination of DFPP at3 μM and at 30 μM....38
Figure 3.7 The titration of PhPyrDC (~50 μM) with DFPP as followed by UV-Vis spectroscopy....40
Figure 3.8 The titration of PhPyrDC (~100 uM) with DFPP as followed by CD spectroscopy....41
Figure 3.9 Fluoride ion analysis of the titration of PhPyrDC with DFPP....42
Figure 4.1 Glycolysis flowchart showing three possible paths for the breakdown of pyruvate....45
Figure 4.2 Reaction scheme of the conversion of pyruvate to ethanol utilizing pyruvate decarboxylase (PDC) and alcohol dehydrogenase (ADH)....46
Figure 4.3 Clustal W sequence alignment of PDC1 and PDC5....51
Figure 4.4 Homology model, constructed using Phyre2, highlighting the difference between PDC1 and PDC5....52
Figure 4.5 Michaelis-Menten plot for the reaction of PDC5 with pyruvate....54
Figure 4.6 Crystal of PDC5 that resulted in an X-ray structure....59
Figure 4.7 Comparison of the active sites of PDC1 and PDC5 in the open conformation from two views....61
Figure 4.8 Comparison of the active sites of PDC1 and PDC5 using the closed conformation of PDC1 from two views....62
Figure 4.9 Clustal W sequence alignment of PDC1 and PDC6....63
Figure 4.10 Clustal W sequence alignment of PDC1 and THI3....67
Figure 4.11 Overlay of THI3 and PDC1....68

Figure Page

Figure 5.1 The number of colonies needed in an NNK versus and NDT site-saturation mutagenesis experiment....75
Figure 5.2 Radar plots illustrating the activities of the variants of the SSM experiments....79
Figure 5.3 Michaelis-Menten plot for the reaction of BFDC T377L-A460Y with pyruvate....81
Figure 5.4 Structure of the expression vectors....84
Figure 5.5 The expression levels of the different promoter systems determined by monitoring β-galactosidase activity....84
Figure 5.6 A flowchart depicting the construction of the p426GPD-Zeo expression vector....87
Figure 5.7 Results from the transformation of pRUL178 into the knockout yeast strain (5.127-17C)....89
Figure 5.8 Results of the p426GPD-Zeo plasmid series transformation....90
Figure 5.9 The knockout PDC yeast was transformed with pRUL178 and was able to grow on SD-URA media....92

LIST OF ABBREVIATIONS

ADH	Alcohol dehydrogenase
AHAS	Acetohydroxyacid synthase
ALS	Acetolactate synthase
AP	4'-aminopyrimidine
BFDC	Benzoylformate decarboxylase
*Bf*PhPyrDC	*Bacteroides fragilis* phosphonopyruvate decarboxylase
BSA	Bovine serum albumin
CAST	Combinational active site saturation mutagenesis
CD	Circular dichroism
CYC1	Cytochrome-c oxidase
DFPP	Difluorophosphonopyruvate
DXPS	1-deoxy-D-xylulose 5-phosphate synthase
epPCR	Error prone polymerase chain reaction
GPD	Glyceraldehyde-3-phosphate dehydrogenase
HLADH	Horse liver alcohol dehydrogenase
IP	1', 4'-iminopyrimidine
IPTG	Isopropylthio-β-galactoside
ISM	Iterative saturation mutagenesis
LB	Luria-Bertani
Ni-NTA	Nickel-nitrilotriacetic acid
NMR	Nuclear Magnetic Resonance

PCR	Polymerase chain reaction
PDB	Protein data bank
PDC	Pyruvate decarboxylase
PDH	Pyruvate dehydrogenase complex
PFRD	Pyruvate ferredoxin oxidoreductase
Phtase	Phosphonoacetaldehyde hydrolase
PhPyrDC	Phosphonopyruvate decarboxylase
POX	Pyruvate oxidase
*Pp*BFDC	*Pseudomonas putida* benzoylformate decarboxylase
*Sc*PDC	*Saccharomyces cerevisiae* pyruvate decarboxylase
SDS-PAGE	Sodium dodecyl sulfate - polyacrylamide gel electrophoresis
SSM	Site-saturation mutagenesis
*Sv*PhPyrDC	*Streptomyces viridochromogenes* phosphonopyruvate decarboxylase
TB	Terrific broth
TEF	Translation elongation factor 1α
ThDP	Thiamin diphosphate
TISAB II	Total ionic strength adjustment buffer II
UV-Vis	Ultraviolet - visible
YADH	Yeast alcohol dehydrogenase
*Zm*PDC	*Zymomonas mobilis* pyruvate decarboxylase

ABSTRACT

Rogers, Megan P. M.S., Purdue University, May 2015. Investigation of the Evolutionary Aspects of Thiamin Diphosphate-Dependent Decarboxylases. Major Professor: Michael J. McLeish.

Thiamin diphosphate (ThDP)-dependent enzymes catalyze a wide range of reactions including the oxidative and nonoxidative decarboxylation of 2-keto acids, carboligation reactions, the cleavage of C-C bonds, and the formation of C-S, C-N, and C-O bonds. Surprisingly, given this diversity, all ThDP-dependent enzyme catalyzed reactions proceed through essentially the same intermediate. This suggests that these enzymes share a common ancestry and have evolved to become the diverse group of enzymes seen today.

Sequence alignments have revealed that all ThDP-dependent enzymes share two common ThDP binding domains, the PYR domain and the PP domain. In addition to these conserved domains, over time, other domains have been added creating further diversity in this superfamily. For instance, the TH3 domain, found in many ThDP-dependent enzymes, serves the function of binding additional cofactors such as FAD in enzymes like acetohydroxyacid synthase (AHAS) but in others, like pyruvate decarboxylase (PDC), it has lost this function completely. The work presented here focuses on ThDP-dependent decarboxylases. In this thesis,

several evolutionary aspects of this group of enzymes will be examined including (i) the characterization of an evolutionary forerunner in the presence of a mechanism-based inhibitor, (ii) the characterization of the minor isozymes of pyruvate decarboxylase from *Saccharomyces cerevisiae*, and (iii) the development of a selection method to increase the efficiency of the site-saturation mutagenesis used to study ThDP-dependent enzyme evolution.

CHAPTER 1. INTRODUCTION

1.1 Thiamin Diphosphate

Organisms require a variety of different nutrients to help maintain optimal functioning. However, not all nutrients can be synthesized and therefore must be obtained through diet. Many of these essential nutrients are known as vitamins; a deficiency in any particular vitamin can lead to disease. For instance, a deficiency in vitamin B1 or thiamin in humans can lead to beriberi, a disease that affects the cardiovascular and nervous systems. Although vitamins have different roles in the body, many, including B1, act as enzyme cofactors.

Thiamin diphosphate (ThDP), the biologically active form of vitamin B1, is utilized as a cofactor by a variety of enzymes carrying out a large range of reactions. Although ThDP is a cofactor, it also acts as a catalyst when it is not bound to an enzyme, but these reactions occur extremely slowly (1).

ThDP (Figure 1.1) consists of a pyrimidine ring (left) that is connected to a thiazolium ring (center) through a methylene carbon. The thiazolium ring is substituted with a methyl group and also a β-ethyldiphosphate group (right).

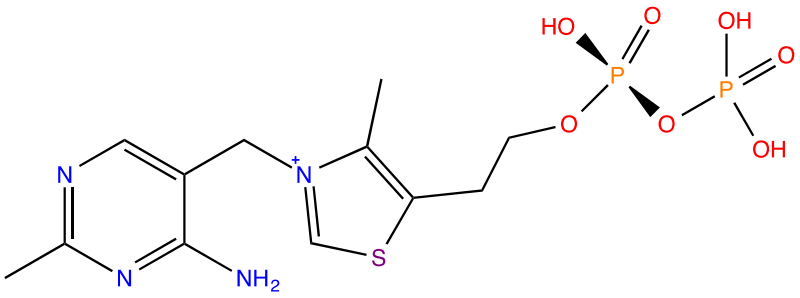

Figure 1.1. Structure of ThDP in the 4'-aminopyrimidine (AP) form.

The structure in Figure 1.1 is known as the 4'-aminopyrimidine (AP) form of ThDP. This form exists in equilibrium with the 1', 4'-iminopyrimidine (IP) form (Figure 1.2) (1). In order for catalysis to occur, ThDP must be in the activated form known as the ylide.

In all ThDP-dependent enzymes, with the exception of glyoxylate carboligase (2), there is a conserved glutamate residue that assists in the formation of the ylide (Figure 1.2). When ThDP is in the IP form, the glutamate residue abstracts a proton from the 1' nitrogen. This permits the hydrogen on the C2 atom to be abstracted by the imino group, leaving a negative charge on the C2 atom and thereby forming the ylide (1).

Figure 1.2. Formation of the ThDP ylide.

Once ThDP is activated, it can perform catalysis (Figure 1.3). The nucleophilic C2 ylide attacks the carbonyl of a 2-keto acid substrate forming 2α-lactyl-ThDP. This intermediate undergoes a rearrangement and, in the process, a molecule of carbon dioxide is expelled forming the 2α-carbanion-enamine intermediate. The addition of a proton to the carbanion forms the 2α-hydroxyethyl-ThDP. After another proton abstraction, the intermediate undergoes further rearrangement causing release of the aldehyde and regeneration of the ylide (1).

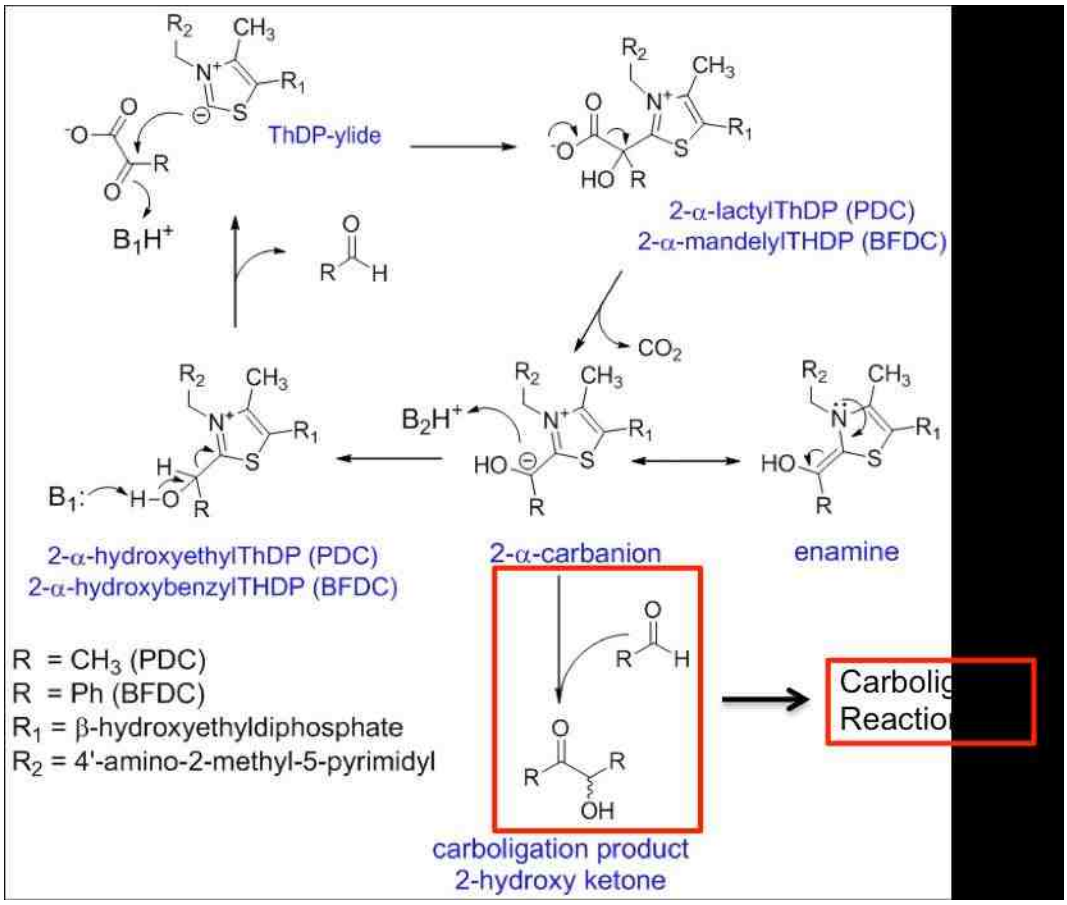

Figure 1.3. The archetypal catalytic mechanism of ThDP-dependent enzymes. Figure adapted from reference (3).

The reactions described above apply to the decarboxylation process. However, an alternative reaction, the carboligation reaction, is also possible. In order for carboligation to occur, a second molecule, such as an aldehyde, ketone, or another molecule of 2-keto acid, is required. In the carboligation reaction, the enamine-carbanion performs a nucleophilic attack on the carbonyl of the second molecule to form the carboligation product (4).

Perhaps the most studied thiamin diphosphate-dependent enzyme is pyruvate decarboxylase. This enzyme will be a major focus of this thesis.

1.2 Pyruvate decarboxylase, a modern ThDP-dependent enzyme

Pyruvate decarboxylase (PDC) (EC 4.1.1.1) is a thiamin diphosphate (ThDP)-dependent enzyme that catalyzes the conversion of pyruvate to acetaldehyde and carbon dioxide. One of the best characterized PDCs is that from baker's yeast, *Saccharomyces cerevisiae* (*Sc*PDC), which is often labeled as the “archetypal” ThDP-dependent enzyme.

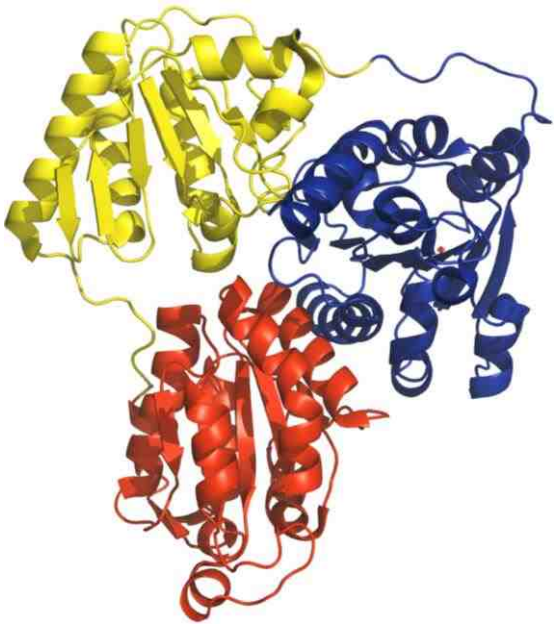

Figure 1.4. Monomer of PDC illustrating the layout of the individual domains, PYR domain (red), TH3 domain (yellow), and PP domain (blue).

PDC consists of four identical monomers of 563 amino acids. Each monomer comprises three domains (Figure 1.4), namely the N-terminal or PYR domain that binds the pyrimidine ring of ThDP, the C-terminal or PP domain that binds the diphosphate group of ThDP, and the middle or TH3 domain that is believed to be involved in controlling the conformational equilibrium between the open and closed state of the enzyme (5-7). These monomers are further

organized into a tetramer. In the case of *Sc*PDC, the minimal functional unit is a dimer. Based on this, the biological assembly is often described as a "dimer of dimers" (Figure 1.5) (8,9).

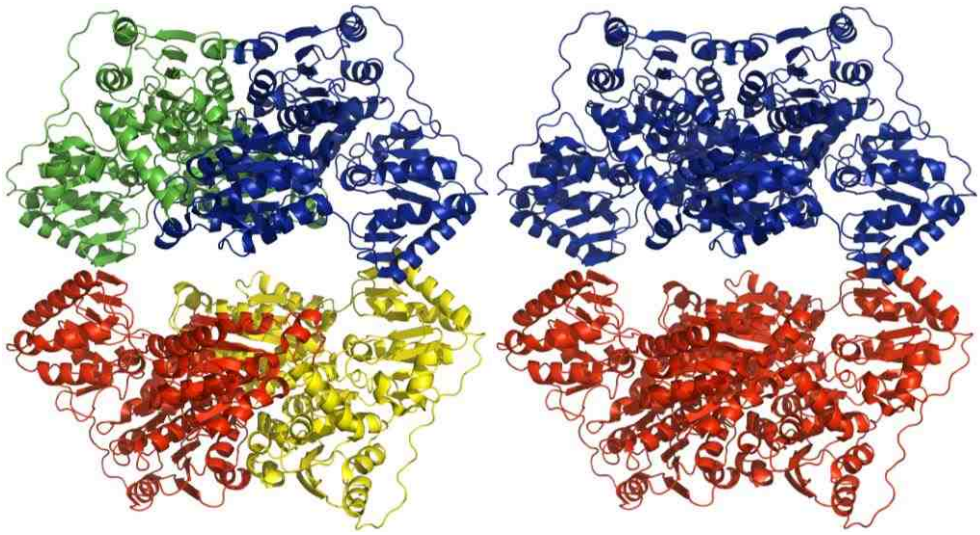

Figure 1.5. Illustration of the PDC tetramer highlighting the individual monomers and the dimer of dimers.

In any ThDP-dependent enzyme, the minimal functional unit must be a dimer because the ThDP binding site and the active site are located at the interface between the monomers. Figure 1.6 shows how ThDP is bound to the enzyme; the PYR domain of one monomer of the dimer is binding the pyrimidine ring while the PP domain of the other monomer binds the diphosphate moiety. This binding symmetry also indicates that there are two ThDP binding sites per dimer and four total in the tetramer (10).

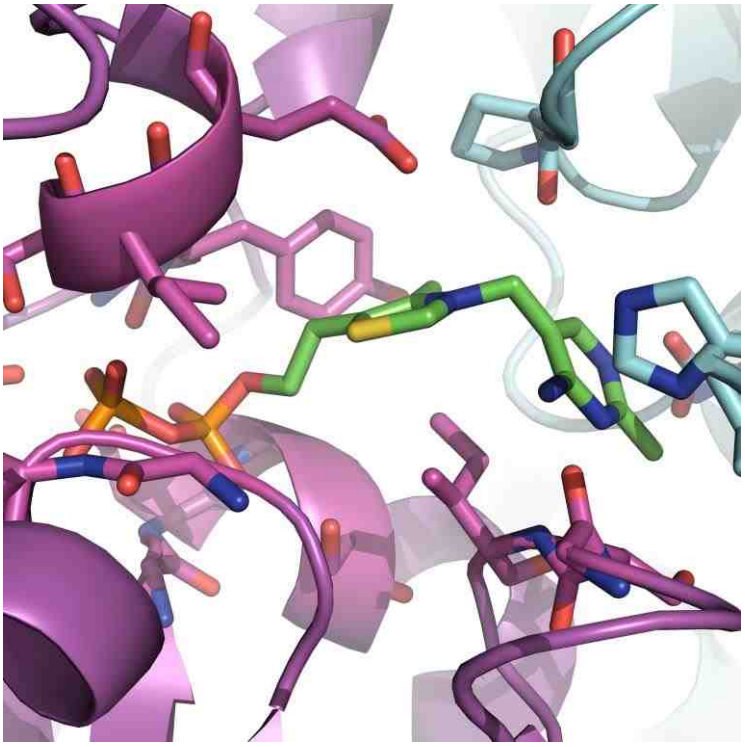

Figure 1.6. ThDP (green and orange) bound at the interface between the monomers. The PYR domain of one monomer (cyan) binds the pyrimidine ring (green) of ThDP while the PP domain of the other monomer (magenta) binds the diphosphate moiety (orange) of ThDP. Figure prepared using PyMOL using data from PDB 1PVD.

1.2.1 The ThDP binding site

A closer look into the ThDP binding site (Figure 1.7) reveals that the diphosphate group is complexed with Mg^{2+} as well as being stabilized via hydrogen bonding to the protein. It is worth noting Asp444 and Asn471; these residues are part of an amino acid sequence, $GDGX_{24\text{-}27}NN$, which has been identified as the ThDP-binding motif and is conserved throughout all ThDP-dependent enzymes (11). Another important feature of ThDP is the conformation it takes within the enzyme, known as the "V" conformation. This conformation is necessary for catalysis because it allows the 4' amino group to

be within interaction distance of the C2 reaction center (10). This conformation is seen only when ThDP is bound to the enzyme and has not been observed in free solution (1). In PDC, Ile415 stabilizes this conformation and mutations to this residue significantly affect catalysis (discussed further in Chapter 4) (12). The apparent need for a bulky hydrophobic residue to maintain this conformation is conserved throughout the ThDP-dependent enzyme family (13).

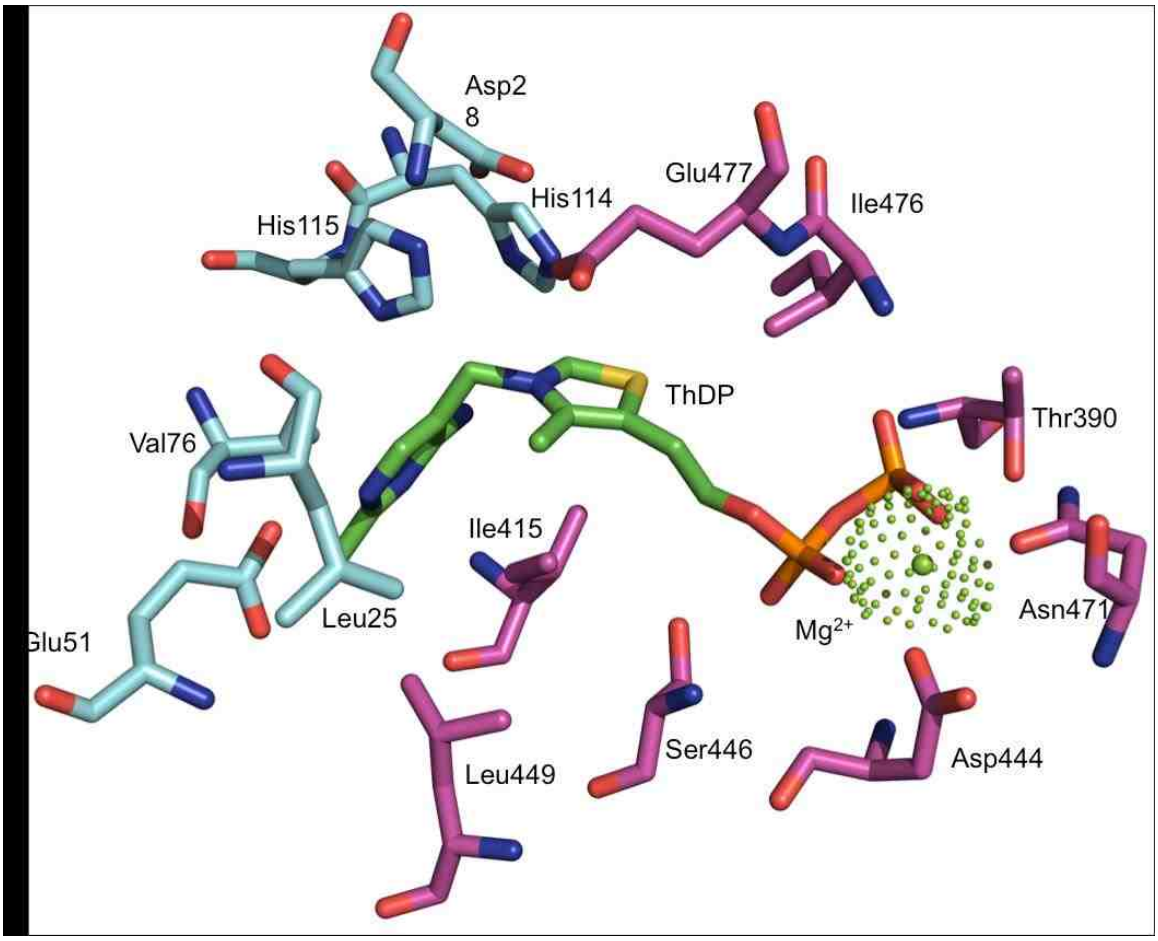

Figure 1.7. Diagram of the residues lining the ThDP binding site as well as residues involved in catalysis. Residues numbered higher than 360 belong to the PP domain of one monomer (magenta), while residues number less then 187 belong to the PYR domain of the other (cyan). Note the position of Ile415 and the V-conformation of the ThDP. Figure prepared using PyMOL using data from PDB 1PVD.

1.2.2 The active site

Like the ThDP-binding site, the active site (Figure 1.8) is also situated at the monomer/monomer interface. In PDC, there are several residues that play an important role in catalysis and stabilization of the substrate. These residues provide the ideal environment for the reaction in Figure 1.3 to take place.

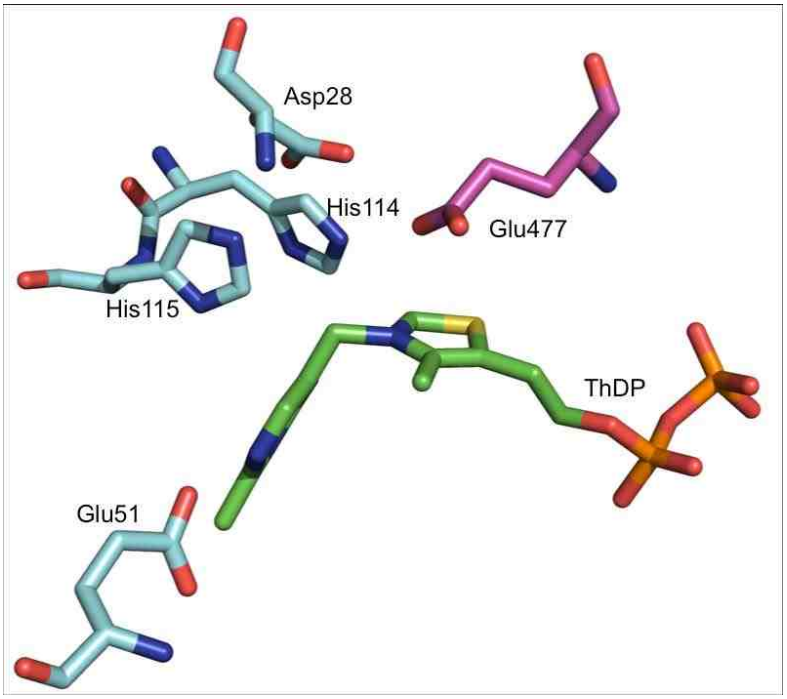

Figure 1.8. Active site of PDC highlighting the catalytic acid/base residues. The residues in magenta are from one monomer while those in cyan are from the other. Figure prepared using PyMOL using data from PDB 1PVD.

The residues that are within van der Waals radii distance of the methyl group of pyruvate include Phe292, Thr388, Ile 476, and Ile 480 (not pictured). It is likely that these residues form a hydrophobic pocket around the methyl group of pyruvate and help to position the substrate for catalysis (4). The residues potentially interacting with the glyoxylate moiety include His114, His115, Asp 28, and Glu477 (Figure 1.8). Mutation of these residues has revealed various

negative effects on k_{cat}, rather than K_m. This suggests that these residues are involved in catalysis and are not present solely for the purpose of stabilizing and positioning the substrate (14,15). Many of the ThDP-dependent enzymes possess residues that occupy similar space too those in PDC. However, the overall conservation in the active sites of ThDP-dependent enzymes is very low (4,16).

1.2.3 The regulatory site

*Sc*PDC is allosterically activated and therefore possesses a separate regulatory binding site (Figure 1.9). This binding site is not found in most other ThDP-dependent enzymes including PDCs from bacteria.

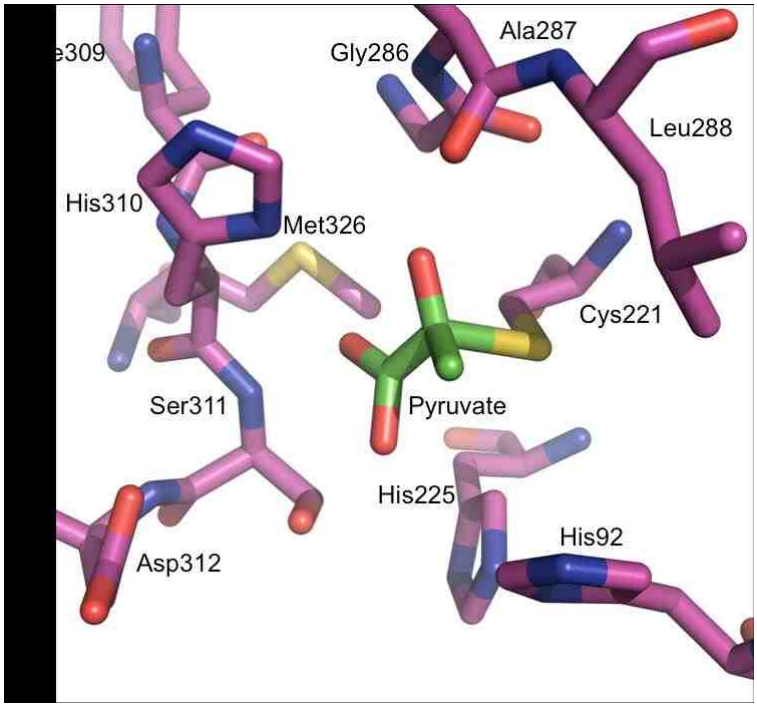

Figure 1.9. Regulatory binding site of PDC showing residues within 5 Å of pyruvate with Cys221 directly interacting with the substrate forming the thiohemiketal. Image prepared using PyMOL using data from PBD 2VK1.

The regulatory site is predominantly lined with positively charged side chains and it is thought that this feature helps to guide the activator molecule into the regulatory site (9). Baburina *et al.* (17) determined that Cys221 is the site at which substrate activation is triggered. Three residues, His92, His225, and His310 line the regulatory site and promote the formation of a thiohemiketal. This results in a conformational change around the regulatory site that leads to the restructuring of several loops in the TH3 and PYR domains. As this occurs, the C-terminal helix closes over the active site (18). The restructuring causes His114, His115, and Asp28 to shift and rotate so that the substrate can interact with the C2 ylide. Interestingly, all side chains that form the active site, with the exception of His114, His115, and Asp28, remain unaffected by the binding of the substrate molecule to the regulatory site (9,18).

Although *Sc*PDC possesses this extra binding site, it still shares many similarities with other ThDP-dependent enzymes. Given the diversity this superfamily possesses, these similarities suggest an evolutionary relationship.

1.3 Evolutionary aspects of ThDP-dependent enzymes

ThDP-dependent enzymes catalyze a wide range of reactions including the oxidative and nonoxidative decarboxylation of 2-keto acids, carboligation reactions, the cleavage of C-C bonds, and the formation of C-S, C-N, and C-O bonds (Figure 1.10). As knowledge about ThDP-dependent enzymes accumulates, it has been found that the variety of reactions catalyzed is increasing while the extent of amino acid similarity is decreasing. In spite of this

catalytic diversity, the three-dimensional structure of these enzymes remains effectively conserved (19). It is not surprising that with this influx of new information, the evolutionary history of this superfamily of enzymes is under investigation to determine how the family has evolved to catalyze so many types of reactions while still utilizing the same cofactor and same intermediate.

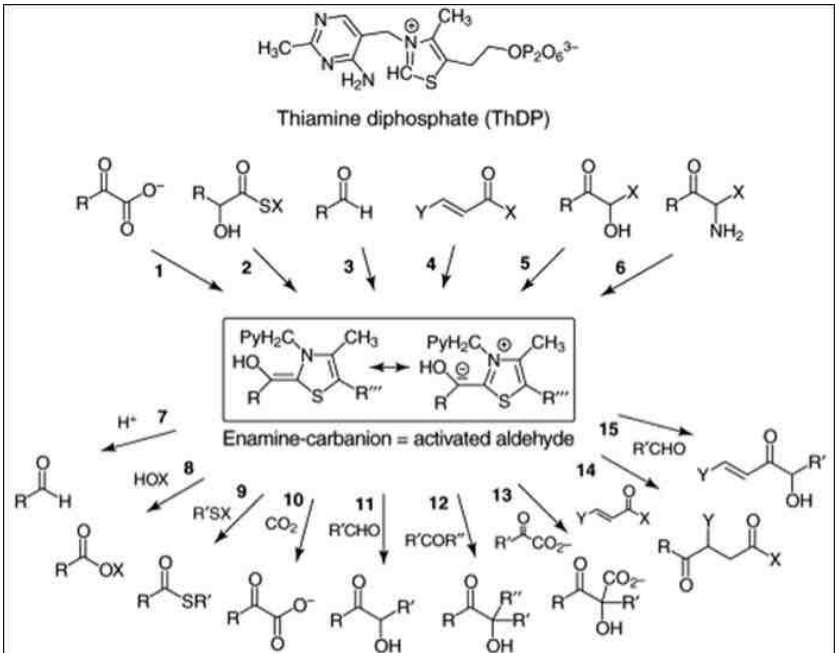

Figure 1.10. The many types of reactions that can be catalyzed utilizing ThDP as a cofactor. Note: all reactions go through the same enamine-carbanion intermediate. Figure from reference (19).

Costelloe *et al.* (20) and Duggleby (6) have performed two separate analyses in which the amino acid sequences and the crystallographic structures of many ThDP-dependent enzymes, respectively, were compared. The common feature of all ThDP-dependent enzymes is the binding of ThDP at the interface of a conserved pyrimidine (PYR) binding domain and a diphosphate (PP) binding domain (20). Based on sequence alignments of the domains, ThDP-dependent

enzymes can be grouped into six main groups (Figure 1.11). These groups include the transketolase (TK)-like family, the 2-oxoisovalerate dehydrogenase (2OXO)-like family, the pyruvate ferredoxin reductase (PFRD)-like family, the sulfopyruvate decarboxylase (SPDC)-like family, the phosphonopyruvate decarboxylase (PhPyrDC)-like family, and the pyruvate decarboxylase (PDC)-like family, which is the largest family and the primary focus of this thesis (6,20).

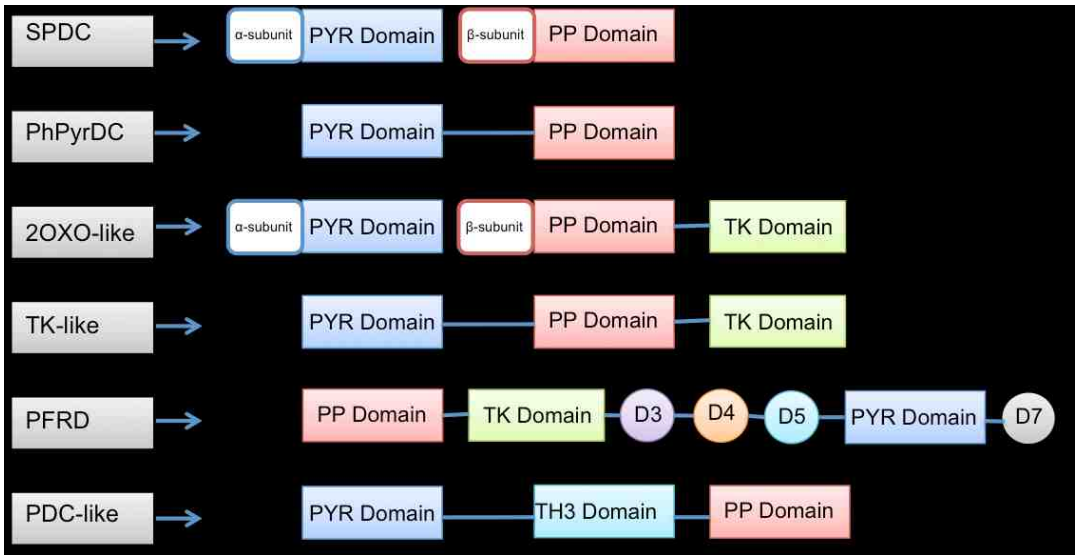

Figure 1.11. The domain arrangements of the six main families of ThDP-dependent enzymes. Adapted from reference (20).

As mentioned previously, enzymes in the PDC-like family consist of three domains: the PYR domain, the PP domain, and the TH3 domain. Each of these domains contain a central, five or six-stranded, parallel β-sheet linked by α helices and connecting loops. Structural alignments of these domains (Figure 1.12) demonstrate a high similarity in structure for the PP and PYR domains, suggesting that there is a common origin. The core structure of these two domains is highly conserved across the families (6).

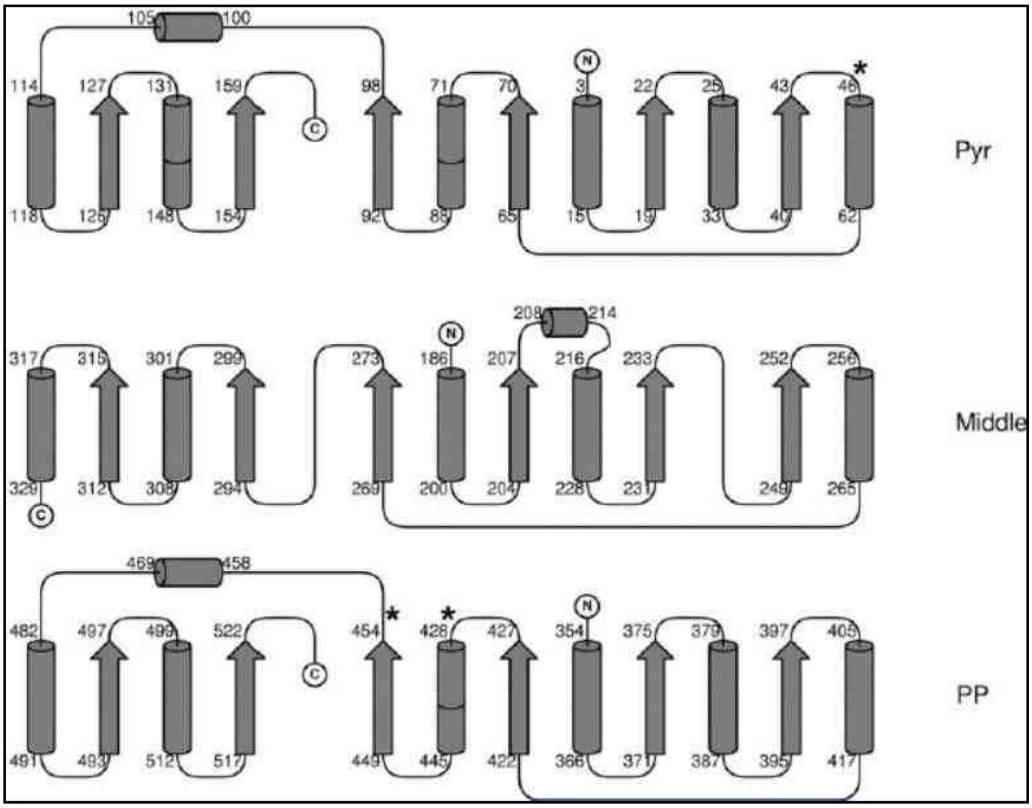

Figure 1.12. Topology cartoon of the three domains of benzoylformate decarboxylase (BFDC) highlighting the structural similarities. Figure from reference (6).

The common ancestor protein would most likely have possessed binding sites for both the pyrimidine ring and the diphosphate moiety, requiring a dimer to form in order for catalysis to occur. Today, the enzyme that most resembles this ancestor is sulfopyruvate decarboxylase, which consists of two single chains that come together to form the active site. The next logical step in the evolution process would be for these two separate chains to fuse into a single protein. Phosphonopyruvate decarboxylase is the modern enzyme that fits this description (6,20).

After the fusion into a single protein, it is probable that the TH3 domain was then also fused to the PYR and PP domains, forming the family of PDC-like enzymes. The TH3 domain is believed to be responsible for the binding of additional cofactors such as NAD^+/NADH or FAD (1). While the TH3 domain in some enzymes, such as acetohydroxyacid synthase (AHAS) and pyruvate oxidase (POX), retains this function, in others, such as PDC, this function has been lost, thereby contributing to further diversity within this group of PDC-like enzymes (6). Based on Figure 1.11, it is clear that all members of the ThDP-dependent enzyme superfamily possess the PYR and PP binding domains but differ in the extra domains that have fused over time, giving this superfamily of enzymes such a diverse range of reactions while maintaining some structural and sequential similarity. For example, PDC, AHAS, acetolactate synthase (ALS), POX, the pyruvate dehydrogenase complex (PDH), pyruvate ferredoxin oxidoreductase (PFRD), and 1-deoxy-D-xylulose 5-phosphate synthase (DXPS) all utilize pyruvate as a substrate. However, these six enzymes fall into four different groups that pass through the same intermediate to yield six different products. This example further highlights the complexity and diversity of this enzyme superfamily.

1.4 Research covered in this work

The research covered in this thesis maintains the theme of studying different evolutionary aspects of ThDP-dependent decarboxylases. In Chapter 3, the activity of phosphonopyruvate decarboxylase, a primitive member of the

ThDP-dependent enzyme superfamily, is examined. Chapter 4 covers the investigation of the potential isozymes of pyruvate decarboxylase from *Saccharomyces cerevisae*. Finally, in Chapter 5, an alternative selection method is investigated for site-saturation mutagenesis to carry out directed evolution of ThDP-dependent decarboxylases.

CHAPTER 2. MATERIALS AND METHODS

2.1 Materials

Primers were synthesized by Integrated DNA Technologies (Coralville, IA) and were designed to use the QuikChange protocol from Agilent Technologies (Santa Clara, CA). *Pfu* Ultra was purchased from Stratagene (La Jolla, CA). Restriction enzymes were purchased from New England BioLabs (Ipswich, MA) and Thermo Fisher Scientific (Waltham, MA). The Champion pETSUMO expression system was purchased from Life Technologies (Grand Island, NY). DNA isolation from *E. coli* was performed utilizing the Zyppy Plasmid Miniprep Kit from Zymo Research (Irvine, CA) or the AxyPrep Plasmid Midiprep Kit from Corning (Tewksbury, MA). Sequencing was carried out at the University of Michigan DNA Sequencing Core Facility (Ann Arbor, MI). YEP broth, YEP agar, and SD Medium-URA were purchased from MP Biomedicals (Santa Ana, CA). Difco™ YPD broth, YPD agar, Luria-Bertani (LB) broth, and LB agar were purchased from Becton, Dickinson and Company (Sparks, MD). Knockout yeast strains were provided by Prof. Jack Pronk (Delft University of Technology, Netherlands) and Prof. Stefan Hohmann (University of Gothenburg, Sweden). Nickel-nitrilotriacetic acid (Ni-NTA) resin was purchased from Qiagen (Valencia CA). A plasmid containing the recombinant alcohol dehydrogenase from *Equus*

caballus was from Prof. Bryce Plapp (University of Iowa, Iowa City, IA) and the pET17HLADH-His plasmid was constructed by Dr. Malea M. Kneen (Indiana University-Purdue University Indianapolis, IN). IPTG, lysozyme, DNaseI, NADH, yeast alcohol dehydrogenase (YADH), and 2-keto acids were purchased from Sigma-Aldrich (St. Louis, MO). Buffers and other reagents were purchased from Sigma-Aldrich or Fisher Scientific (Pittsburgh, PA). Phosphonopyruvate, difluorophosphonopyruvate, and phosphonoacetaldehyde were synthesized by Katharina Pallitsch in the laboratory of Friedrich Hammerschmidt (University of Vienna, Austria). Hampton crystallization screening kit 1 and 2, crystal trays, and cover slips were purchased from Hampton Research (Aliso Viejo, CA). Wizard Classic Screening kits 1-4 were purchased from Rigaku Reagents (Bainbridge Island, WA). Homology models were constructed using the Phyre2 Protein Fold Recognition Server (21). Images were made using PyMOL (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.).

2.2 Transformations

2.2.1 *Escherichia* coli transformation

E. coli transformations were performed using TOP10 or BL21(DE3) chemically competent cells as required. Generally, a 50 µL aliquot of competent cells was allowed to thaw on ice. Once thawed, 5 µL of DNA was added to the cells and the mixture was kept on ice for 30 minutes. The mixture was heat shocked at 42 °C for 1 min 15 s and then cooled for 2 min on ice. LB media (250 µL) was added to the mixture and the cells were allowed to shake (225 rpm) at

37 °C for 30-60 minutes. The cells (250 µL) were then spread onto LB plates containing the appropriate antibiotic. The plates were incubated overnight at 37 °C.

2.2.2 *Saccharomyces cerevisiae* transformation

S. cerevisiae transformations were performed in the knockout strains listed in Table 2.1 using a modification of the Chen *et al.* protocol (22). Competent yeast cells were prepared at the time of transformation. The yeast strains were streaked from frozen stocks onto YEP plates containing 2% EtOH and were incubated at 30 °C for 3 days. A single large colony was selected and resuspended in 1 mL of One Step Transformation Buffer (0.2 M LiAc, 40% (w/v) PEG 4000, 100 mM DTT). This volume was sufficient for ten transformations. The transformation mixture contained 100 µL of cells, 50 – 1000 ng of transforming DNA, and ~15 µg of single stranded salmon sperm DNA (carrier DNA). The mixture was vortexed and incubated at 45 °C for 20 min followed by a 30 min resting period on ice. The transformation was plated directly onto the appropriate selective plate and incubated at 30 °C for 3-4 days.

Table 2.1. *S. cerevisiae* strains used in this study.

Strain	**Genotype**	**Source**
5.127-17C	*a leu2-3/112 ura3-52 trp1-92 pdc1D::LEU2 pdc5D::URA3 pdc6D::TRP1 GAL*	(23)
CEN.PK711-7C	*MATa ura3-52 pdc1Δ pdc5Δ pdc6Δ aro10::loxP-kan-loxP thi3::loxP-kan-loxP*	(24)
IMZ001	*MATa ura3-52 pdc1Δpdc5Δpdc6Δaro10::loxP-kan-loxP thi3::loxP-kan-loxP p426GPD (URA3)*	(24)
IMZ002	*MATa ura3-52 pdc1Δpdc5Δpdc6Δaro10::loxP-kan-loxP thi3::loxP-kan-loxP pUDe001 (URA3TDH3pr-ARO10-CYC1ter)*	(24)

2.3 Construction of expression vectors

2.3.1 Cloning and construction of expression vectors pET17PDC5-His, pET17PDC6-His, and pET17THI3-His

Dr. Malea M. Kneen completed the cloning and amplification of the ScPDC5, ScPDC6, and ScTHI3 genes into the pCRBLUNT vector. These plasmids, pCRBLUNT-PDC5, pCRBLUNT-PDC6, and pCRBLUNT-THI3, were digested with NdeI/XhoI and ligated into a pET17b vector, also digested with NdeI and XhoI. These ligations provided the expression vectors pET17PDC5-His, pET17PDC6-His, and pET17THI3-His.

2.3.2 Construction of a zeocin yeast expression vectors

The p426GPD-Zeo expression vector was constructed via the digestion of the zeocin cassette vector, pTEF1/Zeo (Life Technologies) and p426GPD high expression yeast shuttle vector (25). The p426GPD expression vector first underwent PCR mutagenesis to introduce the SphI restriction site using the primers listed in Table 2.2 and the parameters listed in Table 2.3 and Table 2.4. This resulted in a plasmid called p426GPD-SphI. The p426GPD-SphI vector was digested with SphI/ApaI and ligated with pTEF1/Zeo, also digested with SphI/ApaI. This digestion and ligation resulted in a plasmid called p426GPD-Zeo. The p426GPD-Zeo expression vector was digested with BamHI and religated. This removed one BamHI site and one EcoRI site to form the plasmid p426GPD-Zeo-BamHI. PCR mutagenesis was also done on the p426GPD-Zeo-BamHI expression vector to introduce a HindIII restriction site (Table 2.2) resulting in the p426GPD-Zeo-HindIII plasmid.

Table 2.2. Forward primers used to introduce new restriction sites. Complimentary primers were synthesized based on the forward sequence.

Primer	Sequence
p426GPD-SphI	5'-GTAAATGCATGCATACTAAACTCACAAATTAGAGC
p426GPD-Zeo-HindIII	5'- ACCTTGCTTGAGAAGCTTTTGGGACGCTCGAAGGC

Table 2.3. Site-directed mutagenesis reaction mixture.

Amount	Component
5 µL	10X Ultra AD Reaction Buffer: 100 mM KCl, 100 mM (NH4)2SO4, 200 mM Tris-HCl, 200 mM MgSO4, 1% Triton X-100, 1 mg/ml BSA
1.5 µL	dNTP mix (40 mM)
125 ng	Forward oligonucleotide primer
125 ng	Reverse oligonucleotide primer
50 ng or 200 ng	Template DNA
Up to 50 µL	ddH_2O
1 µL	*Pfu* Ultra High Fidelity DNA Polymerase (2.5 U $\bullet mL^{-1}$)

Table 2.4. Site-directed mutagenesis temperature cycling parameters. Over the first five cycles, the annealing temperature increases from 45 ˚C to 55 ˚C in 2 ˚C increments.

Cycles	Temperature (˚C)	Time (s)
Initial Denaturing	95	120
1-5	95	30
	45	30
	72	360
6-25	95	30
	55	30
	72	360
Final Extension	72	600
Hold	4	∞

2.3.3 Insertion of genes into zeocin yeast expression vector

The genes listed in Table 2.5 were inserted into the zeocin yeast expression vector using the restriction enzymes listed. Figure 2.1 provides a vector map illustrating where the genes were inserted.

Table 2.5. The zeocin yeast expression vector with gene inserts and the restriction enzymes used.

p426GPD-Zeo Expression Vectors	**Restriction Enzymes**
p426GPD-Zeo-PDC1-His	BamHI/MluI
p426GPD-Zeo-BFDC-His	BamHI/HindIII
p426GPD-Zeo-PDC5-His	NdeI/XhoI (into p426GPD-Zeo-PDC1-
p426GPD-Zeo-ZmPDC	BamHI/EcoRI
p426GPDBDFC-Zeo-A460YT377L-	BamHI/HindIII

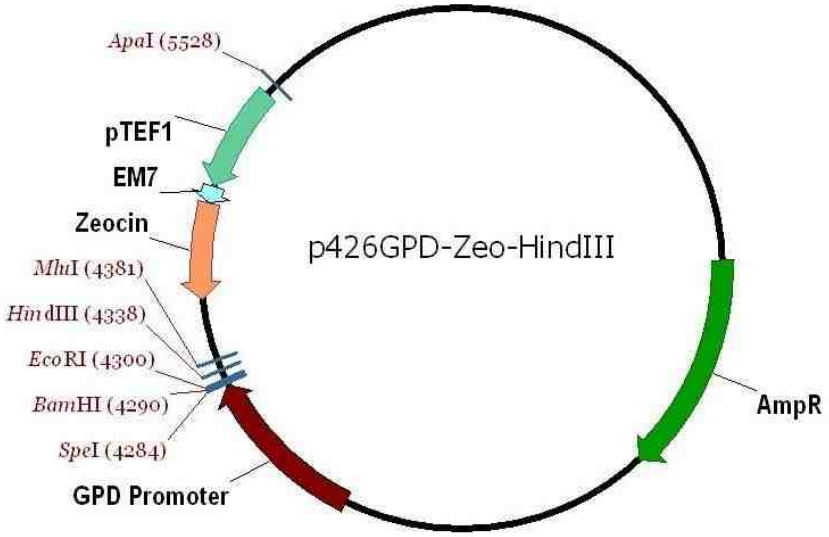

Figure 2.1. The p426GPD-Zeo plasmid map indicating the location of the restriction enzyme sites. The red arrow represents the GPD promoter.

2.4 Protein expression and purification of His-tagged enzymes

2.4.1 Expression of His-tagged enzymes

The expression of all proteins was carried out in BL21 (DE3) competent cells. The starter cultures (50 mL for a 1 L culture) for the expression of PDC1 and PDC5 were grown overnight at room temperature while the starter cultures for all other proteins were grown at 37 °C. The 1 L cell cultures were allowed to grow at 37 °C until an OD_{600} of 0.6-0.8 was reached, at which point the cultures were chilled on ice for 10 min. Expression was induced by the addition IPTG to a final concentration of 1 mM. The cultures were chilled on ice for an additional 10 min. The cell cultures were grown at room temperature overnight (18-20 h) with shaking at 225 rpm.

The cells were harvested by centrifugation at 4 °C and the cell pellet was resuspended in 30 mL of Buffer 1 (50 mM $NaPO_4$, 300 mM NaCl, and 5 mM imidazole, pH 8.0). The cells were stored at -80 °C until required.

2.4.2 General purification of His-tagged proteins

Cells were thawed and lysozyme (200 μg/mL) and DNaseI (5 μg/mL) were added. The cells were incubated on ice for 30 minutes. For PDC1 and PDC5, PMSF (1 mM) was also added. Cell lysis was completed by sonication (2 x 90 s burst at 30% duty cycle with 90 s of rest between bursts). The resulting suspension was centrifuged twice for 30 min at 20,000 × g to remove all cell debris.

The cell free extract was loaded onto an Ni-NTA column (1.0×10 cm) previously equilibrated with Buffer 1 using a Biologic LC system (Bio-Rad, Hercules, CA). After loading, the column was washed with 20 column volumes of Buffer 1 and 8 column volumes of Buffer 2 (50 mM $NaPO_4$, 300 mM NaCl, and 20 mM imidazole, pH 8.0). Protein was eluted with 8 column volumes of Buffer 3 (50 mM $NaPO_4$, 300 mM NaCl, and 250 mM imidazole, pH 8.0) and fractions (5 mL) were collected.

The fractions containing protein were analyzed by SDS-PAGE to determine which fractions contained the desired protein and to assess the level of purity. The cleanest fractions were pooled and concentrated to 3 mL (Amicon Ultra, Millipore, Billerica, MA). The protein was desalted and the buffer exchanged to a storage buffer using an Econo-Pac DG10 desalting column (Bio-Rad) that had been previously equilibrated with the appropriate buffer. Generally, the storage buffer contained 100 mM KPO_4, 1 mM $MgSO_4$, 0.5 mM ThDP, pH 6.0, and 10% (v/v) glycerol. Protein concentrations were determined from the absorbance at 280 nm using the molar extinction coefficients listed in Table 2.6. Concentrations were confirmed by Bradford assay using BSA as the protein standard (26). Proteins were stored at -80 °C.

Table 2.6. Molar extinction coefficients used in calculating protein concentration calculated using Northwestern University's Peptide Calculator 1.00 (http://www.basic.northwestern.edu/biotools/proteincalc.html, accessed February 15, 2015).

Protein	**Extinction Coefficient ($M^{-1} cm^{-1}$)**
PDC1	62070
PDC5	63350
BDFC-A460YT377L	61030
PhPyrDC	21650
Phtase	33240

2.4.3 Purification of phosphonopyruvate decarboxylase and phosphonoacetaldehyde hydrolase

The plasmid containing recombinant phosphonopyruvate decarboxylase (PhPyrDC) from *Bacteroides fragilis* was a gift of Dr. Debra Dunaway-Mariano (University of New Mexico, Albuquerque, NM) and the pET24PhPyrDC-His plasmid was constructed by Dr. Forest H. Andrews (University of Colorado Denver, CO). The plasmid (pET22Phtase-His) containing a recombinant phosphonoacetaldehyde hydrolase (Phtase) from *Pseudomonas putida* was a gift of Dr. Georg A. Sprenger (University of Stuttgart, Germany). Each of these proteins were expressed and purified as previously described except the storage buffer contained 50 mM HEPES, 300 mM NaCl, 10 mM $MgCl_2$, and 1 mM DTT. PhPyrDC was stored at -80 °C while Phtase was stored at 4°C.

2.5 Protein crystallization

2.5.1 Protein expression and purification for crystallization

For crystallization, proteins were expressed and purified in the manner previously described. Proteins were concentrated and the elution buffer (Buffer 3) was exchanged for crystallization buffer (25 mM HEPES, pH 7.0, 0.2 mM ThDP, 0.1 mM $MgCl_2$) using an Econo-Pac DG10 desalting column equilibrated with crystallization buffer. The storage buffer for PhPyrDC contains HEPES so it was purified for crystallization in the same manner as a general purification.

2.5.2 Crystallization of PDC5

PDC5 was concentrated to 15 mg/mL and crystals were grown using the hanging drop diffusion method. The wells contained 200 µL of 2 M ammonium sulfate and 100 mM sodium citrate/citric acid, pH 5.5. The drop mixture contained equal volumes of protein and well solution and was mixed on the cover slide. The crystals were grown at 4 °C. A white precipitate formed overnight and visible crystals appeared within 1-2 days. The crystals were transferred to a cryogenic solution of well buffer mixed with 15% (v/v) glycerol. The crystals were mounted on Hampton CryoLoops and flash frozen in liquid nitrogen.

2.6 Characterization studies

2.6.1 Kinetic characterization of PDC1 and PDC5

The activity of purified PDC1 and PDC5 was determined using a coupled assay. The assay mixture contained, in a total of 1 mL, 50 mM MES (pH 6.0), 2 mM $MgCl_2$, 1 mM ThDP, 0.2 mM NADH, 60 U/mL YADH or 0.25 U/mL HLADH (YADH for straight chain aliphatic substrates), and varying concentrations of the 2-keto acid. In some cases, the MES buffer was replaced by a phosphate buffer (100 mM KPO_4 pH 6.0, 1 mM $MgCl_2$, 0.5 mM ThDP) or an imidazole buffer (40 mM Imidazole-HCl pH 6.0, 2 mM $MgCl_2$, 1 mM ThDP). The reaction was carried out at 30 °C and initiated by the addition of PDC1 or PDC5. The reaction was monitored by a decrease in absorbance at 340 nm. Each assay was performed in triplicate. The kinetic parameters were determined by fitting the initial rate data to the Hill equation (Equation 1) using the enzyme kinetics package from SigmaPlot 12.3.

Equation 1:

$$\nu = \frac{V_{max}[S]^n}{K_H^n + [S]^n}$$

2.6.2 Kinetic characterization of the benzoylformate decarboxylase T377L_A460Y variant

The activity of purified BFDC T377L-A460Y was determined using a coupled assay. The assay mixture contained, in a total of 1 mL, 100 mM KPO_4 (pH 6.0), 1 mM $MgCl_2$, 0.5 mM ThDP, 0.2 mM NADH, 60 U/mL YADH or 0.25 U/mL

HLADH (YADH for straight chain aliphatic substrates), and varying concentrations of the 2-keto acid. The reaction was carried out at 30 °C and initiated by the addition of BFDC T377L-A460Y. The reaction was monitored by a decrease in absorbance at 340 nm. Each assay was performed in triplicate. The kinetic parameters were determined by fitting the initial rate data to the Michaelis-Menten equation (Equation 2) using the enzyme kinetics package from SigmaPlot 12.3.

Equation 2:

$$v = \frac{V_{max}[S]}{K_m + [S]}$$

2.6.3 Steady state analysis of phosphonopyruvate decarboxylase

The activity of purified PhPyrDC was determined using a coupled assay. The assay mixture contained, in a total of 1 mL, 100 mM HEPES (pH 7.0) 5 mM $MgCl_2$, 0.2 mM NADH, 60 U/mL YADH, 100 μg phosphonoacetaldehyde hydrolase, and varying concentrations of phosphonopyruvate (0.5-15 μM). The reaction was carried out at 30 °C and initiated by the addition of PhPyrDC. Each assay was performed in triplicate. The kinetic parameters were determined by fitting the initial rate data to the Michaelis-Menten equation (Equation 2) using the enzyme kinetics package from SigmaPlot 12.3.

2.6.4 IC_{50} and K_i determination for difluorophosphonopyruvate

PhPyrDC was assayed in the presence of difluorophosphonopyruvate (DFPP) to determine the IC_{50} of the inhibitor. The standard assay was used but the concentration of the substrate, phosphonopyruvate, was held constant at 3 μM (K_m) or at 30 μM (10×K_m) while the inhibitor concentration was varied. The IC_{50} was determined by plotting the concentration of inhibitor (μM) against the percentage of activity remaining. The K_i was determined from the IC_{50} value using the Cheng-Prusoff equation (Equation 3).

Equation 3:

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

2.6.5 Reaction of difluorophosphonopyruvate with PhPyrDC monitored by circular dichroism (CD) spectroscopy and UV-VIS spectroscopy

PhPyrDC (~50 μM), in a 2 mL volume of PhPyrDC storage buffer, was titrated with DFPP at concentrations ranging from 0-50 μM. DFPP was introduced in 10 μL aliquots with each addition resulting in a 5 μM increase in inhibitor concentration. Spectra were obtained using a 1 cm path length cell. After the addition of each of the inhibitor aliquots, the solution was gently mixed and allowed to sit for 3 minutes at 25˚C. Initially, the reaction mixtures were monitored on a Cary 50 Bio UV-Vis spectrophotometer at 440 nm to ensure the reaction was complete. The reaction mixture was then scanned from 550 – 250 nm. The A_{440} titration curves were constructed using SigmaPlot 12.3.

This same procedure was repeated using the Jasco J-810 CD spectropolarimeter. The reaction mixture was monitored at 440 nm on a UV-VIS spectrophotometer after each addition of inhibitor to ensure the reaction was complete. The mixture was then scanned from 550 - 250 nm on the CD spectropolarimeter. The data obtained for each inhibitor concentration was the average of five scans. The titration curves for the change in ellipticity at 440 nm were constructed using SigmaPlot 12.3.

2.6.6 Fluoride ion analysis of the reaction of difluorophosphonopyruvate with PhPyrDC

The fluoride ions expected to be released upon reaction of DFPP with the ThDP of PhPyrDC were monitored using a fluoride ion-selective electrode. The samples contained PhPyrDC (~100 μM) and DFPP (0-75 μM) in a total volume of 1 mL. In addition to samples containing enzyme, two controls were used. The first contained PhPyrDC buffer and water, and the second contained PhPyrDC buffer, 75 μM difluorophosphonopyruvate, and water. Both solutions had a total volume of 1 mL. The samples were prepared by gentle mixing and were incubated at room temperature for 30 min. The enzyme was removed from the samples using 10,000 kDa cut-off microspin columns (Microcon, Millipore). Total ionic strength adjustment buffer II solution (TISAB II) was added to the samples for a final volume of 2 mL. The fluoride ion concentration in each sample was determined. This analysis was carried out in the Lippert laboratory at the Indiana University School of Dentistry.

CHAPTER 3. INACTIVATION OF PHOSPHONOPYRUVATE DECARBOXYLASE FROM *BACTEROIDES FRAGILIS* BY DIFLUOROPHOSPHONOPYRUVATE

3.1 Background

Phosphonopyruvate decarboxylase (PhPyrDC) catalyzes the breakdown of phosphonopyruvate to phosphonoacetaldehyde and carbon dioxide (Figure 3.1). As noted in the introduction, PhPyrDC is thought to closely resemble the fused domain common ancestor of all ThDP-dependent enzymes (6,20). Despite its potential importance, PhPyrDC has not been well studied and only a few aspects have been investigated. To date, only the PhPyrDCs from *Bacteroides fragilis* (*Bf*) and *Streptomyces viridochromogenes* Tü494 (*Sv*) have been characterized (27,28).

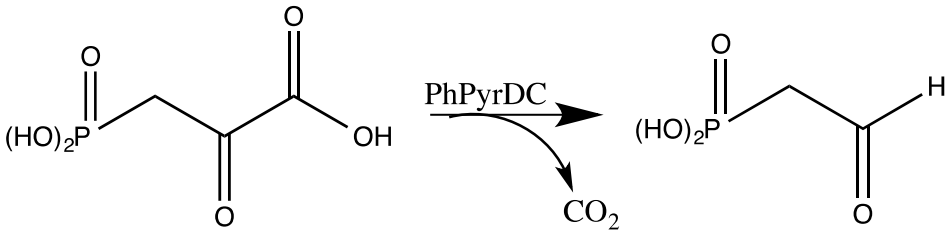

Figure 3.1. Reaction scheme of the conversion of phosphonopyruvate to phosphonoacetaldehyde and carbon dioxide catalyzed by phosphonopyruvate decarboxylase (PhPyrDC).

*Bf*PhPyrDC was the first to be characterized by Zhang *et al*. (28). Based on sequence analysis, the group determined the molecular weight of *Bf*PhPyrDC to be ~40 kDa and confirmed this experimentally. Johnen *et al.* (27), whose studies focused on *Sv*PhPyrDC, determined the molecular weight to be ~44 kDa. Interestingly, the oligomeric state of *Sv*PhPyrDC was determined to be homodimeric while that of *Bf*PhPyrDC was found to be homotrimeric (27,28). To date, no crystal structure of PhPyrDC has been solved.

Using sequence alignments, both groups compared PhPyrDC to variants from other organisms, as well as other ThDP-dependent enzymes, to determine which residues were conserved. The conserved active site glutamate residue (Chapter 1) is also found in PhPyrDCs from both organisms. Johnen *et al.* (27) carried out mutagenesis at this site (Glu 48) to confirm the residue's importance in catalysis. They prepared the Glu48Gln variant and found that, while it expressed to the same extent as wild-type, it was devoid of activity. Not surprisingly, both PhPyrDCs possessed the ThDP binding motif ($GDGX_{24\text{-}27}NN$) and mutagenesis was performed to confirm the essential nature of these residues (27,28). Furthermore, all PhPyrDCs used in the sequence alignment by Johnen *et al.* (27) possessed a conserved serine residue at position 25 (*Sv*PhPyrDC numbering). In some ThDP-dependent enzymes, such as benzoylformate decarboxylase, there is a serine in this position while in enzymes such as pyruvate decarboxylase, it is an aspartate (16). Johnen *et al.* (27) used mutagenesis to explore the function of this residue and determined that Ser25 is important in substrate binding. *Sv*PhPyrDC also possesses the

characteristic HH-motif (His110 and His111) seen in PDC-like enzymes. Mutations to His110 resulted in a loss of activity, but changes to His111 had no significant influence on catalysis (27). Johnen *et al.* (27) also determined that Asp297 played an important role in substrate binding.

Zhang *et al.* (28) also performed mutagenesis on *Bf*PhPyrDC to confirm the residues important to catalysis and ThDP binding. However, they also determined the kinetics for pyruvate as a substrate in addition to phosphonopyruvate. For phosphonopyruvate, the K_m was found to be 3.2 ± 0.2 μM and the k_{cat} to be 10.2 ± 0.3 s^{-1}. Pyruvate was a slow substrate with a K_m of 25 mM and a k_{cat} of 0.05 s^{-1}. In addition, sulfopyruvate and phosphonoacetaldehyde were shown to be competitive inhibitors of *Bf*PhPyrDC (28). Given that (1) sulfopyruvate is structurally similar to phosphonopyruvate, (2) phosphonoacetaldehyde possesses the "phosphono" group but not the carboxylate moiety, and (3) the low activity with pyruvate, Zhang *et al.* (28) concluded that, rather than the carboxylate moiety, the "phosphono" group must be of primary significance in substrate binding (28).

While, sequence alignments have been helpful in identifying some important residues, a crystal structure will aid in obtaining a more complete picture of this enzyme. In addition, it will answer the question of the enzyme's oligomeric state. Here, in order to further the understanding of this enzyme, PhPyrDC was characterized in the presence of a mechanism-based inhibitor. In addition, crystallization trials were carried out.

3.2 Results and Discussion

3.2.1 Purification and kinetic characterization of PhPyrDC

PhPyrDC from *Bacteroides fragilis* was expressed and purified as described in section 2.4 using a plasmid that was a gift from Dr. Debra Dunaway-Mariano (University of New Mexico). The steady state kinetics of *Bf*PhPyrDC were determined using a modified coupled assay. In a traditional coupled assay for ThDP-dependent decarboxylases, once the aldehyde is released it becomes the substrate for a yeast alcohol dehydrogenase (YADH) that converts it to an alcohol with the concomitant oxidation of NADH. The reaction is monitored by observing the conversion of NADH to NAD^+, which results in a loss of absorbance at 340 nm. In the case of PhPyrDC, the aldehyde that is released, phosphonoacetaldehyde, is not a substrate for YADH and, in order for the reaction to continue, the phosphono group must be removed. Phosphonoacetaldehyde hydrolase (Phtase) removes that group so the coupled assay can continue (Figure 3.2). Accordingly, Phtase was also expressed and affinity purified using a plasmid that was provided by Dr. Georg A. Sprenger (University of Stuttgart).

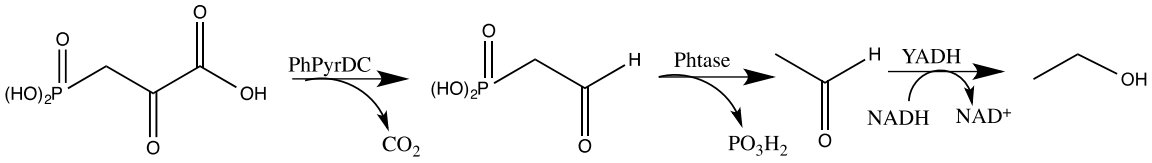

Figure 3.2. Scheme of the coupled assay used to monitor the conversion of phosphonopyruvate to ethanol.

PhPyrDC exhibited Michaelis-Menten type kinetics (Figure 3.3) with its natural substrate, phosphonopyruvate. PhPyrDC did not exhibit any allosteric activation like that seen with yeast PDC (29). The K_m was determined to be 2.9 ± 0.2 μM with a k_{cat} value of 10.1 ± 0.3 s^{-1}. This provides a k_{cat}/K_m of 3500 $mM^{-1}s^{-1}$. By comparison, yeast PDC has a k_{cat}/K_m value of 45 $mM^{-1}s^{-1}$, which is almost 80-fold lower than that of PhPyrDC, and PDC from *Zymomonas mobilis* has a k_{cat}/K_m value of 440 $mM^{-1}s^{-1}$, which is 8-fold lower (4). In addition, pyruvate was also tested as a substrate but no activity was detected under the conditions of our assay. Overall, these values are in good agreement with those determined by Johnen *et al*. (27).

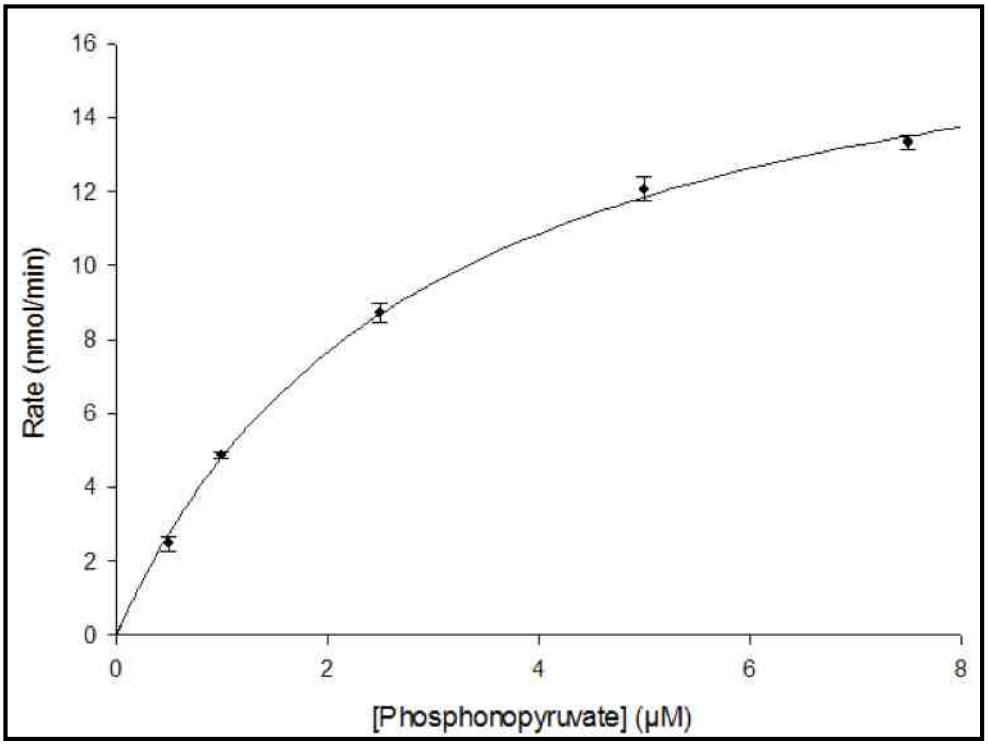

Figure 3.3. Michaelis-Menten plot for the reaction of PhPyrDC with phosphonopyruvate.

3.2.2 Design and synthesis of difluorophosphonopyruvate

PhPyrDC has not been studied in depth and much of the information gathered about the enzyme is obtained by comparisons with other ThDP-dependent decarboxylases. This includes a prediction of the enzyme's catalytic mechanism. It would be reasonable to assume that the catalytic mechanism is similar to the mechanism in Figure 1.3 seen the introduction. It has been shown that fluoropyruvate can be broken down by pyruvate decarboxylase and that a fluoride ion is released along with the generation of acetyl-ThDP (30). With that knowledge, a mechanism-based inhibitor, difluorophosphonopyruvate (DFPP), was designed based on the mechanism in Figure 3.4.

Figure 3.4. The proposed mechanism of action of difluorophosphonopyruvate.

The proposed mechanism of action of DFPP involves the initial binding of the inhibitor to ThDP. The decarboxylation step proceeds as normal and the resultant 2α-carbanion undergoes rearrangement to lose a fluoride ion. However, unlike the fluoropyruvate reaction with PDC, there is another fluorine

attached to C2, making any further rearrangements very slow, and thereby rendering the enzyme effectively inactive.

DFPP was synthesized using the scheme in Figure 3.5 by Katharina Pallitsch in the laboratory of Friedrich Hammerschmidt (University of Vienna).

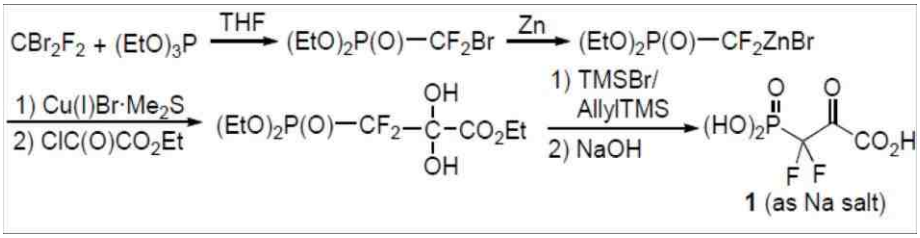

Figure 3.5. Scheme of difluorophosphonopyruvate synthesis (1) (31).

To confirm this proposed mechanism, a series of experiments was carried out.

3.2.3 IC_{50} and K_i determination for DFPP

In a control experiment, the inhibitor was tested against the two coupling enzymes, Phtase and YADH. It was determined that DFPP did not inhibit the activity of these enzymes. However, it quickly became apparent that DFPP was a good inhibitor of PhPyrDC.

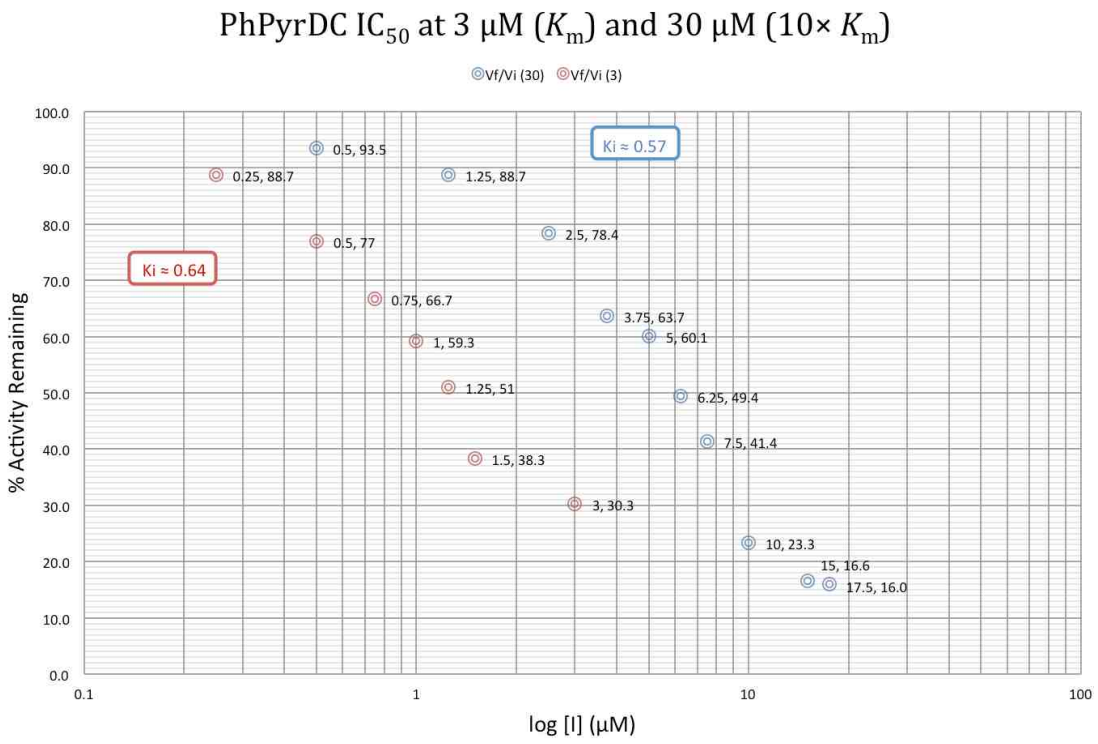

Figure 3.6. The IC_{50} determination of DFPP at 3 μM and at 30 μM. The values used in this plot are the mean of 3 independent trials. The red data points represent values determined at 3 μM and the blue data points represent values determined at 30 μM.

The IC_{50} of DFPP was determined using two concentrations of phosphonopyruvate, 3 μM (K_m) and 30 μM (10×K_m). The data from each of these concentrations was plotted (Figure 3.6) and the IC_{50} was determined for each condition based on the plot. Using the IC_{50} values, the average K_i was calculated to be 0.60 ± 0.05 μM. Clearly, this value indicates that DFPP is a potent inhibitor of PhPyrDC. However, questions still remain as to (i) whether the inhibitor is reversible and (ii) what is its mechanism of action.

3.2.4 Dilution studies of PhPyrDC with DFPP

In order to address the question of the reversibility of the enzyme, a dilution study was performed. The enzyme-inhibitor mix was incubated at 30 °C and, at appropriate time intervals, a sample of the mix was diluted 1:100 into an assay mixture containing 30 μM phosphonopyruvate. Due to the efficiency of the DFPP as an inhibitor of PhPyrDC, the reaction during these dilution studies occurred too quickly to monitor a time course for inactivation. In fact, after incubation with the inhibitor, no activity was seen at any of the time points tested. This clearly indicates that the inhibitor does bind in an irreversible manner.

3.2.5 Reaction of DFPP with PhPyrDC monitored by circular dichroism (CD) spectroscopy and UV-Vis spectroscopy

To begin to understand the mechanism of action of DFPP, PhPyrDC was titrated with varying concentrations of DFPP. This resulted in the appearance of a yellow color. Consequently, this titration was monitored via UV-Vis spectroscopy and it was found that the color change was reflected by an increase in absorbance at 440 nm (Figure 3.7). The intensity increased linearly until a plateau was reached (Figure 3.7 inset). The plateau occurred at approximately the same concentration as that of the enzyme, indicating that the enzyme was saturated with inhibitor.

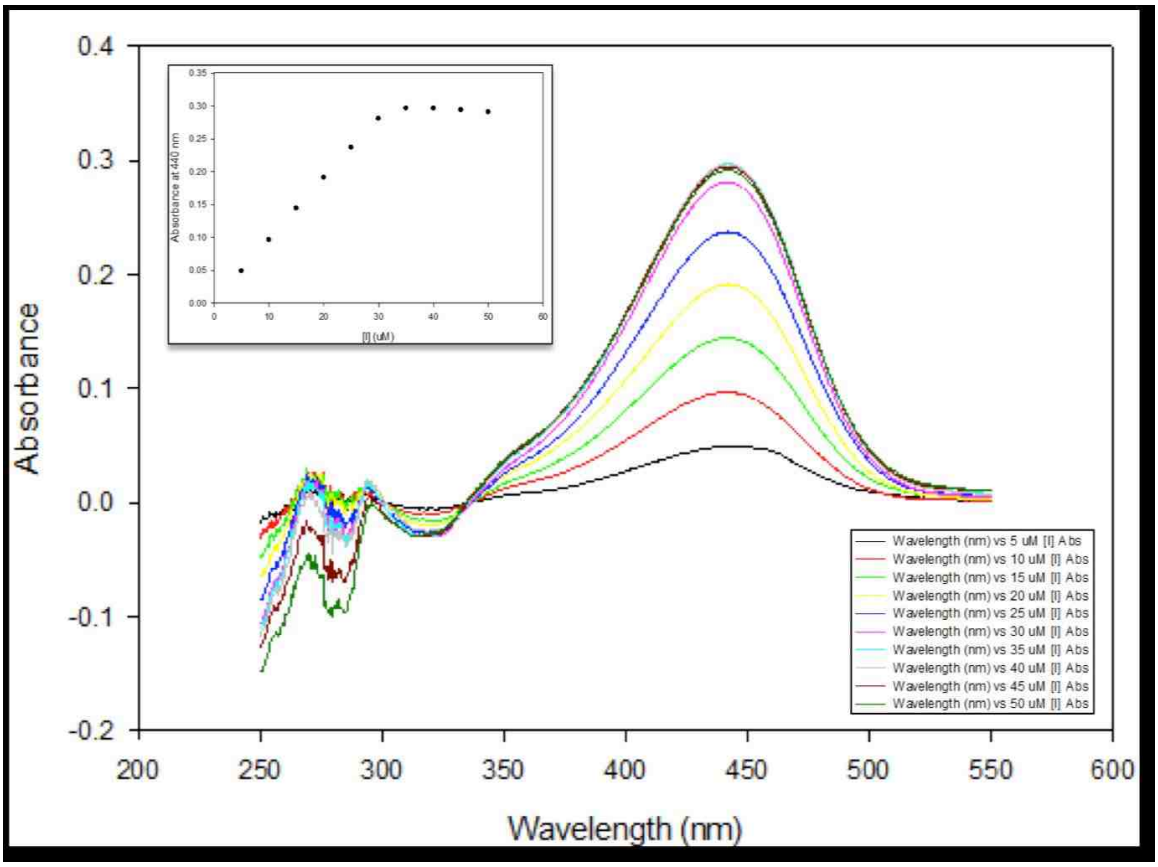

Figure 3.7. The titration of PhPyrDC (~50 μM) with DFPP as followed by UV-Vis spectroscopy. The inset is the plot of A_{440} vs. [I].

It has been shown that, on ThDP-dependent enzymes, ThDP and any of its derivatives are chiral (32,33). This means the DFPP titration could also be monitored by CD spectroscopy. The titration was therefore repeated, under conditions identical to those used in the UV experiments. As can be seen in Figure 3.8, there is an increase in ellipticity at 440 nm that also correlates with the observed color change and UV-Vis data. Further, there is again evidence of saturation, consistent with the UV data. Finally there is an increase in ellipticity at ~310 nm. This is consistent with a change from the AP form of ThDP to the IP

form seen with ThDP-derived intermediates (32,33). These data too suggest that the enzyme concentration used in the experiment is ~32 μM.

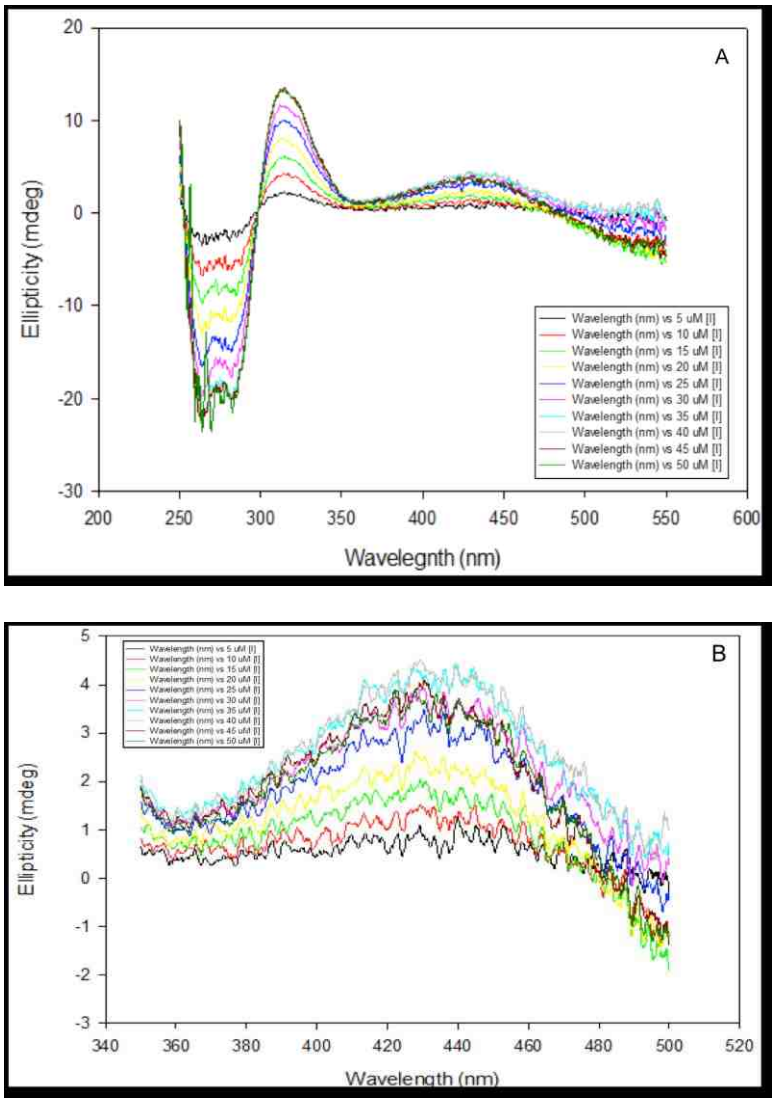

Figure 3.8. The titration of PhPyrDC (~100 uM) with DFPP as followed by CD spectroscopy. Figure A shows the monitoring from 550-250 nm, while Figure B highlights the changes occurring at 440 nm.

CD spectroscopy has been performed on other ThDP-dependent enzymes in the presence of a variety of inhibitors (34). Based on those earlier spectra, the

increase in ellipticity observed at 440 nm is likely to involve charge transfer transitions, but the basis for those transitions is yet to be determined (34).

3.2.6 Fluoride ion analysis of the reaction of DFPP with PhPyrDC

The proposed mechanism of the inhibitor (Figure 3.4) requires the release of a fluoride ion. To test this hypothesis, a fluoride ion-selective probe was used to analyze the products of the enzyme-inhibitor titration. Four different concentrations of inhibitor were tested, as well as two controls, as described in section 2.6.6. For this analysis, each reaction mixture contained 100 μM PhPyrDC. Based on the inhibitor concentrations used and the proposed mechanism of the inhibitor, we would expect to see a linear relationship between the amount of inhibitor and the amount of fluoride ion released.

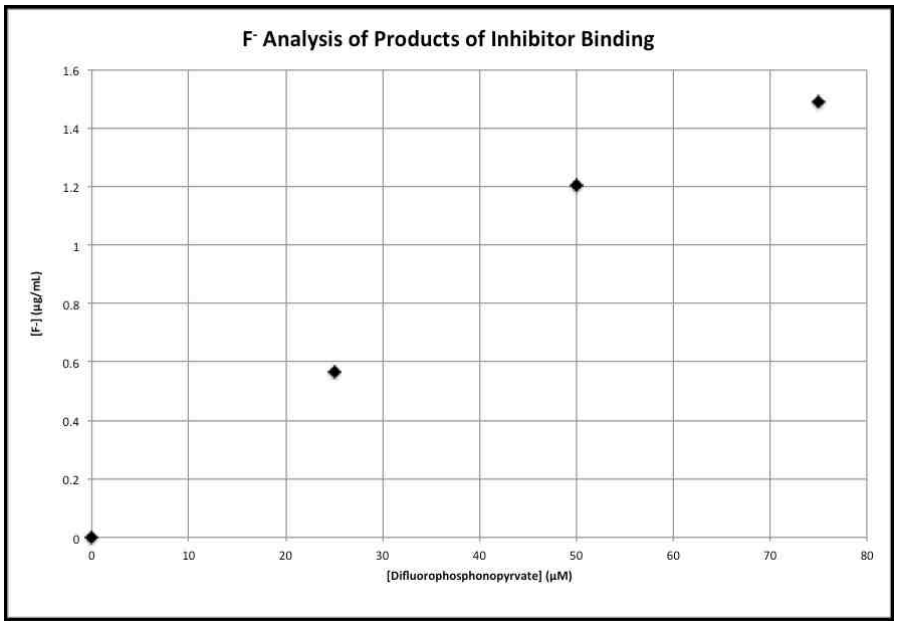

Figure 3.9. Fluoride ion analysis of the titration of PhPyrDC (100 μM) with DFPP.

Looking at Figure 3.9, we see a linear relationship between 0-50 mM DFPP, followed by a tapering off at a DFPP concentration between 50 μM and 75 μM. More likely this is an indication that the enzyme is saturated at around 65 μM rather than 100 μM. In Figure 3.9, we see start to see a tapering off around 65 μM and in the inset of Figure 3.7 the plateau begins at around 32 μM. Given that the enzyme concentration was doubled for the fluoride ion analysis, these UV and fluoride ion data are in good agreement. It certainly is possible that our calculated enzyme concentration is lower than the actual value, especially given that enzyme concentrations are calculated using the Bradford method.

The other possibility of error in this series of experiments is that the concentration of DFPP is not perfectly calculated. Since DFPP was synthesized and titrated to the sodium salt, calculating the exact molar mass difficult. The concentration of inhibitor may well be higher than expected.

Overall, based on the proposed mechanism of action of DFPP, we would expect to see a 1:1 ratio of fluoride ion to DFPP. Taking the concentration errors into account, this analysis confirms that a fluoride ion is being released and the data is a good indication that this reaction is indeed happening in a 1:1 ratio.

3.2.7 Crystallization studies of PhPyrDC

Our laboratory has previously shown the ThDP-dependent enzymes will often crystalize when the cofactor has been irreversibly inactivated (35). Here, we have demonstrated that DFPP does, in fact, irreversibly bind ThDP. However, initial crystallization trials with and without this inhibitor were not successful.

3.3 Conclusions and Future Direction

We have shown that difluorophosphonopyruvate is a mechanism-based inhibitor of PhPyrDC. By our studies to prove this, we have also demonstrated that the mechanism of the inhibition strongly correlates with the catalytic mechanism of PhPyrDC that has been proposed based on comparisons with other ThDP-dependent enzymes. In the future, stopped-flow experiments as well as 19F NMR will be used to investigate the intermediates of the PhPyrDC reaction. In addition, more inhibitors will be synthesized for further crystallization trials of PhPyrDC.

CHAPTER 4. INVESTIGATION AND CHARACTERIZATION OF THE MINOR ISOZYMES OF PYRUVATE DECARBOXYLASE ISOLATED FROM BAKER'S YEAST

4.1 Background

Glycolysis is the metabolic pathway utilized by most organisms to produce energy from the catabolism of glucose to form pyruvate. Once pyruvate has been formed, depending on the organism, there are three possible pathways for the breakdown of pyruvate: 1) aerobic oxidation where pyruvate is converted to acetyl-CoA, which then goes on to enter the citric acid cycle, 2) anaerobic homolactic fermentation to form lactate, or 3) anaerobic alcoholic fermentation, where ethanol and carbon dioxide are the final products (Figure 4.1).

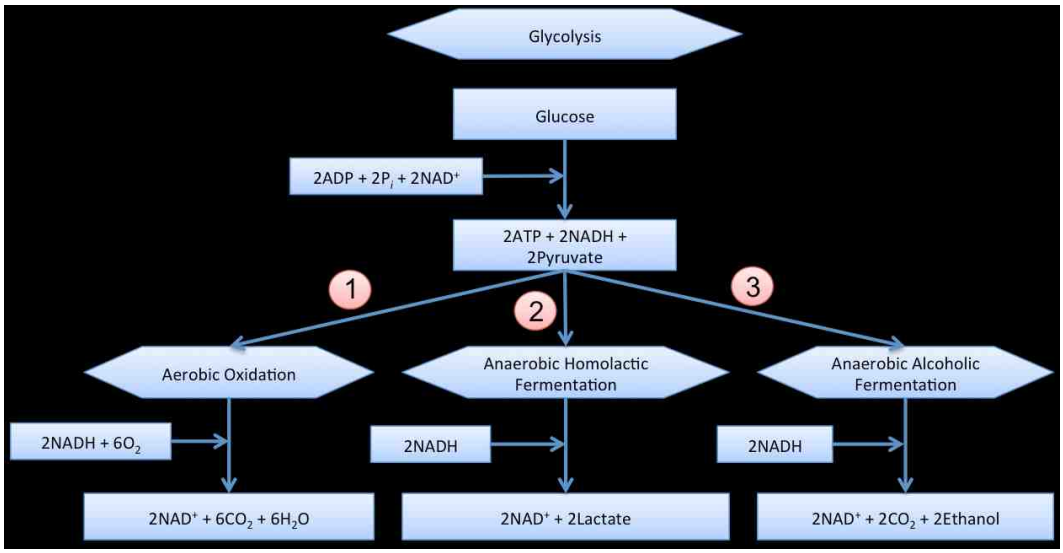

Figure 4.1. Glycolysis flowchart showing the three possible paths for the breakdown of pyruvate including the necessary reactants and the final products.

In *Saccharomyces cerevisiae*, also known as Baker's yeast, alcoholic fermentation is the principal pathway for the breakdown of pyruvate (36). The first step in alcoholic fermentation is the nonoxidative decarboxylation of pyruvate to acetaldehyde and carbon dioxide catalyzed by pyruvate decarboxylase (EC 4.1.1.1). This is then followed by the breakdown of the resulting aldehyde by alcohol dehydrogenase (ADH) to form ethanol (Figure 4.2).

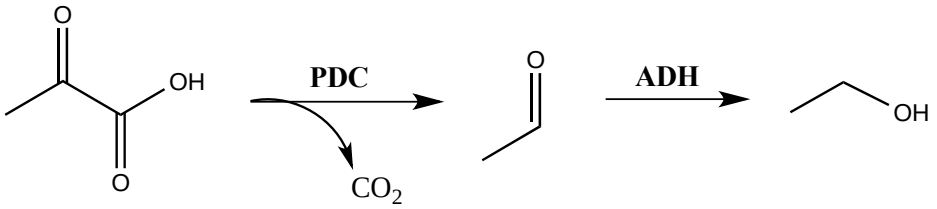

Figure 4.2. Reaction scheme of the conversion of pyruvate to ethanol utilizing pyruvate decarboxylase (PDC) and alcohol dehydrogenase (ADH).

In *S. cerevisiae*, pyruvate decarboxylase isozyme 1 (PDC1) is the primary isozyme that is expressed but, *S. cerevisiae* also possesses homologs for three other potential pyruvate decarboxylase isozymes, PDC5, PDC6, and THI3 (37). These will be the subject of the chapter.

4.1.1 PDC5

The first potential isozyme, PDC5, shares 88% sequence identity with PDC1. It was discovered by happenstance when researchers deleted or replaced *pdc1* and the resulting mutants were still able to ferment glucose, contrary to what the researchers predicted would happen (37,38). Seeboth *et al.* (38) determined that, in wild-type cells, *pdc5* is transcribed in a small amount or not

at all. However, when *pdc1* is knocked out, *pdc5* is transcribed at the same level as *pdc1* in wild-type cells (38). It is unlikely that *pdc1* would be completely knocked out under normal circumstances so researchers investigated the factors affecting the expression and regulation of *pdc5* (39).

They confirmed that the expression of both PDC1 and PDC5 is controlled by an autoregulation mechanism (38,39). Muller *et al.* (40) hypothesized that the deletion of *pdc1* could lead to higher ThDP levels in the cell, and that those higher levels could contribute to the stimulation of *pdc5*. Surprisingly, they found that ThDP is involved in PDC5 expression but high levels of ThDP cause repression rather than stimulation. They also determined that ThDP levels only affect the expression of PDC5 and not that of PDC1 (40). Eberhardt *et al.* (41) have determined that the autoregulation is independent of both sugar catabolism and the catalytic activity of PDC1. This suggests that PDC1 possesses a property beyond catalysis that is necessary to mediate the autoregulation. Further investigation utilizing *pdc1* mutants, which encoded a PDC1 that was catalytically inactive, showed that expression of PDC5 was still repressed, suggesting that these mutants were still active as regulatory proteins (41). Konig *et al.* (42) have shown that in PDC1, ThDP is necessary for the stability of the protein, and the binding of ThDP causes conformational changes in the structure of the enzyme for oligomerization purposes. Taken this information into account, Muller *et al.* (40) suggested that under ThDP limited conditions, ThDP does not bind to PDC1 and the active conformation does not form. This inactivated conformation may be the signal for the expression of PDC5 (40).

To date, the expression of *pdc5,* and the factors that affect it, are the only aspects that have been studied for this particular isozyme. A rationale for the isozyme's presence has been presented but not confirmed and the function of the enzyme has yet to be determined. To expand our knowledge of this isozyme, our laboratory has purified, characterized, and obtained an X-ray crystal structure of PDC5.

4.1.2 PDC6

As discussed in the previous section, the deletion of *pdc1* causes an enhancement in the expression of PDC5. PDC6 is another apparent isozyme of PDC1. However, when *pdc1* and *pdc5* were deleted, no detectable activity was found for PDC6 and the deletion yeast strain was unable to ferment glucose (23,39). Generally, the expression of PDC6 is extremely low, but when strains are grown on a non-fermentable carbon source such as ethanol, expression increases (23,43). Hohmann *et al.* (23) discovered that the level of PDC6 mRNA in the wild-type strain is increased in comparison to the *pdc1* knockout when the cells are grown on ethanol. This finding suggested that the expression of PDC6 is somehow connected to, at least, the presence of *pdc1*. In a later study, Hohmann *et al.* (43) also determined that the codon usage for PDC6 is typical for a weakly expressed protein, and that the pattern of the codon usage suggested that PDC6 had evolved to be a weakly expressed version of PDC1. These early findings provided limited information about PDC6 and none alluded to any possible reasoning for its presence.

In more recent studies, it has been shown that physiological changes cause large and genome-wide modifications in the gene expression of bacteria, yeast, and other eukaryotes (44). Gasch *et al.* (45) have determined that yeast cells have specific genomic "programs" that are utilized for each unique stress that the organism may encounter and those "programs" involve the alteration of thousands of genes. They believe that isozymes are involved in this response, hypothesizing that isozymes may possess different properties that make them more beneficial in a given situation (45). Fauchon *et al.* (44) determined that this was indeed the case concerning PDC6.

In their study, Fauchon *et al.* (44) exposed yeast to cadmium and show that, in order for the yeast to survive, they must remove it. Therefore, they must synthesize glutathione, a sulfur-containing compound, which is used to sequester the cadmium. To achieve the amount of sulfur needed, the yeast modifies its genomic expression so that the only genes that are expressed are those that require a limited amount of sulfur (44). Fauchon *et al.* (44) have shown that in response to cadmium, there is increased expression of PDC6 mRNA. At the same time, the expression of PDC1 mRNA is decreased. In a later study, it was also shown that PDC6 is expressed when sulfur is limited in the growth media, confirming the role of the expression of PDC6 as a response to sulfur limitation (46).

These findings have shed light on the potential role of PDC6. However, information about this enzyme is still very limited. Our laboratory has taken on the study of the purified enzyme in an attempt to characterize it.

4.1.3 THI3

In *S. cerevisiae,* ThDP is synthesized *de novo,* or external thiamin is converted to ThDP (47,48). The enzymes involved in synthesis and the transcription factors involved in the regulation of these pathways have been termed the THI regulatory system (49). These THI genes are some of the most strongly expressed in yeast, and are controlled by the level of thiamin present with higher concentrations repressing expression (47). Nishimura *et al.* (50) identified *thi3* as a positive regulatory gene. Mutations to this gene caused a loss of the yeast's ability to transport thiamin, as well as preventing the expression of many of the other genes involved in the THI regulatory system (49,50).

The product of *thi3*, THI3, contains 609 amino acids and shares 52% sequence identity with PDC1. THI3 contains the amino acid sequence, $GDGX_{24\text{-}27}NN$, that has been identified as a ThDP-binding motif which is conserved throughout ThDP-dependent enzymes (11). Nosaka *et al.* (49) hypothesized that the presence of this motif would allow THI3 to sense the presence of ThDP in the cell and expression or repression would be adjusted accordingly. Using mutagenesis to disrupt the binding of ThDP, it was found that the THI3 mutants were unable to detect high levels of ThDP and downregulate the expression of the THI genes thus confirming the hypothesis (49).

Despite its apparent function, THI3 is not similar to any other known transcriptional activator (49). Taking into account its high sequence similarity with PDC1, our laboratory has started an investigation into whether THI3 has

the capability of performing as a pyruvate decarboxylase in addition to its potential role as a regulatory protein.

4.2 Results and Discussion

4.2.1 Analysis of PDC5

4.2.1.1 Sequence alignments and homology model comparison of PDC5 and PDC1

PDC5 and PDC1 share the greatest sequence identity of the three potential isozymes at 88%. PDC5 retains all of the active site residues that have been determined as necessary for functioning (Figure 4.3). There are several stretches where the amino acid differences occur more heavily, for example, residues 340-360. However, these regions are not near to any of the active site residues.

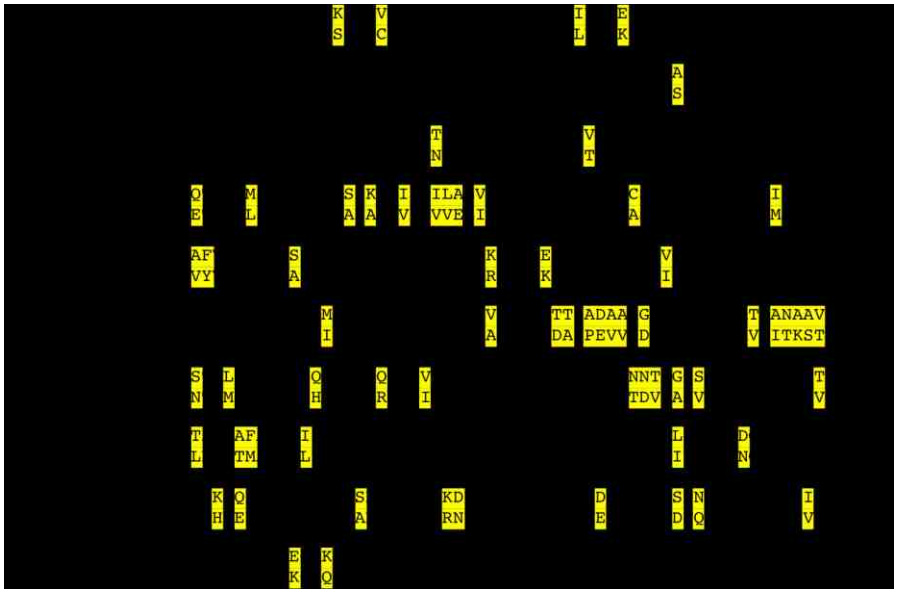

Figure 4.3. Clustal W (51) sequence alignment of PDC1 and PDC5. The yellow highlights indicate where the sequences differ.

Homology models, based on PDC1, were constructed using the Phyre2 Protein Fold Recognition Server. For the comparison of PDC1 and PDC5, PDB ID 2VK8 was used as the template. This comparison is shown in Figure 4.4. This figure clearly emphasizes the fact that the ThDP-binding site and the active (substrate-binding) sites of PDC1 and PDC5 are essentially identical.

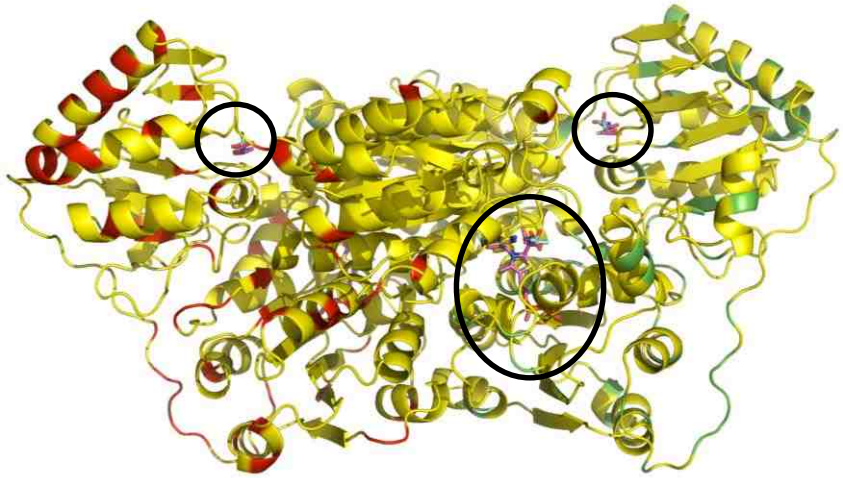

Figure 4.4. Homology model, constructed using Phyre2, highlighting the difference between PDC1 and PDC5. Red indicates differences on one monomer while green indicates areas of difference on the other monomer. The black circles highlight pyruvate in the regulatory site and ThDP.

Table 4.1 provides a closer look at the residue differences between PDC1 and PDC5 around the binding sites out to 20 Å. In fact, the first significant difference occurs around 8 Å from the ThDP binding site with the Asp472Asn change, which takes a negatively charged residue to a neutral residue. The next significant change is around 12 Å from the substrate binding site and 20Å from the ThDP site. The Glu551Lys mutation changes a negative charge to a positive charge. Also within that 12 Å area around the substrate binding site is the positive to neutral Lys554Gln alteration. These appear to be the only major

changes close to the active and ThDP binding sites. There are other changes further out as can be seen in Table 4.1 but, if there is any reduction in activity, residue changes closest to these sites should be investigated first. Right now, it is not clear what these changes mean but there is emerging evidence that changes in sites that are distant from the active site can have an affect on the overall activity of an enzyme (52).

Table 4.1. Differences surrounding binding sites between PDC1 and PDC5 (using PDC1 numbering). Determined by using homology models.

	ThDP	**Regulatory**	**Substrate**
6 Å	No Differences	Val 285 Ile	No Differences
8 Å	Asp 472 Asn	Cys 222 Ala	No Differences
12 Å	Ser 408 Val	Ser 251 Ala	Ser 408 Val
	Thr 419 Val	Met 314 Ile	Glu 551 Lys
	Ile 538 Val		Lys 554 Gln
20 Å	Lys 269 Arg	Val 158 Thr	Gly 406 Ala
	Asn 356 Thr	Met 187 Leu	Thr 419 Val
	Leu 365 Met	Ser 196 Ala	Leu 466 Ile
	Gln 373 His	Lys 198 Ala	Asp 472 Asn
	Val 383 Ile	Ile 201 Val	Lys 484 His
	Gly 406 Ala	Asp 202 Arg	Gln 486 Glu
	Thr 422 Leu	Ile 204 Val	Ile 538 Val
	Ala 426 Thr	Leu 205 Val	
	Leu 466 Ile	Asp 208 Ile	
	Lys 484 His	Ile 235 Met	
	Gln 486 Glu	Ala 242 Val	
	Ser 497 Ala	Phe 243 Tyr	
	Asp 519 Glu	Val 329 Ala	
	Glu 551 Lys	Gly 406 Ala	
	Lys 554 Gln	Ser 408 Val	
	Ile 37 Leu	Ala 106 Ser	
	Ala 106 Ser		
	Phe 427 Met		
	Lys 505 Arg		

4.2.1.2 Expression and kinetic characterization of PDC5

PDC5 was expressed and purified as described in section 2.4 resulting in 20 mg/L of protein. Once purified, the protein could be stored at -80 °C for several months without significant loss in activity. The purified enzyme was screened under steady-state conditions using the standard coupled assay system typically used for ThDP-dependent decarboxylases. This meant that in the initial assays for this enzyme, a phosphate buffer was utilized. It was immediately clear that PDC5 exhibited allosteric activation similar to PDC1 as evidenced by the sigmoidal curve in Figure 4.5.

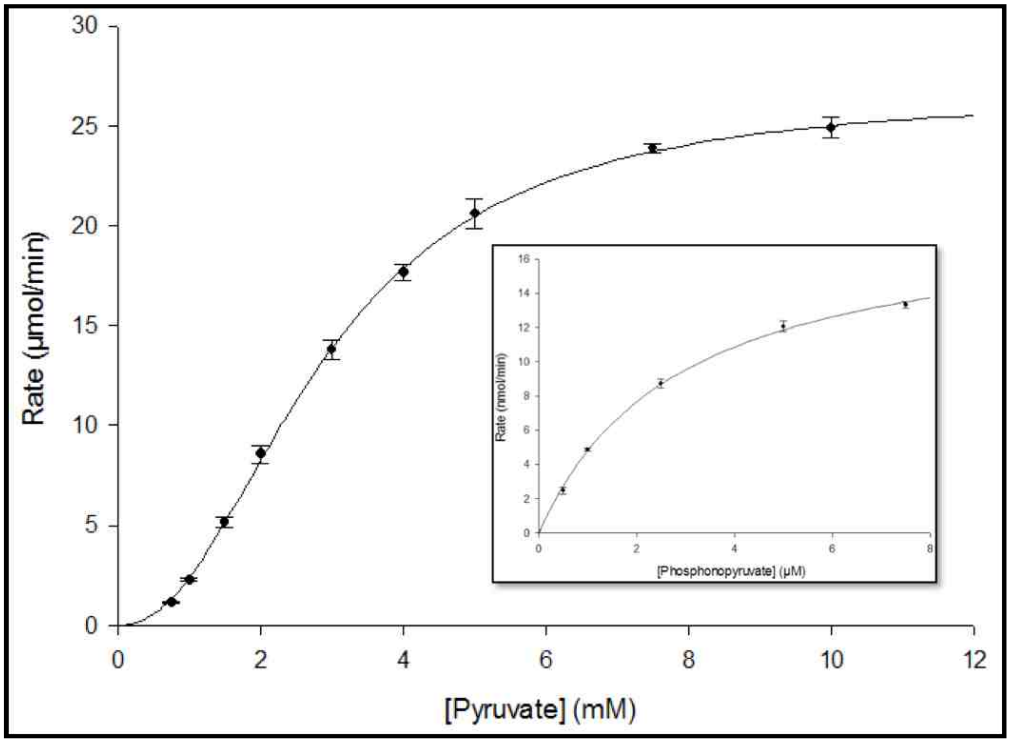

Figure 4.5. Michaelis-Menten plot for the reaction of PDC5 with pyruvate. Note the sigmoidal curvature not seen with enzymes displaying Michaelis-Menten kinetics. Inset is the plot of the PhPyrDC reaction with phosphonopyruvate for comparison. This indicates that PDC5 was an allosteric enzyme and, accordingly, kinetic data were fitted using the Hill equation.

As seen in Table 4.2, PDC5 has a minimum level of activity with all the substrates tested except benzoylformate. Surprisingly, by comparison to the literature values for PDC1 with pyruvate (S_{05} = 1.8 mM, k_{cat} = 38 s^{-1} (53)), the PDC5 S_{05} value is greatly increased while the k_{cat} is significantly reduced. Given that PDC1 and PDC5 share 88% sequence identity and identical active sites, these numbers are surprisingly different. Upon further review of the literature, it was discovered that, unlike most ThDP-dependent enzymes, PDC is inhibited by phosphate. In fact, phosphate reduces the enzyme's affinity for its substrate while also enhancing cooperativity (54). It is believed that this inhibition helps to regulate fermentation in addition to substrate activation (11).

Table 4.2. Kinetic Characterization of PDC5 carried out in a phosphate buffer[a]

Substrate	S_{05} (mM)	k_{cat} (s^{-1})	k_{cat}/S_{05} ($mM^{-1}s^{-1}$)	*n Value*
Pyruvate	35 ± 4	10 ± 1	0.30	1.7
2-Ketobutanoate	12± 1	5.3 ± 0.2	0.45	1.9
2-Ketopentanoate	19 ± 1	4.6 ± 0.1	0.24	1.9
2-Ketohexanoate	17 ± 1	2.0 ± 0.1	0.12	2.3
3-Methyl-2-ketobutanoate	21 ± 1	2.1 ± 0.1	0.099	1.9
4-Methyl-2-ketopentanoate	46 ± 5	1.7 ± 0.2	0.038	2.1
4-Methylthiol-2-ketobutanoate	13 ± 1	0.88 ± 0.07	0.071	2.0
Phenylpyruvate	6.2 ± 0.9	0.57 ± 0.09	0.092	1.4
Benzoylformate	NAD[b]	NAD[b]	NAD[b]	NAD[b]

[a] Values are the mean of 3 independent determinations ± standard error
[b] NAD: No activity detected

Based on the literature, the most popular buffers for the assay of PDC were a MES buffer and an imidazole buffer (14,24). In light of this information, these buffers were investigated to determine which yielded the greatest activity.

Table 4.3. Comparison of PDC1 and PDC5 in various buffers utilizing pyruvate as a substrate.

Enzyme - Buffer	S_{05} (mM)	k_{cat} (s^{-1})	k_{cat}/S_{05} ($mM^{-1}s^{-1}$)	*n Value*
PDC1 - MES	1.3 ± 0.1	60 ± 2	45.3	2.1
PDC1 - Imidazole	8.9 ± 0.7	74 ± 4	8.3	2.0
PDC1- Phosphate	19 ± 3	66 ± 9	3.4	2.0
PDC5 - MES	2.9 ± 0.1	64 ± 1	22.2	2.2
PDC5 - Imidazole	12 ± 1	75 ± 3	6.1	2.3
PDC5 - Phosphate	35 ± 4	10 ± 1	0.30	1.7

From Table 4.3, it can be seen that PDC1 exhibited the lowest S_{05} in the MES buffer at 1.3 mM. In addition, PDC5 also exhibited the lowest S_{05} in the MES buffer at 2.9 mM. Looking at the k_{cat} for PDC1 across the buffers tested, the value is slightly varied but largely unaffected. Interestingly, for PDC5, the phosphate buffer significantly decreases k_{cat}. The cooperativity of both enzymes remains unaffected.

In reviewing the literature, it was noted that a full characterization of PDC1 with an array of 2-keto acids had not been completed. For purposes of comparison against PDC5, this characterization was carried out using the optimum MES buffer system. The data is shown in Table 4.4.

Table 4.4. Kinetic Characterization of PDC1 utilizing a MES bufffer[a]

Substrate	S_{05} (mM)	k_{cat} (s^{-1})	k_{cat}/S_{05} ($mM^{-1}s^{-1}$)	*n Value*
Pyruvate	1.3 ± 0.1	60 ± 2	45	2.1
2-Ketobutanoate	0.49 ± 0.02	25 ± 1	49	2.7
2-Ketopentanoate	1.1 ± 0.1	27 ± 2	24	2.1
2-Ketohexanoate	1.5 ± 0.1	15 ± 1	9.8	1.9
3-Methyl-2-ketobutanoate	0.73 ± 0.03	6.3 ± 0.2	8.6	2.5
3-Methyl-2-ketopentanoate	8.7 ± 0.5	22 ± 1	2.5	2.2
4-Methyl-2-ketopentanoate	5.0 ± 0.2	10 ± 1	2.0	2.0
4-Methylthiol-2-ketobutanoate	0.80 ± 0.05	5.2 ± 0.3	6.6	2.4
Phenylpyruvate	0.23 ± 0.01	0.87 ± 0.02	3.7	2.5
Benzoylformate	NAD[b]	NAD[b]	NAD[b]	NAD[b]

[a] Values are the mean of 3 independent determinations ± standard error
[b]NAD: No activity detected

PDC1 had the best k_{cat}/S_{05} values for pyruvate and 2-ketobutanoate; these values are the same within error. Interestingly, PDC1 had a lower S_{05} value (2-fold) for 2-ketobutanoate than for pyruvate but had a much higher k_{cat} value for pyruvate. PDC1 had the lowest k_{cat}/S_{05} value for 4-methyl-2-ketopentanoate while no activity was exhibited with benzoylformate.

A similar kinetic characterization was carried out for PDC5, and the results shown in Table 4.5.

Table 4.5. Kinetic Characterization of PDC5 utilizing a MES buffer[a]

Substrate	S_{05} (mM)	k_{cat} (s^{-1})	k_{cat}/S_{05} ($mM^{-1}s^{-1}$)	*n Value*
Pyruvate	2.9 ± 0.1	64 ± 1	22	2.2
2-Ketobutanoate	1.0 ± 0.1	43 ± 1	44	2.5
2-Ketopentanoate	5.8 ± 0.3	24 ± 1	4.2	2.3
2-Ketohexanoate	3.4 ± 0.1	15 ± 1	4.4	2.5
3-Methyl-2-ketobutanoate	2.3 ± 0.1	13 ± 1	5.7	2.6
3-Methyl-2-ketopentanoate	16 ± 1	26 ± 2	1.6	2.7
4-Methyl-2-ketopentanoate	12 ± 1	12 ± 1	1.1	2.3
4-Methylthiol-2-ketobutanoate	1.8 ± 0.1	6.8 ± 0.2	3.8	3.2
Phenylpyruvate	0.5 ± 0.1	2.0 ± 0.1	4.2	2.4
Benzoylformate	NAD[b]	NAD[b]	NAD[b]	NAD[b]

[a] Values are the mean of 3 independent determinations ± standard error
[b]NAD: No activity detected

The data showed that 2-ketobutanoate had the best k_{cat}/S_{05} value at 44 $mM^{-1}s^{-1}$ with pyruvate following at 22 $mM^{-1}s^{-1}$. In almost all instances, within error, PDC1 had higher activity with the substrates tested. Intriguingly, with the exception of benzoylformate, both enzymes exhibited some level of activity with all substrates tested. Overall, the result of the comparison of these enzymes is not surprising. The high level of sequence identity and the suggested function of PDC5 as a "back-up" to PDC1 in certain circumstances would lead to the expectation that these enzymes would and should function in a similar manner for cell survival.

4.2.1.3 Crystallization, X-ray structure and structure analysis for PDC5

PDC5 was expressed and purified for crystallization as described in Chapter 2. The purified protein was concentrated to a final concentration of 15 mg/mL. Seeding was used to facilitate crystal growth utilizing the hanging drop diffusion method. Screening kits from Hampton Research and Rigaku were used to test various potential crystallization conditions. Once the crystals began to appear, those conditions were tested with the crystallization buffer alone and with varying concentrations of enzyme to optimize the conditions and to help determine whether the observed crystals were salt or protein.

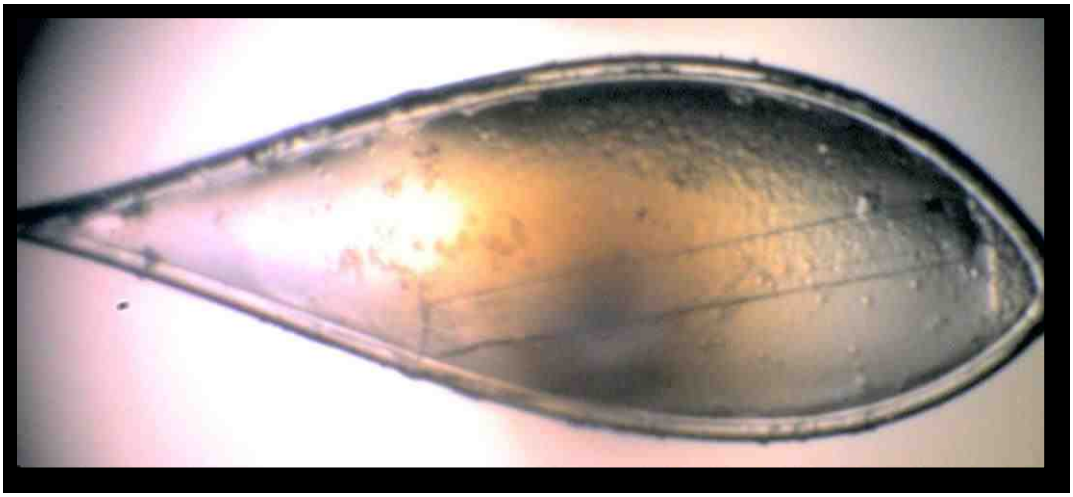

Figure 4.6. Crystal of PDC5 that resulted in an X-ray structure.

Protein crystals were obtained using a 2 M ammonium sulfate and 100 mM sodium citrate/citric acid, pH 5.5 solution and crystals formed in the absence of substrate or a substrate analog. These crystals were utilized for the x-ray diffraction (Figure 4.6) and found to diffract to 2.8 Å. The relatively poor diffraction is likely due the crystal having two space groups for one crystal, also known as being twinned. Another possible cause for the poor diffraction is that the cryogenic conditions were not optimized, due to the limited number of crystals available.

Dr. Forest H. Andrews completed the refined model. Refinement statistics can be seen in Table 4.6. The best model resulted in an R_{free} value around 40% and an R_{work} of 33%.

Table 4.6. X-ray diffraction data collection and refinement statistics

	PDC5
Data collection	APS GM/CA-23-ID-B
Space group	
Cell dimensions	P212121
a, *b*, *c* (Å)	90.0, 90.0, 90.0
α, β, γ (°)	114.0, 151.6, 194.8
Resolution (Å)	2.78-18.0 (2.78-2.83)
R_{sym} or R_{merge}	15.6 (68.0)
$I/\sigma I$	5.7 (2.5)
Completeness (%)	98.8 (99.5)
Redundancy	
Refinement	
Resolution (Å)	2.78
No. reflections	340138
R_{work}/ R_{free}	33.4/40.6
No. atoms	17490
Protein	17208
Water	282
B-factors	**37.23**
R.m.s deviations	
Bond lengths (Å)	0.010
Bond angles (º)	1.443

The active site of PDC5 was then compared to the active site of PDC1 in the open and closed conformations (PDB IDs 1PVD and 2W93, respectively). As can be seen in Figure 4.7, the overlay of the active sites of these two enzymes in the open conformation is almost identical. This would be expected since PDC5 was not crystalized with any substrate analogs. It is clear that Ile415, Glu51, His115, and Glu477 align almost perfectly, and that there is very little difference between the two structures. However, although Asp28 and His114 both occupy the same spatial area, these PDC5 residues appear to be different rotamers than those in PDC1. This is not uncommon and can often be seen between structures of the same enzyme.

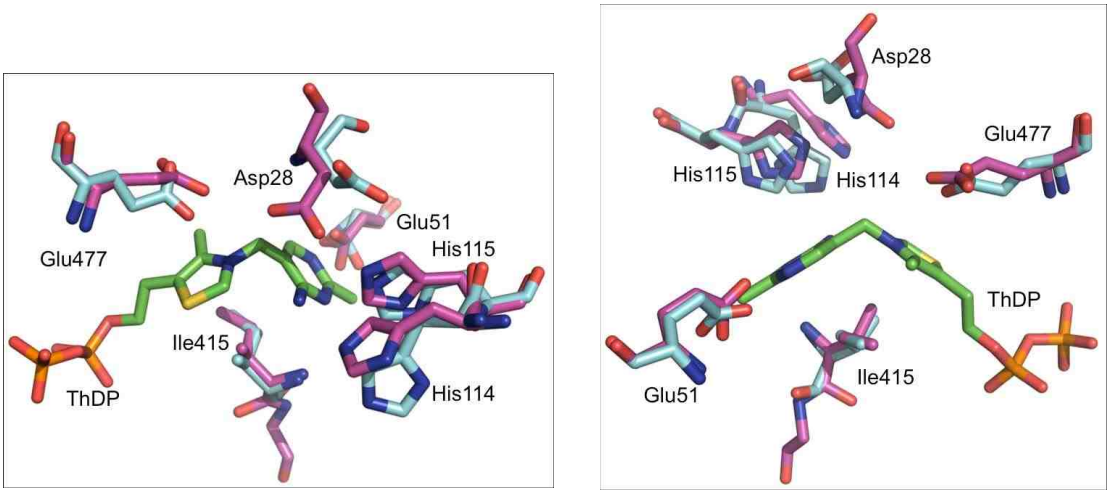

Figure 4.7. Comparison of the active sites of PDC1 (magenta) and PDC5 (cyan) in the open conformation from two views. Figures were generated using PyMOL with the PDC1 data from PDB ID 1PVD.

The PDC5 structure was also compared with that of PDC1 in the closed conformation, (Figure 4.8). Here, Ile415, and Glu51, and Glu477 (Gln477 in the PDC1 variant used) remain almost unchanged between the open (PDC5) and the closed (PDC1) conformations. However, looking at His114, His115, and Asp28,

there is a clear shift in the active site between PDC1 (closed conformation) and PDC5 (open conformation). The HH-motif appears to shift closer to the substrate while Asp28 moves over to accommodate the shifting histidine residues. This comparison is in good agreement with the restructuring that occurs when substrate binds to the regulatory site as described in the introduction.

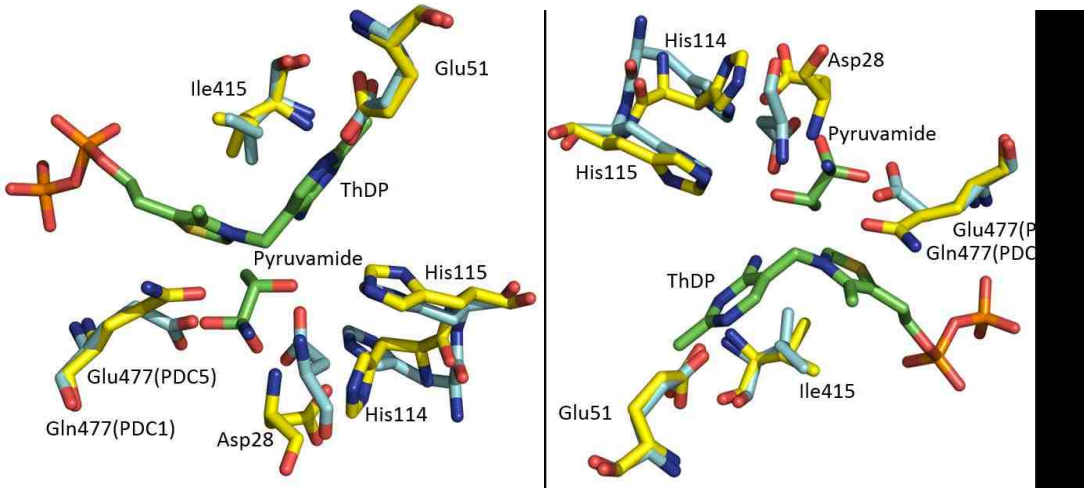

Figure 4.8. Comparison of the active sites of PDC1 (yellow) and PDC5 (cyan) using the closed conformation of PDC1 from two views. Figures were generated using PyMOL with data from PDB ID 2W93 used for PDC1. Note that, in order to generate the closed conformation, a PDC1 mutant, Glu477Gln in complex with pyruvamide was used.

4.2.2 Analysis of PDC6

4.2.2.1 Sequence alignment and homology model comparison of PDC6 and PDC1

PDC1 and PDC6 share 83% sequence identity and a purpose for this isozyme has been determined as a "sulfur-limited" response. In fact, PDC6 has substitutions for 11 of the 17 sulfur-containing residues within the sequence. The sequence alignment of these enzymes can be seen in Figure 4.9.

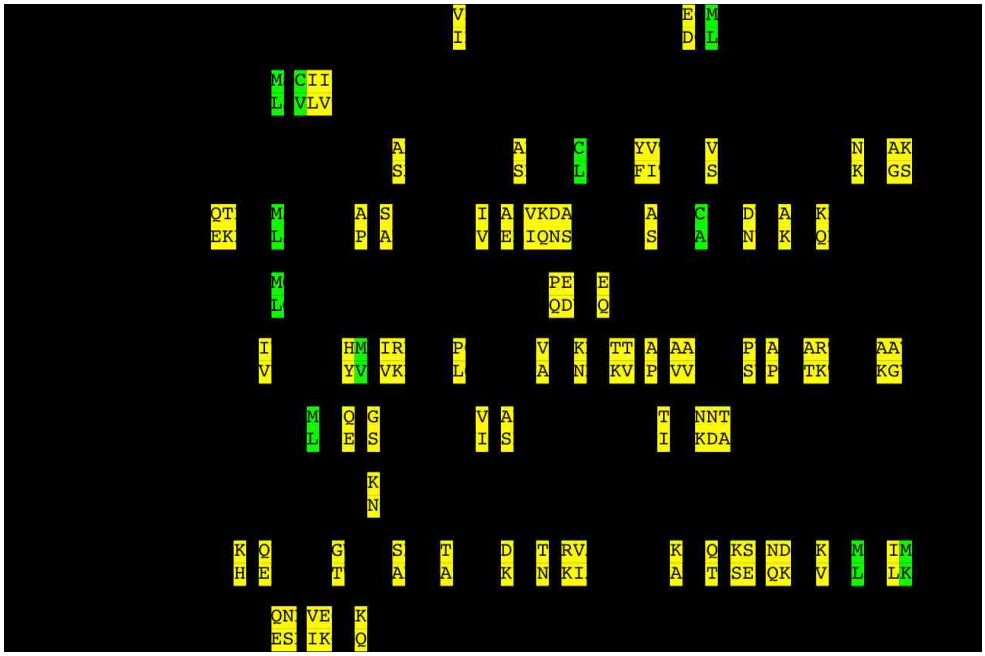

Figure 4.9. Clustal W sequence alignment of PDC1 and PDC6. The yellow highlights indicate amino acid changes from PDC1. In addition, those amino acids highlighted in green indicate where a sulfur-containing residue has been replaced.

When comparing these sequences, it should be noted that PDC6 retains all the conserved acid/base residues shown in Figure 1.8. To enable further characterization, PDC6 was expressed and purified as described in section 2.4 (see below for more detail). Unlike PDC5, the yield of enzyme was relatively poor and, more importantly, the purified enzyme was inactive. This leads to the question of what residue differences could lead to an inactive enzyme. Using a homology model, constructed using Phyre2 as described earlier, the structures of these enzymes were examined, paying particular attention to each of the binding sites: ThDP, regulatory, and substrate.

Table 4.7 lists the differences found between the two enzymes around each of these binding sites out to a distance of 20 Å. Since the activity of the

enzyme is in question, it would make sense that residue changes closest to the substrate binding site or ThDP binding site are going to be the most important to look at.

Table 4.7. Differences surrounding binding sites between PDC1 and PDC6 (PDC1 numbering). The differences in sulfur-containing residues are in green.

	ThDP	**Regulatory**	**Substrate**
6 Å	No Differences	His 313 Tyr	No Differences
8 Å	No Differences	Cys 222 Ala	No Differences
		Met 247 Leu	
12 Å	Met 370 Leu	Tyr 157 Phe	Ile 71 Val
	Ala 386 Ser	Val 158 Ile	Ala 386 Ser
	Gln 486 Glu	Ala 218 Ser	Asn 548 Ser
	Gly 492 Thr	Asp 226 Asn	Val 550 Ile
	Ile 70 Leu	Ile 306 Val	Glu 551 Lys
	Ile 71 Val	Met 314 Val	Lys 554 Gln
	Ile 538 Leu	Pro 322 Leu	
	Met 539 Lys	Val 329 Ala	
	Gln 547 Glu		
	Asn 548 Ser		
	Val 550 Ile		
20 Å	Met 247 Leu	Met 67 Leu	Val 22 Ile
	Gln 373 Glu	Cys 152 Leu	Cys 69 Val
	Gly 375 Ser	Val 163 Ser	Ile 70 Leu
	Val 384 Ile	Met 187 Leu	Met 247 Leu
	Thr 399 Ile	Ala 194 Pro	Met 370 Leu
	Thr 404 Ala	Ser 196 Ala	Val 384 Ile
	Lys 484 His	Ile 204 Val	Lys 484 His
	Ser 497 Ala	Asp 208 Val	Gln 486 Glu
	Thr 501 Ala	Ala 229 Lys	Gly 492 Thr
	Arg 511 Lys	Lys 232 Gln	Ile 538 Leu
	Val 512 Ile	Ile 316 Val	Met 539 Lys
	Val 22 Ile	Arg 317 Lys	Gln 547 Glu
	Met 43 Leu	Lys 332 Asn	
	Cys 69 Val	Val 384 Ile	
	Ala 137 Ser	Ala 386 Ser	
	Val 163 Ser	Thr 404 Ala	
	Glu 551 Lys		
	Lys 554 Gln		

Interestingly, no differences occur until at least 12 Å from any of those sites. Perhaps the most interesting change is Glu551Lys, where a negative charge is substituted by a positive charge. This occurs around 12 Å from the substrate binding site and 20 Å from the ThDP binding site. Further out from these differences are changes where a neutral amino acid is replaced by a charged amino acid or vice versa, as is the case in Lys554Gln, Gln486Glu, and Gln547Glu. Not taking into consideration these changes, all other changes seem minor. However, it is possible that the sum of these changes is enough to result in the inactive enzyme that was produced in this study (see below).

4.2.2.2 Expression conditions for PDC6

PDC6 was expressed using LB media and IPTG induction, however, the amount of protein obtained was <5 mg/L and the enzyme was virtually inactive. In an attempt to increase the yield of protein and possibly the activity of the enzyme, auto-induction media and terrific broth (TB) with IPTG induction were tried. The auto-induction media yielded similar amounts of inactive protein. Interestingly, when using the terrific broth, and growing the cultures from start to finish at room temperature, with the exception of the 2- 3 hours at 37 °C to reach an OD of 0.6, the amount of purified protein more than doubled reaching >10 mg/L.

In this case, the enzyme remained inactive. However, it should be noted that the protein that is expressed is soluble and does not form inclusion bodies. This would indicate that, at the very least, the enzyme is folding in a somewhat

correct manner and is not likely the direct cause of inactivity. More work is required to identify the problem so that characterization can be carried out.

4.2.3 Analysis of THI3

4.2.3.1 Sequence alignment and homology model comparison of THI3 and PDC1

THI3 and PDC1 share 52% sequence identity (Figure 4.10), the lowest of the three potential isozymes. When comparing these sequences, THI3 maintains the conserved ThDP binding site, $GDGX_{24\text{-}27}NN$, mentioned in the previous section. THI3 also possesses the essential glutamic acid residue (Glu51), the HH-binding motif, and Cys221, the residue that is essential for regulation. However, THI3 differs from PDC1 in three important active site residues, i.e., 1) Asp28 has been replaced by Glu, 2) Ile415 has been replaced by Val, and 3) Glu477 has been replaced by Asp.

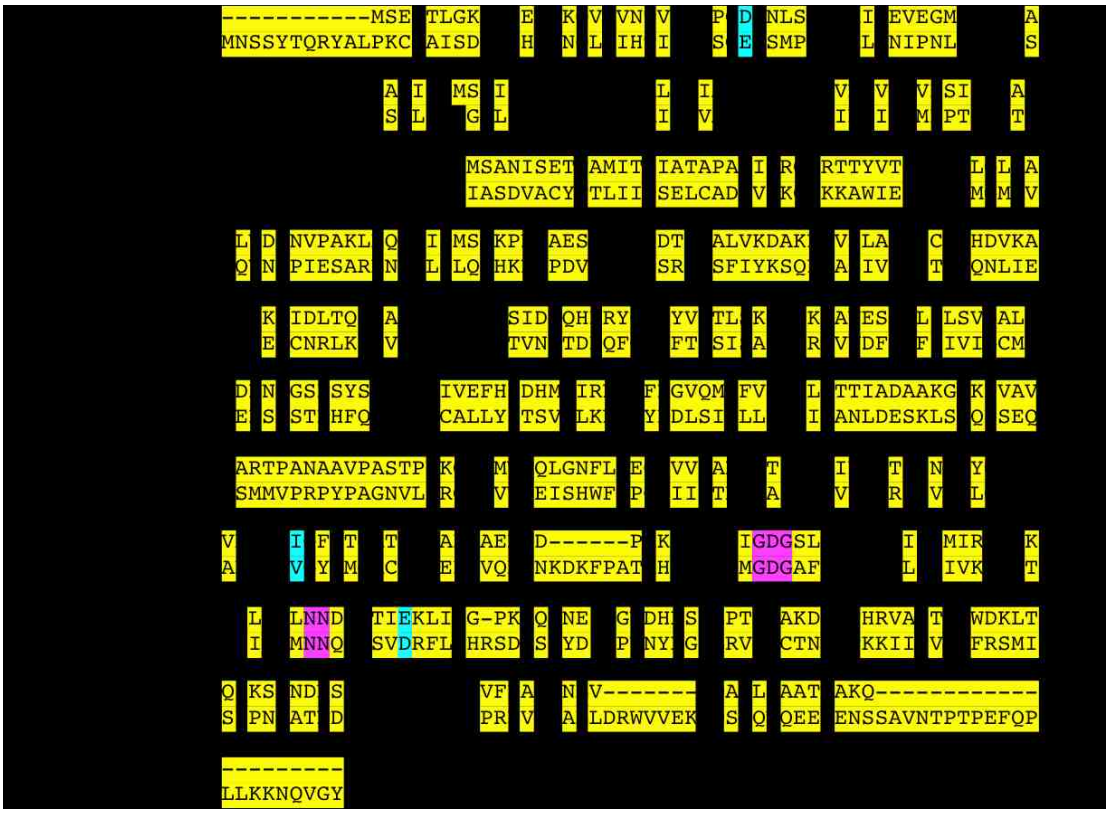

Figure 4.10. Clustal W sequence alignment of PDC1 and THI3. The yellow highlighting indicates the differences between the sequences. The purple highlights indicate the beginning and ending of the conserved ThDP binding sequence, $GDGX_{24\text{-}27}NN$. The blue highlighting indicates changes in residues known to be critical for the functioning of PDC1.

Homology models based on PDC1 were constructed using Phyre2 Protein Fold Recognition Server. A comparison of the active site differences between PDC1 and THI3 (Figure 4.11) was prepared using PDB ID 2VK8 (PDC1) as the template. These models help to provide insight into how these changes may affect the positioning of the substrate and cofactor and also how catalytic ability may be affected. Figure 4.11A illustrates the change at position 415 (PDC1 numbering) from an isoleucine in PDC1 to a valine in THI3.

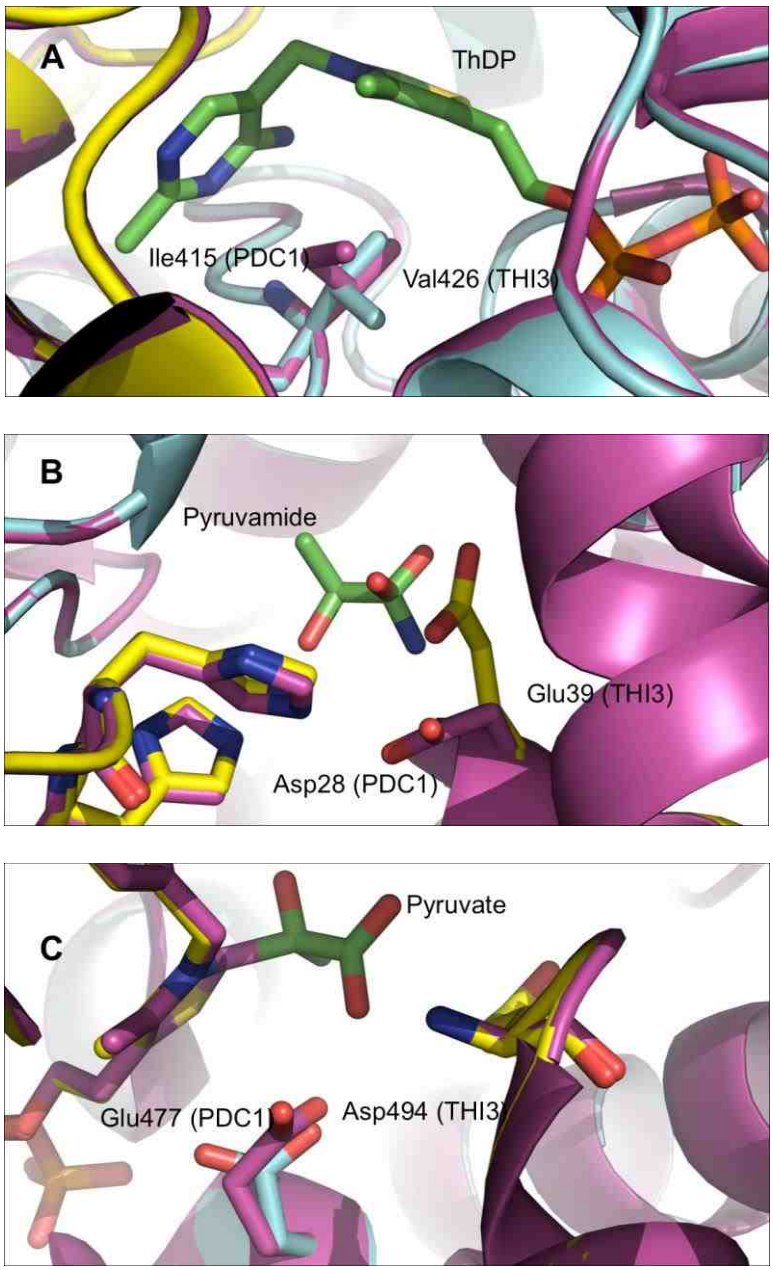

Figure 4.11. Overlay of THI3 (cyan and yellow for each monomer) and PDC1 (magenta). A) Ile415 replacement by Val, B) Asp28 replacement by Glu, C) Glu477 replacement by Asp. Figures constructed using PyMOL with data from PDB ID 2VK1 and 2W39.

In the introduction, it was noted that the V-conformation of ThDP is necessary for catalysis, and it was believed that this conformation was maintained by the presence of a bulky, hydrophobic residue (Ile415 in PDC1). Andrews *et al.* (13) recently showed that this is not the case in BFDC, and that the V-conformation seemed to be attained essentially independently. However, the identity of the residue at that position did play a large role in enzyme activity. Guo *et al*. (12) have studied this position extensively in PDC and found that mutations to this position have a large effect on catalytic activity. In fact, when the Ile415Val mutation was tested, it was found that k_{cat} was significantly diminished with the mutant only retaining 6% of wild-type activity. Conversely, $S_{0.5}$ was only slightly affected (12). This significant reduction in activity suggests that this I415V mutation may be one of the contributing factors to why THI3 appears to be catalytically inactive.

Figures 4.11B and 4.11C illustrate the changes at position 28 from an aspartate in PDC1 to a glutamate in THI3, and at position 477 from a glutamate in PDC1 to an aspartate in THI3. In PDC1, Asp28 and Glu477 play pivotal roles in catalysis due to their ability to form hydrogen bonds with the carboxylate moiety of pyruvate and position the substrate. When these sites have undergone mutation, the resulting enzymes have been almost catalytically inactive (for D28A k_{cat} = 0.07 s^{-1}, for E377Q k_{cat} = 0.09 s^{-1}) (14,55,56). Presumably these changes would also render THI3 inactive.

4.2.3.2 Attempted expression of THI3

The expression of THI3 utilizing the pET17THI3-His plasmid was carried out as described in Chapter 2. In addition to using LB media with IPTG induction, expression was also attempted in an auto-induction medium, as well as terrific broth (TB) with IPTG induction. Expression cultures were grown at both 37 °C and at room temperature. However, no expression of the protein was seen under any of the conditions tested.

Under my supervision, a pETSUMO-THI3 expression vector was constructed by Ahmed Malik to determine whether use of a fusion protein (57) could help to facilitate the expression of THI3. However, when expression was reattempted, unfortunately, there was no evidence of protein expression.

Alternatively, the pKK223 is under a *tac* promoter system while the pET system uses a T7 promoter system. This difference in promoter systems was utilized to again attempt expression. Again, no expression was seen with the conditions tested.

4.3 Conclusions and Future Directions

We have fully characterized PDC5 and have confirmed its similarity to PDC1. However, the role of PDC5 as a "back-up" to PDC1 still remains to be proven. We were unable to characterize PDC6 due to the inability to obtain active protein. It is possible, given the successful expression of active PDC6 in yeast (24), that *E.coli* is not a suitable model organism for the expression of active PDC6. In contrast, THI3 has not been expressed in *E. coli.* Nosaka *et al.*

(49) have reported some expression in yeast but these levels are negligible (~100 μg/L). Taking into account the three major residue differences between THI3 and PDC1 and our unsuccessful expression, our laboratory is currently making the necessary variants to determine whether expression and activity can be achieved.

CHAPTER 5. INCREASING THE EFFICIENCY OF THE SCREENING AND SELECTION REQUIRED FOR SITE-SATURATION MUTAGENESIS OF PDC VARIANTS

5.1 Background

5.1.1 Site-saturation mutagenesis

Directed evolution, a term used to describe a range of molecular biology techniques that allow natural evolutionary processes to be mimicked in the lab, has emerged as a useful way to engineer the catalytic properties of enzymes (58,59). These techniques include error-prone PCR (epPCR), DNA shuffling, and site-saturation mutagenesis. Error-prone PCR, the most frequently used technique, introduces random point mutations and is thought to address the gene as a whole (59,60). DNA shuffling allows random recombination between two or more genes and is often used in conjunction with epPCR (58,61). Site-saturation mutagenesis (SSM) allows for an individual site to be targeted for mutagenesis wherein all 20 natural amino acids can be substituted into that site (62). This last technique gives the ability to generate protein variants that are novel, revealing functional capabilities of an enzyme that would not have been discovered based on using rational design alone. Based on this, SSM has been used to engineer changes in substrate specificity, enantiospecificity, to improve thermostability, and to introduce new enzyme functions (3).

To begin using SSM, degenerate mutagenic primers are designed. This means that the primers synthesized are a mixture of oligonucleotides that are identical except in the position that has been targeted for mutagenesis. The identical parts of the primers allow annealing to occur to the parental DNA while the degenerate position creates diversity in the amplified product (63). This amplified product is then transformed into *E. coli* or another model organism and a library is created and screened for activity. Using this technique, two problems can arise. First, SSM can be experimentally time consuming because multiple rounds of mutagenesis are required to create variants that have been successfully evolved to perform the desired function. Second, the number of colonies that have to be screened to ensure that every amino acid has been covered at a site increases exponentially as each additional site is targeted. To address the first issue, the iterative saturation mutagenesis (ISM) and combinational active site saturation mutagenesis (CAST) techniques have been developed (64,65). In ISM, sites are chosen, generally based on structural data or homology models, and SSM is carried out on one site. Single SSM is carried out for each of the chosen spots. The multiple libraries generated are then screened and the mutants that are classified as "hits" are then put through another round of SSM at one of the next spots chosen. This cycle continues until the desired results are achieved (64). The CAST technique, on the other hand, targets multiple sites at one time (65). By combining one or both of these techniques with rational design, the experimental time can be reduced but these techniques

alone fail to address the fact that the genetic code is redundant and results in large numbers of colonies needing to be screened.

The most important aspect of a SSM experiment has to deal with the codon randomness. A completely randomized codon results in a library size of 64 different sequences that encode all 20 amino acids and 3 stop codons (63). This is also known as the NNN library where N can be adenine (A), cytosine (C), thymine (T), or guanine (G). Using the NNN library will result in 64 different combinations of nucleotides at a single selected position and, to ensure that full coverage is achieved, around 200 colonies will need to be screened (59). If the number of sites being targeted is increased to three, there are 262,144 different combinations possible, thus the requirement for colony screening increases exponentially (63). In order to reduce these numbers but still maintain full coverage of all the amino acids, the NNK library was developed where K is either a T or G. By eliminating some of the redundancy in the genetic code, the library size is reduced (59,63). For instance, in the NNN library, arginine has six codons that code for it but, in the NNK library, there are only two codons for arginine. Another library, the NDT library (where D can be A, G, or T), reduces the possible number of amino acids to 12 from the native 20. However, this library does include a balanced mix of polar, nonpolar, aliphatic, aromatic, negatively charged, and positively charged amino acids. Structurally similar amino acids have been excluded and there are no stop codons, thereby significantly reducing the number of colonies needing to be screened (59). Reetz *et al.* have compared the NNK and NDT libraries to demonstrate the differences in the number of

colonies that need to be screened (Figure 5.1). Reetz *et al.* (59) experimentally tested the difference between the NNK and NDT libraries and found that the NDT library produced higher quality results. Overall, the NDT library had more positive and active hits than the NNK library, and the redundancy of the positive hits was significantly lower than expected (59).

No.[a]	NNK		NDT	
	Codons	Transformants needed	Codons	Transformants needed
1	32	94	12	34
2	1028	3066	144	430
3	32768	98163	1728	5175
4	1048576	3141251	20736	62118
5	33554432	100520093	248832	745433
6	$>1.0\times10^{9}$	$>3.2\times10^{9}$	$>2.9\times10^{8}$	$>8.9\times10^{8}$
7	$>3.4\times10^{10}$	$>1.0\times10^{11}$	$>3.5\times10^{7}$	$>1.1\times10^{8}$
8	$>1.0\times10^{12}$	$>3.3\times10^{12}$	$>4.2\times10^{8}$	$>1.3\times10^{9}$
9	$>3.5\times10^{13}$	$>1.0\times10^{14}$	$>5.1\times10^{9}$	$>1.5\times10^{10}$
10	$>1.1\times10^{15}$	$>3.4\times10^{15}$	$>6.1\times10^{10}$	$>1.9\times10^{11}$

[a] Number of aa positions at one site.

Figure 5.1. The number of colonies needed in an NNK versus and NDT site-saturation mutagenesis experiment. Data from Reetz *et al.* (59).

Unfortunately, reducing the redundancy in the genetic coding used for SSM is the only way to decrease the number of colonies needing to be screened for the desired change. However, the addition of a selection step can also be used as a way to reduce the number of colonies to be investigated.

5.1.2 Selection vs. screening

In selection, it is assumed that the host organism experiences a growth or survival type advantage if it contains a gene that has been changed in some way to provide that advantage (62). The premise behind selection is to automatically eliminate the variants that are nonfunctional. Often times selection is focused on genes that provide some sort of nutrient that is absolutely required by the cell (66). However, this does not always have to be the case and this method may not work for every scenario. For example, while Reetz *et al.* (67) have used directed evolution techniques to change the enantioselectivity of lipases, they have also found that traditional selection techniques can be difficult to use to address enantioselectivity. To overcome this, Reetz *et al.* (66,67) developed a selection method wherein a mix of enantiomers were provided to the cells and, if the lipase variant catalyzed the hydrolysis of the "wrong" enantiomer, a poison resulted and the cells died. If the "correct" enantiomer was hydrolyzed, it acted as a carbon source (66,67). This method is doubly effective because lipases with the wrong enantioselectivity are eliminated by cell death and those with the correct enantioselectivity provide the carbon source that is necessary for growth. Clearly, using either one of these independently would be an effective way to select for the desired mutation.

Despite reducing the number of colonies, selection still presents a problem. Once an organism is able to maintain growth using this method, there is no pressure for that organism to evolve further to improve its activity or sustainability. The McLeish laboratory has carried out some evolutionary studies

on ThDP-dependent enzymes (see below). Further progress now requires either extensive screening, or development of an efficient selection method. This chapter will detail some progress towards the latter goal.

5.1.3 Evolution of benzoylformate decarboxylase to a pyruvate decarboxylase

Benzoylformate decarboxylase (BFDC) catalyzes the conversion of benzoylformate to benzaldehyde and carbon dioxide. BFDC from *Pseudomonas putida* shares a similar structure and mechanism to all ThDP-dependent decarboxylases including that of *Sc*PDC and the PDC from *Zymomonas mobilis* (*Zm*PDC). Mechanistically, BFDC differs from PDC only in its preference for the phenyl moiety of benzoylformate rather than the methyl group of pyruvate.

Structurally, a comparison of BFDC, *Zm*PDC, and *Sc*PDC reveals the enzymes are similar in overall protein fold, and the residues that are involved in the binding of ThDP. However, when looking at the residues that make up the active site, there is little conservation between BFDC and the two PDCs (68). Given the similarity in catalytic mechanism, where the only difference for any ThDP-dependent decarboxylase is the bulk of the preferred substrate, it would appear that evolution of substrate specificity is related the evolution of the active site architecture (16,68). Previously, Siegert *et al.* (69) had begun to explore this issue and attempted to convert a BFDC into a PDC (and vice versa).

In that study, residues were selected for mutagenesis based on sequence comparisons and an overlay of the structures of the active sites. In *Zm*PDC, there are two conserved residues, Ile472 and Ile476 (Ile476 and Ile480 in *Sc*PDC),

which are thought to play a role in substrate specificity (4). BFDC possesses two residues, Ala460 and Phe464, which could be considered direct counterparts to those in the two PDCs and these residues were the targeted sites for site-directed mutagenesis. The single and double mutants were prepared but overall, the experiment was largely unsuccessful (69).

Subsequently, instead of using site-directed mutagenesis, Yep and McLeish (68,70) decided to employ SSM. SSM is most successful when residues are within 5-6 Å of the substrate (3). Based on this premise, 15 residues from BFDC were identified that had counterparts in PDC. Of these, 12 residues were used in the SSM experiment (70). Following mutagenesis, the resulting colonies were screened using a microplate assay that had been designed to identify BFDC variants with changes in activity (68). Once the activities were determined with a range of substrates, the data was plotted using radar plots (Figure 5.2) and two variants, T377L and A460Y were identified as having a favorable change in activity towards pyruvate (70).

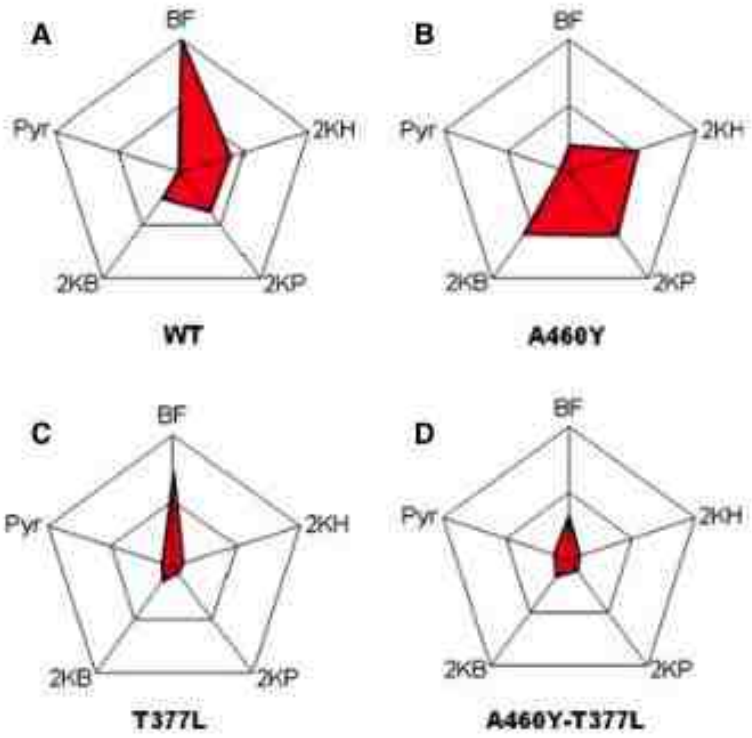

Figure 5.2. Radar plots illustrating the activities of the variants of the SSM experiments. The shape of the radar plot (red) is indicative of the substrate specificity change for each variant where (A) is wild-type BFDC, (B) is the A460Y variant, (C) is the T337L variant, and (D) is the T377L-A460Y double variant. The substrates tested include benzoylformate (BF), pyruvate (Pyr), 2-ketobutanoate (2KB), 2-ketopentanoate (2KP), and 2-ketohexanoate (2KH). Figure taken from Yep and McLeish (70).

Looking at Figure 5.2, it can be seen when comparing the individual radar plots that there is a clear decrease in the activity of wild-type BFDC compared to the variants. Wild-type BFDC (A) has better activity with the large substrates and none at all with pyruvate. The A460Y variant (B) had reduced activity with benzoylformate but significantly more activity with the smaller aliphatic substrates, 2-ketopentanaote and 2-ketobutanoate. The T377L variant (C) showed less of an effect with benzoylformate but did show enhanced activity with pyruvate. A second round of mutagenesis provided the T377L-A460Y

double variant (D) which has reduced activity overall but has the highest activity with pyruvate. This variant had a K_m value for pyruvate of 2.0 mM and a k_{cat} value of 1.8 s^{-1} (70). The K_m value is within the range of values seen for native PDCs, but the k_{cat} value is considerably (20 to 50-fold) lower. Nonetheless this is a solid beginning to evolving BFDC to a PDC. It remains to be seen whether the activity that has been developed is sufficient to replace PDC *in vivo*, or whether further optimization will be required. Both of these will require the development of a selection method. This chapter describes the initial approaches to that goal.

5.2 Results and Discussion

5.2.1 The substrate spectrum of BFDC T377L-A460Y

BFDC T377L-A460Y was expressed and purified (section 2.4) with a yield of 30 mg/L active enzyme. Once purified, the protein could be stored for months at -80 °C without significant loss in activity. The purified enzyme was then assayed with a variety of substrates under steady-state conditions using the standard coupled assay (section 2.6.2). The results are listed in Table 5.1.

The full steady-state kinetic study of BFDC T377L-A460Y demonstrated that this variant exhibited Michaelis-Menten type kinetics (Figure 5.3), which is consistent with wt BFDC. There was no evidence of allosteric activation nor was there any evidence of substrate inhibition.

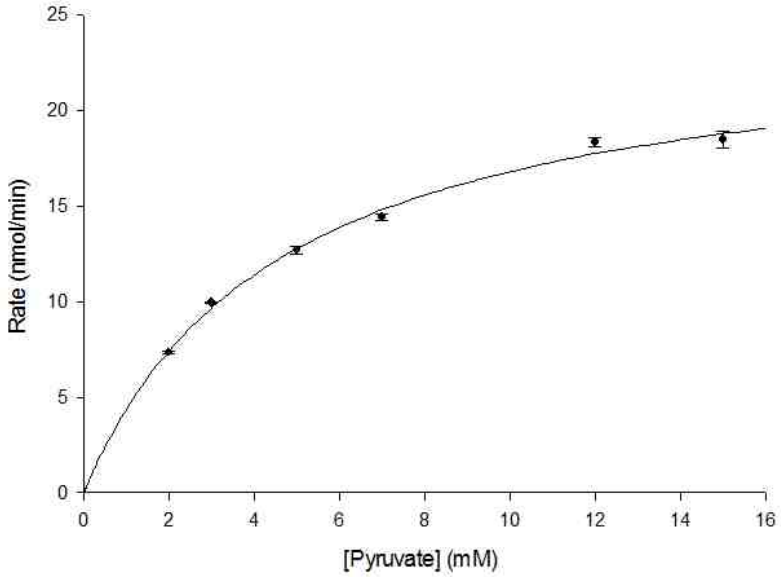

Figure 5.3. Michaelis-Menten plot for the reaction of BFDC A460Y-T-377L with pyruvate.

Table 5.1. Steady-state Kinetic Characterization of BFDC T377L-A460Y[a]

Substrate	K_m **(mM)**	k_{cat} **(s^{-1})**	k_{cat}/K_m **($mM^{-1}s^{-1}$)**
Pyruvate	4.6 ± 0.3	3.3 ± 0.1	0.72
2-Ketobutanoate	0.39 ± 0.02	6.0 ± 0.1	16
2-Ketopentanoate	0.15 ± 0.01	1.1 ± 0.1	7.5
2-Ketohexanoate	0.14 ± 0.01	5.9 ± 0.2	42
3-Methyl-2-ketobutanoate	0.77 ± 0.06	12 ± 0.4	15
3-Methyl-2-ketopentanoate	1.2 ± 0.1	16 ± 0.7	13
4-Methyl-2-ketopentanoate	0.14 ± 0.02	2.7 ± 0.2	20
4-Methylthio-2-ketobutanoate	0.14 ± 0.01	4.6 ± 0.1	33
Phenylpyruvate	0.82 ± 0.06	2.0 ± 0.1	2.5
Benzoylformate	0.45 ± 0.04	15 ± 0.6	34
[a] Values are the mean of 3 independent determinations ± standard error.			

Overall, BFDC T377L-A460Y exhibited the greatest activity with 2-ketohexanoate with a k_{cat}/K_m value of 42 $mM^{-1}s^{-1}$. Benzoylformate and 4-methylthio-2-ketobutanoate (a product of methionine metabolism), had similar activity. Interestingly, the double mutant exhibited some level of activity with all substrates tested and, with the exception of pyruvate, the K_m values for all substrates are very similar. There was more variation in the k_{cat} values but, even then, they are all within an order of magnitude.

Table 5.2. Kinetic values for wild-type BFDC and BFDC T377L-A460Y determined by Yep and McLeish. (70).

	wt BFDC			**BFDC T377L-A460Y**		
Substrate	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1}s^{-1}$)	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($mM^{-1}s^{-1}$)
Pyruvate	-	-	0.012	2.0±0.1	3.6± 0.3	1.8
2-ketobutanoate	7.5±0.6	4±1	0.5	0.42±0.02	5.8± 0.2	14
2-ketopentanoate	6.0±0.5	11±1	1.8	0.10±0.01	0.81 ± 0.05	8.1
2-ketohexanoate	4.1±0.1	5.3± 0.3	1.3	0.15±0.02	1.2± 0.1	8
Benzoylformate	0.27±0.02	320± 5	1180	1.5±0.2	25±2	16

The kinetic constants obtained for pyruvate, 2-ketobutanoate, 2-ketopentanoate, 2-ketohexanoate, and benzoylformate in this study are broadly similar to those values determined by Yep and McLeish. (Table 5.2) (70). Taken together, it is clear that wt BFDC does not really accept pyruvate as a substrate but the double mutant displayed reasonable activity albeit significantly less than *Sc*PDC (K_m=1.3±0.1 mM, k_{cat}=60±2 s^{-1}, k_{cat}/Km=45 $mM^{-1}s^{-1}$) and *Zm*PDC

(K_m=1.1±0.1 mM, k_{cat}=486±49 s^{-1}, k_{cat}/K_m=440 $mM^{-1}s^{-1}$)(4). It would seem likely that further improvement will be required to fully convert BFDC into a good PDC. Potentially, this can be achieved by the use of saturation mutagenesis of multiple sites. However, a better selection method is needed.

5.2.2 Construction of yeast expression vectors

As noted above, one of the downfalls of using SSM is the large number of colonies that must be screened to ensure that full coverage of every amino acid has been obtained at the site that was targeted for mutagenesis. Using selection over screening can mitigate this issue. Our lab has begun developing a method utilizing yeast vectors for controlling expression levels in order to enhance the selection process and thereby address the numbers problem in SSM.

Mumberg *et al.* (25) first described a convenient series of yeast expression vectors that are useful in that they possess multiple restriction sites as well as the ability for controlling the level of gene expression. Figure 5.4 depicts the structure of the p4XX vector series, the various promoters, and the marker genes (HIS3, TRP1, LEU2, URA3) that can be used. Each plasmid series, depending on the marker, shares the same restriction sites but vary in those that are single cut sites. The promoters that Mumberg *et al.* (25) selected to construct the vectors were derived from genes encoding cytochrome-c oxidase (CYC1), alcohol dehydrogenase (ADH), translation elongation factor 1α (TEF), and glyceraldehyde-3-phosphate dehydrogenase (GPD).

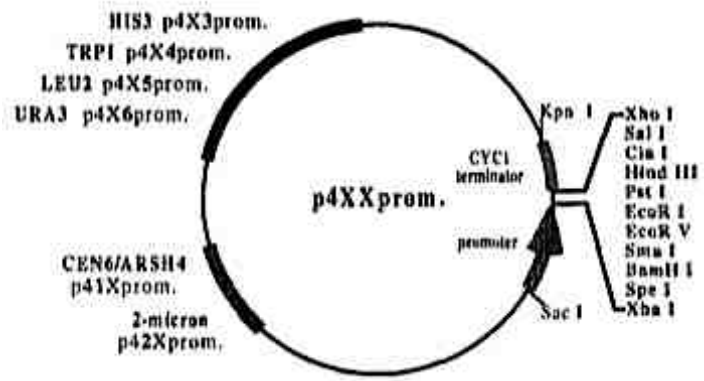

Figure 5.4. Structure of the expression vectors. The plasmid map shows the location of the restriction sites between the promoter and terminator (left) and the marker genes (right) for each of the types of plasmid. Figure taken from Mumberg *et al.* (25).

Once the plasmids were constructed, the expression levels were determined by cloning a *lacZ* gene into the vectors and determining the β-galactosidase activity under the control of each promoter (Figure 5.5). Using the p426 plasmid series as an example, it can be seen that the GPD promoter is the strongest followed by the TEF promoter, ADH promoter, and finally the CYC1 promoter, which has the lowest activity.

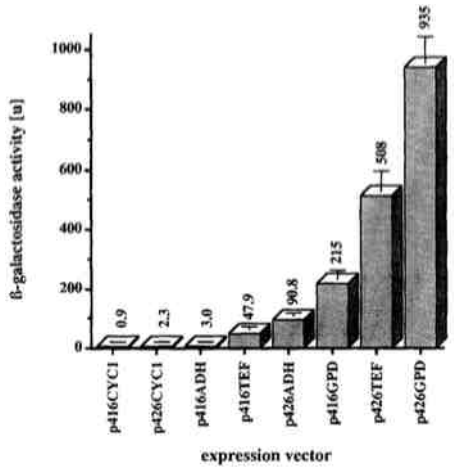

Figure 5.5. The expression levels of the different promoter systems determined by monitoring β-galactosidase activity. Figure taken from Mumberg *et al.* (25).

This range of promoter activity is important, as one of the drawbacks of selection is that once an organism achieves a level of sustainability, there is no longer a need for the organism to evolve and improve. To alleviate this problem, our laboratory is developing a way to use the p426 vector system to select for activity after a SSM experiment. In theory, if a SSM experiment is carried out using the p426GPD expression vector then expression of the gene undergoing mutagenesis should be very high. Thus a yeast knockout could be viable with a relatively low level of enzyme activity. However, if the gene is placed behind the p426CYC1 promoter, then only those mutants with high activity should be able to survive. In the case of yeast, a pyruvate decarboxylase is required to be able to grow on a glucose media. As mentioned previously, *S. cerevisiae* possesses three genes, PDC1, PDC5, and PDC6, which are or can be capable of acting as pyruvate decarboxylases. If a knockout is prepared that lacks all three PDCs, then another enzyme that has some PDC activity, such as the BFDC T377L-A460Y variant, could potentially be used to regain viability on glucose. If this, or any other variant placed in the p426GPD expression vector can successfully replace PDC, it can be tried for activity with a weaker promoter system. These BFDC variants could be subject to successive rounds of SSM, with the best variants being placed under more stringent conditions Eventually, only those variants with the highest activity would be obtained since they would have enough activity to overcome the lower expression levels.

This process stands to greatly reduce the number of colonies needing to be screened in SSM by selecting for only those colonies with high levels of

activity and eliminating those with little to no activity. As can be seen from Figure 5.5, this series of vectors could potentially provide a 1000-fold change in expression levels.

To explore this possibility, a series of p426GPD expression vectors was constructed using the methodology outlined in Figure 5.6. The p426GPD expression vector was the initial vector of choice as it used the strongest promoter. However, the available yeast PDC knockout strains used the same URA3 marker and, therefore, the p426 series of plasmids could not be used for selection without some modification. Consequently, PCR was performed to insert an SphI site. This enables the URA3 marker gene to be excised and the zeocin resistance gene inserted in its place via an SphI/ApaI digestion and ligation to form the p426GPD-Zeo plasmid.

Following the addition of the zeocin cassette to the p426GPD expression vector, the p426GPD-Zeo plasmid was digested with BamHI and religated to remove extraneous EcoRI and BamHI sites. This provided the p426GPD-Zeo-BamHI plasmid. Subsequently, a HindIII site was added via PCR mutagenesis to form the p426GPD-Zeo-HindIII plasmid. After this was complete, the genes for *Sc*PDC1, *Sc*PDC5, *Zm*PDC, BFDC, and BFDC T377L-A460Y were inserted into the p426GPD-Zeo-HindIII plasmid as described in section 2.3.3. Once the plasmids were successfully constructed, they were used in yeast growth experiments.

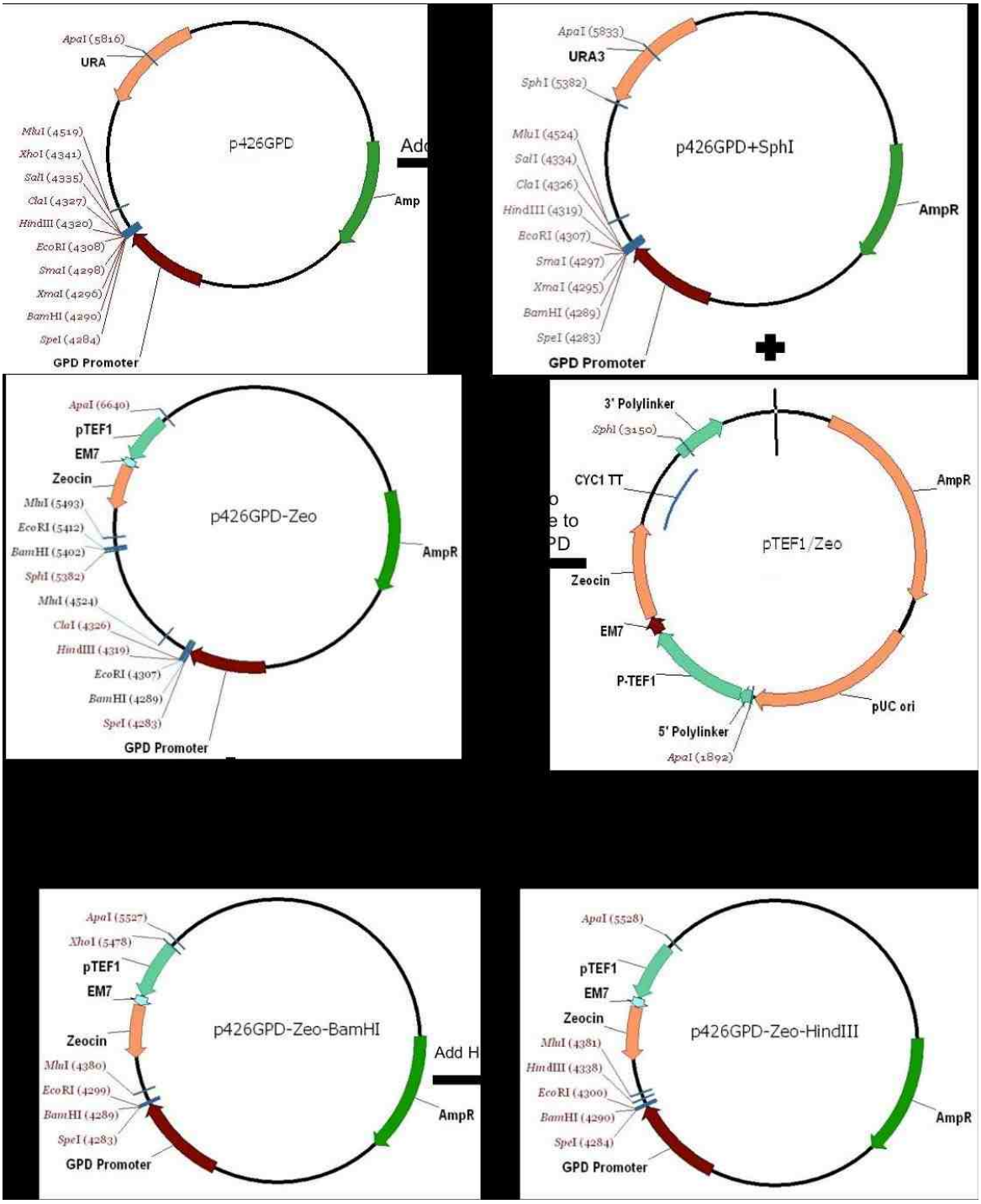

Figure 5.6. A flowchart depicting the construction of the p426GPD-Zeo expression vector. *E.coli* was used to complete all DNA work.

5.2.3 Yeast growth experiments

Yeast PDC knockout strains (section 2.2.2) were stored at -80 °C. Fresh colonies were obtained by streaking on YEP plates containing 2% ethanol. These colonies were then streaked onto fresh YEP/2% ethanol plates as well as YPD plates, which contain glucose (dextrose). The knockout strains should not be able to grow on the YPD plates because they lack pyruvate decarboxylase activity and, therefore, cannot metabolize glucose. Indeed, no growth was observed on the YPD plates but growth was observed on the YEP/2% ethanol plates. This confirmed that the knockouts were indeed functional. At this point, a single colony was selected and a transformation (Chapter 2) was carried out using the plasmid pRUL178, which contains the *Sc*PDC gene. The transformed cells were plated on YEP/2% ethanol and YPD plates. The results, provided in Figure 5.7, show the knockout yeast are unable to grow on YPD (left) but can maintain growth on YEP/2% ethanol (right). However, those yeast transformed with the plasmid containing the *Sc*PDC are clearly able to grow on the YPD medium.

Figure 5.7. Results from the transformation of pRUL178 into the knockout yeast strain (5.127-17C). The 500 represent the strain number used in the laboratory and the 500+ indicate the knockout plus pRUL178. The YEP/2% ethanol plate (right) has growth with and without pRUL178 plasmid and the YPD plate (left) only has growth with the addition of pRUL178 plasmid.

Subsequently, the knockout yeast (strain 5.127.17C) was transformed with the newly constructed plasmids and plated on YPD plates. No colonies were observed with the PDC5, BFDC, and BFDC T377L-A460Y transformations. However, transformations with the PDC1 and *Zm*PDC genes were successful (Figure 5.8). In case transformation efficiency was a problem, the transformations were repeated using several different protocols but the end result was the same each time.

The results from this experiment answer several questions. First, the presence of transformants confirms that the plasmid constructs were successfully created. Second, as noted in the introduction, PDCs from yeast are allosterically activated but bacterial PDCs are not. Based on the success of the *Zm*PDC transformation, it would seem that an allosteric PDC is not required for successful growth of yeast.

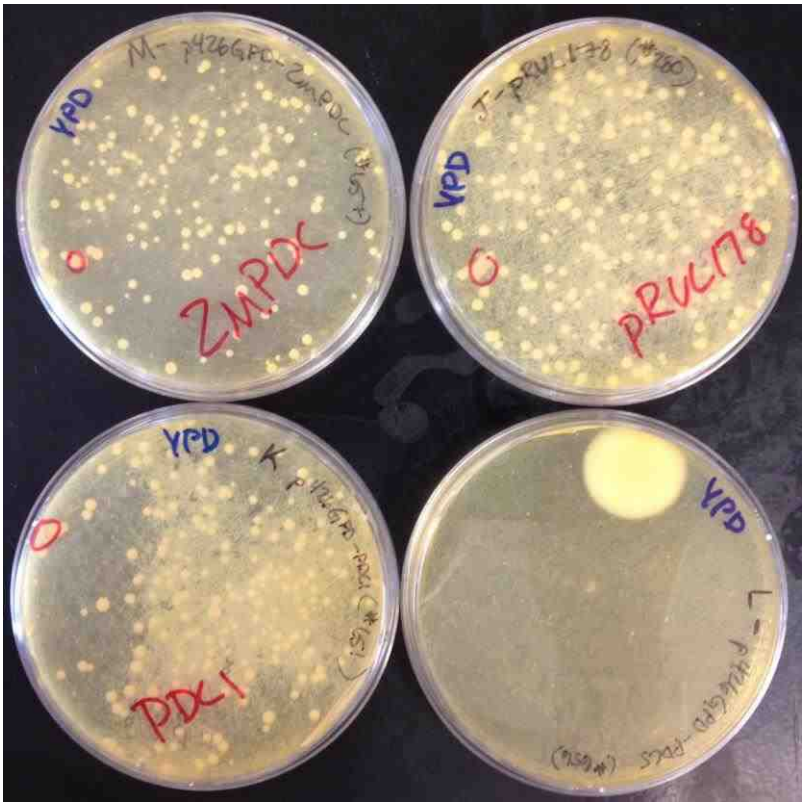

Figure 5.8. Results of the p426GPD-Zeo plasmid series transformation. The top left is *Zm*PDC, top right is the untagged PDC in the pRUL178 plasmid, bottom left is PDC1, and bottom right the unsuccessful PDC5. There is slight contamination on the PDC5 transformation but the contamination is very different in appearance from the yeast colonies observed on the other 3 plates.

It is not surprising that the BFDC transformation was unsuccessful given that BFDC has no activity with pyruvate. It was more disappointing, but not entirely unexpected, that the BFDC T377L-A460Y variant does not possess enough PDC activity to support the growth of yeast. Despite this, given that we have confirmed that an allosteric enzyme is not required for yeast growth, it would appear that an appropriately evolved BFDC variant does have the possibility of replacing PDC. Clearly, further rounds of mutagenesis are necessary to increase activity with pyruvate. Finally, perhaps the most

surprising/disappointing result was the lack of success of the PDC5 transformation. Based on the activities determined for PDC5 (Chapter 4), this isozyme should be able to support growth but did not. Right now, the reasons for this are not clear.

In addition to the transformations described above, the same series of transformations, including the various conditions tried, were plated on YEP/2% ethanol plates with the added antibiotic (zeocin) at recommended concentrations (25-300 μg/mL). These transformations were not successful with any of the plasmids nor with any of the variety of zeocin concentrations used. Intriguingly, wild-type baker's yeast transformed with only a zeocin containing plasmid was also not able to grow on zeocin containing plates. These plasmids are shuttle vectors and when used with *E. coli*, transformants were able to grow on zeocin/LB plates. This suggests that the zeocin resistance factor was functional, just not in yeast. Given this problem, and that with the URA3 marker described earlier, a number of alternative knockout strains were obtained (CEN.PK711-7C, IMZ001, and IMZ002). These strains were checked on YPD and YEP/2% ethanol plates to ensure that they were not able to grow on glucose. A second check was carried out by transforming with the pRUL178 plasmid and plating on YPD as described above. Each of the strains could grow on glucose only if pRUL178 (containing PDC1) was present.

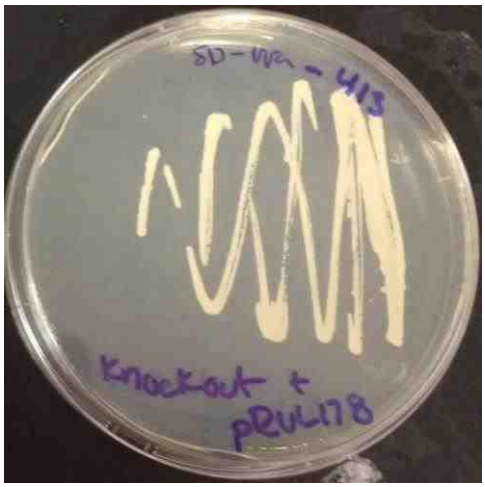

Figure 5.9. The knockout PDC yeast was transformed with pRUL178 and was able to grow on SD-URA media.

The pRUL178 plasmid also contains the URA3 marker gene. The knockout yeast containing the pRUL178 plasmid was plated on SD media lacking uracil (-URA) and 2% ethanol. The URA3 gene, which encodes orotidine 5'-phosphate decarboxylase, allows the synthesis of uracil, a component of RNA. If yeast is lacking the URA3 marker gene and is plated on media without uracil, they will not be able to grow. However, the yeast transformed with pRUL178 were able to grow as seen in Figure 5.9. This implies both the PDC1 and the URA3 gene were operational. Currently, plasmid construction of PDC1, PDC5, *Zm*PDC, BFDC, and BFDC T377L-A460Y with a URA3 marker instead of the zeocin cassette is underway.

5.3 Conclusions and Future Directions

We have fully characterized BFDC T377L-A460Y and demonstrated a shift in the substrate specificity that is in agreement with that observed by Yep

and McLeish (70). Further, we have presented a plausible plan utilizing selection for reducing the number of colonies that need to be screened during a SSM experiment. We have also demonstrated that allosteric activation is not a requirement for yeast growth utilizing glucose as a carbon source. In the future, further work is needed to optimize the yeast transformation procedure to ensure the most efficient transformation possible. Also, plasmids will be constructed replacing the zeocin cassette with the URA3 gene to determine if that will aid in a more efficient selection method. After those steps have been completed, multi-site saturation mutagenesis will be employed to make new variants of BFDC T377L-A460Y. The newly developed selection system will be used to identify variants with enhanced PDC activity.

LIST OF REFERENCES

LIST OF REFERENCES

1. Kluger, R., and Tittmann, K. (2008) Thiamin diphosphate catalysis: enzymic and nonenzymic covalent intermediates. *Chemical Reviews* **108**, 1797-1833

2. Kaplun, A., Binshtein, E., Vyazmensky, M., Steinmetz, A., Barak, Z., Chipman, D. M., Tittmann, K., and Shaanan, B. (2008) Glyoxylate carboligase lacks the canonical active site glutamate of thiamine-dependent enzymes. *Nature Chemical Biology* **4**, 113-118

3. Andrews, F. H., and McLeish, M. J. (2013) Using site-saturation mutagenesis to explore mechanism and substrate specificity in thiamin diphosphate-dependent enzymes. *FEBS Journal* **280**, 6395-6411

4. Andrews, F. H., and McLeish, M. J. (2012) Substrate specificity in thiamin diphosphate-dependent decarboxylases. *Bioorganic Chemistry* **43**, 26-36

5. Furey, W., Arjunan, P., Chen, L., Sax, M., Guo, F., and Jordan, F. (1998) Structure-function relationships and flexible tetramer assembly in pyruvate decarboxylase revealed by analysis of crystal structures. *Biochimica et Biophysica Acta* **1385**, 253-270

6. Duggleby, R. G. (2006) Domain relationships in thiamine diphosphate-dependent enzymes. *Accounts of Chemical Research* **39**, 550-557

7. Joseph, E., Wei, W., Tittmann, K., and Jordan, F. (2006) Function of a conserved loop of the beta-domain, not involved in thiamin diphosphate binding, in catalysis and substrate activation in yeast pyruvate decarboxylase. *Biochemistry* **45**, 13517-13527

8. Killenberg-Jabs, M., Jabs, A., Lilie, H., Golbik, R., and Hubner, G. (2001) Active oligomeric states of pyruvate decarboxylase and their functional characterization. *European Journal of Biochemistry* **268**, 1698-1704

9. Kutter, S., Spinka, M., Koch, M. H., and Konig, S. (2007) The influence of protein concentration on oligomer structure and catalytic function of two pyruvate decarboxylases. *The Protein Journal* **26**, 585-591

10. Dyda, F., Furey, W., Swaminathan, S., Sax, M., Farrenkopf, B., and Jordan, F. (1993) Catalytic centers in the thiamin diphosphate dependent enzyme pyruvate decarboxylase at 2.4-A resolution. *Biochemistry* **32**, 6165-6170

11. Candy, J. M., and Duggleby, R. G. (1998) Structure and properties of pyruvate decarboxylase and site-directed mutagenesis of the *Zymomonas mobilis* enzyme. *Biochimica et Biophysica Acta* **1385**, 323-338

12. Guo, F., Zhang, D., Kahyaoglu, A., Farid, R. S., and Jordan, F. (1998) Is a hydrophobic amino acid required to maintain the reactive V conformation of thiamin at the active center of thiamin diphosphate-requiring enzymes? Experimental and computational studies of isoleucine 415 of yeast pyruvate decarboxylase. *Biochemistry* **37**, 13379-13391

13. Andrews, F. H., Tom, A. R., Gunderman, P. R., Novak, W. R., and McLeish, M. J. (2013) A bulky hydrophobic residue is not required to maintain the V-conformation of enzyme-bound thiamin diphosphate. *Biochemistry* **52**, 3028-3030

14. Liu, M., Sergienko, E. A., Guo, F., Wang, J., Tittmann, K., Hubner, G., Furey, W., and Jordan, F. (2001) Catalytic acid-base groups in yeast pyruvate decarboxylase. 1. Site-directed mutagenesis and steady-state kinetic studies on the enzyme with the D28A, H114F, H115F, and E477Q substitutions. *Biochemistry* **40**, 7355-7368

15. Sergienko, E. A., and Jordan, F. (2001) Catalytic acid-base groups in yeast pyruvate decarboxylase. 2. Insights into the specific roles of D28 and E477 from the rates and stereospecificity of formation of carboligase side products. *Biochemistry* **40**, 7369-7381

16. Hasson, M. S., Muscate, A., McLeish, M. J., Polovnikova, L. S., Gerlt, J. A., Kenyon, G. L., Petsko, G. A., and Ringe, D. (1998) The crystal structure of benzoylformate decarboxylase at 1.6 A resolution: diversity of catalytic residues in thiamin diphosphate-dependent enzymes. *Biochemistry* **37**, 9918-9930

17. Baburina, I., Dikdan, G., Guo, F., Tous, G. I., Root, B., and Jordan, F. (1998) Reactivity at the substrate activation site of yeast pyruvate decarboxylase: inhibition by distortion of domain interactions. *Biochemistry* **37**, 1245-1255

18. Koenig, S. S., Michael; Kutter, Steffen. (2009) Allosteric activation of pyruvate decarboxylases. A never-ending story? in *Journal of Molecular Catalysis B: Enzymatic*

19. Pohl, M., Sprenger, G. A., and Muller, M. (2004) A new perspective on thiamine catalysis. *Current Opinion in Biotechnology* **15**, 335-342

20. Costelloe, S. J., Ward, J. M., and Dalby, P. A. (2008) Evolutionary analysis of the TPP-dependent enzyme family. *Journal of Molecular Evolution* **66**, 36-49

21. Kelley, L. A., and Sternberg, M. J. (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nature Protocols* **4**, 363-371

22. Chen, D. C., Yang, B. C., and Kuo, T. T. (1992) One-step transformation of yeast in stationary phase. *Current Genetics* **21**, 83-84

23. Hohmann, S. (1991) Characterization of PDC6, a third structural gene for pyruvate decarboxylase in *Saccharomyces cerevisiae*. *Journal of Bacteriology* **173**, 7963-7969

24. Romagnoli, G., Luttik, M. A., Kotter, P., Pronk, J. T., and Daran, J. M. (2012) Substrate specificity of thiamine pyrophosphate-dependent 2-oxo-acid decarboxylases in *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology* **78**, 7538-7548

25. Mumberg, D., Muller, R., and Funk, M. (1995) Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* **156**, 119-122

26. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248-254

27. Johnen, S., G.A. Sprenger. (2009) Characterization of recombinant thiamine diphosphate-dependent phosphonopyruvate decarboxylase from *Streptomyces viridochromogenes Tu494*. *Journal of Molecular Catalysis B: Enzymatic* **61**, 39-46

28. Zhang, G., Dai, J., Lu, Z., and Dunaway-Mariano, D. (2003) The phosphonopyruvate decarboxylase from *Bacteroides fragilis*. *The Journal of Biological Chemistry* **278**, 41302-41308

29. Hubner, G., Weidhase, R., and Schellenberger, A. (1978) The mechanism of substrate activation of pyruvate decarboxylase: a first approach. *European Journal of Biochemistry* **92**, 175-181

30. Sun, S. S., Gregory S.; O'Leary, Marion H.; Schowen, Richard L. (1997) The linkage of catalysis and regulation in enzyme action. Fluoropyruvate as a probe of regulation in pyruvate decarboxylases. *Journal of the American Chemical Society.* **119**, 1507-1515

31. Burton, D. J. S., Lee G. (1988) Preparation of difluorophosphonoacetic acid and its derivatives. *Journal of Organic Chemistry* **53**, 1523-1527

32. Nemeria, N., Baykal, A., Joseph, E., Zhang, S., Yan, Y., Furey, W., and Jordan, F. (2004) Tetrahedral intermediates in thiamin diphosphate-dependent decarboxylations exist as a 1',4'-imino tautomeric form of the coenzyme, unlike the michaelis complex or the free coenzyme. *Biochemistry* **43**, 6565-6575

33. Nemeria, N. S., Chakraborty, S., Balakrishnan, A., and Jordan, F. (2009) Reaction mechanisms of thiamin diphosphate enzymes: defining states of ionization and tautomerization of the cofactor at individual steps. *FEBS Journal* **276**, 2432-2446

34. Patel, H., Nemeria, N. S., Andrews, F. H., McLeish, M. J., and Jordan, F. (2014) Identification of charge transfer transitions related to thiamin-bound intermediates on enzymes provides a plethora of signatures useful in mechanistic studies. *Biochemistry* **53**, 2145-2152

35. Brandt, G. S., Kneen, M. M., Chakraborty, S., Baykal, A. T., Nemeria, N., Yep, A., Ruby, D. I., Petsko, G. A., Kenyon, G. L., McLeish, M. J., Jordan, F., and Ringe, D. (2009) Snapshot of a reaction intermediate: analysis of benzoylformate decarboxylase in complex with a benzoylphosphonate inhibitor. *Biochemistry* **48**, 3247-3257

36. Flikweert, M. T., de Swaaf, M., van Dijken, J. P., and Pronk, J. T. (1999) Growth requirements of pyruvate-decarboxylase-negative *Saccharomyces cerevisiae*. *FEMS Microbiology Letters* **174**, 73-79

37. Schaaff, I., Green, J. B., Gozalbo, D., and Hohmann, S. (1989) A deletion of the PDC1 gene for pyruvate decarboxylase of yeast causes a different phenotype than previously isolated point mutations. *Current Genetics* **15**, 75-81

38. Seeboth, P. G., Bohnsack, K., and Hollenberg, C. P. (1990) pdc1(0) mutants of *Saccharomyces cerevisiae* give evidence for an additional structural PDC gene: cloning of PDC5, a gene homologous to PDC1. *Journal of Bacteriology* **172**, 678-685

39. Hohmann, S., and Cederberg, H. (1990) Autoregulation may control the expression of yeast pyruvate decarboxylase structural genes PDC1 and PDC5. *European Journal of Biochemistry* **188**, 615-621

40. Muller, E. H., Richards, E. J., Norbeck, J., Byrne, K. L., Karlsson, K. A., Pretorius, G. H., Meacock, P. A., Blomberg, A., and Hohmann, S. (1999) Thiamine repression and pyruvate decarboxylase autoregulation independently control the expression of the *Saccharomyces cerevisiae* PDC5 gene. *FEBS Letters* **449**, 245-250

41. Eberhardt, I., Cederberg, H., Li, H., Konig, S., Jordan, F., and Hohmann, S. (1999) Autoregulation of yeast pyruvate decarboxylase gene expression requires the enzyme but not its catalytic activity. *European Journal Of Biochemistry* **262**, 191-201

42. Konig, S. (1998) Subunit structure, function and organisation of pyruvate decarboxylases from various organisms. *Biochimica et Biophysica Acta* **1385**, 271-286

43. Hohmann, S. (1991) PDC6, a weakly expressed pyruvate decarboxylase gene from yeast, is activated when fused spontaneously under the control of the PDC1 promoter. *Current Genetics* **20**, 373-378

44. Fauchon, M., Lagniel, G., Aude, J. C., Lombardia, L., Soularue, P., Petat, C., Marguerie, G., Sentenac, A., Werner, M., and Labarre, J. (2002) Sulfur sparing in the yeast proteome in response to sulfur demand. *Molecular Cell* **9**, 713-723

45. Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B., Storz, G., Botstein, D., and Brown, P. O. (2000) Genomic expression programs in the response of yeast cells to environmental changes. *Molecular Biology of the Cell* **11**, 4241-4257

46. Boer, V. M., de Winde, J. H., Pronk, J. T., and Piper, M. D. (2003) The genome-wide transcriptional responses of *Saccharomyces cerevisiae* grown on glucose in aerobic chemostat cultures limited for carbon, nitrogen, phosphorus, or sulfur. *Journal of Biological Chemistry* **278**, 3265-3274

47. Hohmann, S., and Meacock, P. A. (1998) Thiamin metabolism and thiamin diphosphate-dependent enzymes in the yeast *Saccharomyces cerevisiae*: genetic regulation. *Biochimica et Biophysica Acta* **1385**, 201-219

48. Begley, T. P. (1996) The biosynthesis and degradation of thiamin (vitamin B1). *Natural Product Reports* **13**, 177-185

49. Nosaka, K., Onozuka, M., Konno, H., Kawasaki, Y., Nishimura, H., Sano, M., and Akaji, K. (2005) Genetic regulation mediated by thiamin pyrophosphate-binding motif in *Saccharomyces cerevisiae*. *Molecular Microbiology* **58**, 467-479

50. Nishimura, H., Kawasaki, Y., Kaneko, Y., Nosaka, K., and Iwashima, A. (1992) A positive regulatory gene, THI3, is required for thiamine metabolism in *Saccharomyces cerevisiae*. *Journal of Bacteriology* **174**, 4701-4706

51. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948

52. Reetz, M. T. (2011) Laboratory evolution of stereoselective enzymes: a prolific source of catalysts for asymmetric reactions. *Angewandte Chemie* **50**, 138-174

53. Li, H., and Jordan, F. (1999) Effects of substitution of tryptophan 412 in the substrate activation pathway of yeast pyruvate decarboxylase. *Biochemistry* **38**, 10004-10012

54. Boiteux, A., and Hess, B. (1970) Allosteric properties of yeast pyruvate decarboxylase. *FEBS Letters* **9**, 293-296

55. Kutter, S., Weiss, M. S., Wille, G., Golbik, R., Spinka, M., and Konig, S. (2009) Covalently bound substrate at the regulatory site of yeast pyruvate decarboxylases triggers allosteric enzyme activation. *Journal of Biological Chemistry* **284**, 12136-12144

56. Balakrishnan, A., Gao, Y., Moorjani, P., Nemeria, N. S., Tittmann, K., and Jordan, F. (2012) Bifunctionality of the thiamin diphosphate cofactor: assignment of tautomeric/ionization states of the 4'-aminopyrimidine ring when various intermediates occupy the active sites during the catalysis of yeast pyruvate decarboxylase. *Journal of the American Chemical Society* **134**, 3873-3885

57. Butt, T. R., Edavettal, S. C., Hall, J. P., and Mattern, M. R. (2005) SUMO fusion technology for difficult-to-express proteins. *Protein Expression and Purification* **43**, 1-9

58. Dalby, P. A. (2011) Strategy and success for the directed evolution of enzymes. *Current Opinion in Structural Biology* **21**, 473-480

59. Reetz, M. T., Kahakeaw, D., and Lohmer, R. (2008) Addressing the numbers problem in directed evolution. *Chembiochem : A European Journal of Chemical Biology* **9**, 1797-1804

60. Leung, D. W., Chen, E., and Goeddel, D. V. (1989) A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Technique* **1**, 11-15

61. Stemmer, W. P. (1994) Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370**, 389-391

62. Reetz, M. T. (2012) Laboratory evolution of stereoselective enzymes as a means to expand the toolbox of organic chemists. *Tetrahedron* **68**, 7530-7548

63. Siloto, R. M. P., Weselake, R.J. (2012) Site saturation mutagenesis: Methods and applications in protein engineering *Biocatal Agric Biotechnol* **1**, 181-189

64. Reetz, M. T., and Carballeira, J. D. (2007) Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nature Protocols* **2**, 891-903

65. Reetz, M. T., Carballeira, J. D., Peyralans, J., Hobenreich, H., Maichele, A., and Vogel, A. (2006) Expanding the substrate scope of enzymes: combining mutations obtained by CASTing. *Chemistry* **12**, 6031-6038

66. Leemhuis, H., Kelly, R. M., and Dijkhuizen, L. (2009) Directed evolution of enzymes: Library screening strategies. *IUBMB Life* **61**, 222-228

67. Reetz, M. T., Hobenreich, H., Soni, P., and Fernandez, L. (2008) A genetic selection system for evolving enantioselectivity of enzymes. *Chemical Communications*, 5502-5504

68. Yep, A., Kenyon, G. L., and McLeish, M. J. (2008) Saturation mutagenesis of putative catalytic residues of benzoylformate decarboxylase provides a challenge to the accepted mechanism. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 5733-5738

69. Siegert, P., McLeish, M. J., Baumann, M., Iding, H., Kneen, M. M., Kenyon, G. L., and Pohl, M. (2005) Exchanging the substrate specificities of pyruvate decarboxylase from *Zymomonas mobilis* and benzoylformate decarboxylase from *Pseudomonas putida*. *Protein engineering, design & selection : PEDS* **18**, 345-357

70. Yep, A., and McLeish, M. J. (2009) Engineering the substrate binding site of benzoylformate decarboxylase. *Biochemistry* **48**, 8387-8395