2010

Structure-Function Relationship Studies of the UDP-Glucose Pyrophosphorylase From Escherichia Coli

Agnieszka Maria Orlof

Loyola University Chicago

Recommended Citation
http://ecommons.luc.edu/luc_theses/558

This Thesis is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Master’s Theses by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License.
Copyright © 2010 Agnieszka Maria Orlof
LOYOLA UNIVERSITY CHICAGO

STRUCTURE-FUNCTION RELATIONSHIP
STUDIES OF
THE UDP-GLUCOSE PYROPHOSPHORYLASE
FROM *ESCHERICHIA COLI*

A THESIS SUBMITTED TO
THE FACULTY OF THE GRADUATE SCHOOL
IN CANDIDACY FOR THE DEGREE OF
MASTER OF SCIENCE

PROGRAM IN CHEMISTRY

BY

AGNIESZKA M. ORLOF

CHICAGO, ILLINOIS

AUGUST 2010
Copyright by Agnieszka M. Orlof, 2010
All rights reserved.
# TABLE OF CONTENTS

## LIST OF FIGURES

v

## LIST OF ABBREVIATIONS

vii

## CHAPTER ONE: INTRODUCTION

1

- Lipopolysaccharides
- UDP-glucose in Prokaryotes
- UDP-glucose in Eukaryotes
- Kinetic mechanism UDP-Glc PPase
- Three dimensional structure of *E. coli’s* UDP-Glc PPase
- Important residues for the substrate binding
- In vivo regulation
- Evolution of UDP-Glc PPase
- Tissue and subcellular localization
- GalF protein

## CHAPTER TWO: GOAL OF THIS PROJECT

14

- Overall goal of the project
- Part I: Biochemical exploration of UDP-Glucose Pyrophosphorylase (GalU) in *E.coli*
- Part II: Investigation of putative UDP-Glucose Pyrophosphorylase (GalF) in *E.coli*

## CHAPTER THREE: RESULTS AND DISCUSSION

20

- Part I: Expression, purification and characterization of GalU
  - Kinetic analysis
  - Magnesium curve
  - Kinetic analysis of GalU mutants
- Conclusions
- Part II: Expression, purification and characterization of GalF
  - Kinetic analysis
  - Magnesium curve
  - Kinetic comparisons of GalU, GalF and GalF mutant
- Conclusions

## CHAPTER FOUR: MATERIALS AND METHODS

44

- Structure prediction and homology modeling
- Cloning of galU and galF genes from *E. coli* genome
- Primers for cloning of galU gene from *E.coli* genome
- Primers for cloning of galF gene from *E.coli* genome
- Site-directed mutagenesis
Primers used for site-directed mutagenesis 49
  galU 49
  galF 49
Flanking primers for galU and galF genes 49
pMCSG9 vector 49
Subcloning genes into pMCSG9 vectors 52
Colony PCR procedure 53
DNA sequence visualization 53
pET vector 53
Subcloning genes into pET vectors 54
Expression and purification of GalU and GalF from pET vectors 54
Expression of GalU and GalF from pMCSG9 vectors 55
Purification of GalU and GalF from pMCSG9 vectors 56
Measuring protein concentration 57
SDS-PAGE preparation 57
Enzyme assay: Pyrophosphorolysis direction 58
Calculation of kinetic constants 60

APPENDIX A: KINETIC GRAPHS OF UDP-GLC PPASES 61

APPENDIX B: SUPPLEMENTARY INFORMATION ABOUT PROTEIN MODELS 85

REFERENCES 90

VITA 95
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Arrangement of gram negative bacteria’s outer membrane</td>
<td>3</td>
</tr>
<tr>
<td>2. Model of UDP-Glucose synthesis in bacteria</td>
<td>5</td>
</tr>
<tr>
<td>3. Kinetic mechanism of UDP-Glc PPase</td>
<td>8</td>
</tr>
<tr>
<td>4. Crystal Structure of UDP-Glc PPase tetramer (Protein Data Bank code is 2E3D)</td>
<td>9</td>
</tr>
<tr>
<td>5. Sequence alignment of UDP-Glc PPase (GalU) from <em>E. coli</em> with UDP-Glc PPase (GalU) from <em>C. glutamicum</em></td>
<td>16</td>
</tr>
<tr>
<td>6. Model of UDP-Glc PPase (GalU) with UDP-Glc and magnesium ion from <em>C. glutamicum</em> UDP-Glc PPase</td>
<td>17</td>
</tr>
<tr>
<td>7. Sequence alignment of <em>E.coli’s</em> UDP-Glc PPase from galU and galF genes</td>
<td>19</td>
</tr>
<tr>
<td>8. SDS-PAGE of GalU expressed using pET28c vector and purified using 10 ml DEAE Sepharose column</td>
<td>21</td>
</tr>
<tr>
<td>9. SDS-PAGE of GalU expressed using pMCSG9 vector and purified using immobilized metal ion affinity chromatography (IMAC)</td>
<td>22</td>
</tr>
<tr>
<td>10. Specific activity and $K_m$ comparisons of GalU enzymes</td>
<td>25</td>
</tr>
<tr>
<td>11. $S_{0.5} (Mg^{2+})$ comparison of GalU and GalF enzymes</td>
<td>25</td>
</tr>
<tr>
<td>12. Models of GalU with UDP-Glc and magnesium ion from <em>C. glutamicum</em> UDP-Glc PPase (GalU) illustrating six residues mutated in this study</td>
<td>28</td>
</tr>
<tr>
<td>13. Specific activity ($V_{max}$) comparison of UDP-Glc PPase (GalU) constructs</td>
<td>31</td>
</tr>
<tr>
<td>14. $K_m$ comparison of GalU constructs while varying UDP-Glucose concentrations</td>
<td>32</td>
</tr>
<tr>
<td>15. $K_m$ comparison of GalU constructs while varying $PP_i$ concentrations</td>
<td>33</td>
</tr>
</tbody>
</table>
16. SDS-PAGE of GalF expressed using pET24a vector and purified using 10 ml DEAE Sepharose column

17. SDS-PAGE of GalF expressed using pMCSG9 vector and purified using immobilized metal ion affinity chromatography (IMAC)

18. UDP-Glc PPase sequence comparison

19. Homology model of GalF using E. coli GalU as a template and UDP-Glc, magnesium ion from C. glutamicum GalU

20. $V_{\text{max}}$ comparison of UDP-Glc PPases

21. $K_m$ comparison of UDP-Glc PPases using different UDP-Glucose concentrations

22. $K_m$ comparison of UDP-Glc PPases using different $P_i$ concentrations

23. Example of 1.2 % agarose gel electrophoresis showing galF in Strataclone vector

24. Simple model of pMCSG9 vector with the galU gene

25. Spectrophotometric coupled enzyme essay of UDP-Glucose in the pyrophosphorolysis direction
LIST OF ABBREVIATIONS

ADP-glucose ADP-Glc
ADP-glucose Pyrophosphorylase ADP-Glc PPase
Bovine serum albumin BSA
Diethylaminoethyl Sepharose DEAE
Dithiothreitol DTT
Ethylene diaminetetraacetic acid EDTA
Glucose-1-phosphate Glc1P
Isopropyl-β-D-thiogalactoside IPTG
Inorganic pyrophosphate PPi
Luria-Bertani media LB
Maltose-binding protein MBP
N-acetylglucosamine-1-phosphate uridylyltransferase GlmU
Phosphoglucomutase PGM
Protein Data Bank PDB
Sodium dodecyl sulfate SDS
Tabacco etch virus TEV
Uridine 5’-diphosphoglucone UDP-Glc
Uridine 5’-diphosphoglucone pyrophosphorylase UDP-Glc PPase
CHAPTER ONE
INTRODUCTION

*Escherichia coli* were discovered in 1885 by a pediatrician, Theodor Escherich. His bacteriologic studies of infant intestinal flora allowed him to characterize *E. coli* and their relation to the physiology of digestion [34]. The harmless strains of *Escherichia coli* have a habitat in the bowel of humans and animals. Their primary function in the gut is to provide the host with vitamin K and prevent the colonization of pathogenic bacteria. Pathogenic *E. coli* contain strains that produce toxins resulting in food poisoning. They can be found in the soil and contaminated water, and are generally ingested with unclean food or impure drink causing a gastro-intestinal infection. The strains of *E. coli* associated with infectious diseases include Enterotoxigenic *E. coli* (ETEC), Enteroinvasive *E. coli* (EIEC), Enterohemorrhagic *E. coli* (EHEC), Enteropathogenic *E. coli* (EPEC) and Uropathogenic *E. coli* (UPEC) groups. For example, Enteroinvasive *E. coli* (EIEC) is found in water polluted with human feces. The transmission usually occurs through hand-to-hand or hand-to-mouth in unhygienic settings and/or through ingestion of contaminated food. Usually, the symptoms of bacterial infection are associated with mucosal abdominal pain, diarrhea, vomiting, and fever. The immunity against pathogenic *E. coli* depends on the composition of the bacterial cell surface made of lipopolysaccharides [32]. As a result, there have been investigations done on the
pathways and participants involved in the production of lipopolysaccharides in different strains of *E.coli*.

**Lipopolysaccharides**

Lipopolysaccharides (LPS) are glycolipids in Gram-negative bacteria. They are responsible for many surface characteristics of bacterial cells including non-specific host defenses and the resistance to detergents, dyes and antibiotics. LPS’s chemistry and their three dimensional structures prevent many compounds from crossing the outer membrane and gaining access to the periplasm or peptidoglycan. Lipopolysaccharides are amphiphilic because of a hydrophobic moiety called lipid A and the hydrophilic nature of saccharides. Specifically, the backbone of lipid A consists of two \( \beta\)-1,6 linked glucosamine residues which are esterified via the hydroxyl group to fatty acids. The core of lipopolysaccharides contains oligosaccharides which are connected to lipid A via 3-deoxy-D-manno-octulosonate (KDO). Lastly, O-antigens or O-polysaccharides projecting out of the core are carbohydrates made of repeating units of sugars (Figure 1). The composition of these sugars differs between strains of bacteria. These oligosaccharides play a structural role in recognition sites in many biological processes such as antibodies, toxins, and cell recognition [33]. In this thesis, we study the role of UDP-Glucose Pyrophosphorylase (UDP-Glc PPase) involved in bacterial oligosaccharide metabolism. UDP-Glc PPase is encoded by a galU gene in both prokaryotes and eukaryotes. The role of this enzyme is to produce UDP-Glucose (Uridine diphosphate glucose).
Figure 1. Arrangement of gram negative bacteria’s outer membrane. C represents the Core; O symbolizes oligosaccharides.
UDP-Glucose in Prokaryotes

Sugar nucleotides such as UDP-Glucose play a role in constructing oligosaccharides in *E. coli*. UDP-glucose is a starting point for the production of other UDP-sugars such as UDP-galactose (Figure 2). One of the enzymes involved in the synthesis of UDP-glucose is UDP-Glucose Pyrophosphorylase (GalU). UDP-glucose is also a substrate for the synthesis of UDP-glucuronic acid and required for interconversion of galactose and glucose by the Leloir pathway. There has been a model presented for UDP-glucose production in *E. coli* illustrated in Figure 2 [35]. One study revealed that Mesophilic Aeromonas’ UDP-glucose Pyrophosphorylase (GalU) mutants reduced the structure of lipopolysaccharides. Specifically, these mutants were unable to produce O34-antigen and resulted in reduction of pathogenic features [36]. Moreover, *Streptococcus pneumoniae*’s GalU had been reported to be essential for capsule formation and virulence. *Acetobacter xylinum*’s UDP-glucose takes part in cellulose synthesis which is used as a substitute for plants’ cellulose in the production of many commercial products such as paper and cotton textiles [9]. UDP-glucose is also involved in the production of disaccharides such as trehalose in gram positive bacterium, *Corynebacterium glutamicum*, which is used in a wide range of applications in the food industry [10].
Figure 2. Model of UDP-Glucose synthesis in bacteria. Enzymes present are as follows: 1) glucokinase; 2) glucose-6-phosphate-1-dehydrogenase; 3) 6-phosphogluconate dehydrogenase; 4) ribose-phosphate diphosphokinase; 5) orotate phosphoribosyltransferase and orotidine-5′-phosphatedecarboxylase; 6) uridylate kinase; 7) UDP kinase; 8) phosphoglucomutase; 9) UDPglucose-4-epimerase; 10) lipopolysaccharide 3-alpha-galactosyltransferase
UDP-Glucose in Eukaryotes

Uridine diphosphate glucose (UDP-Glucose) is one of the most important sugar nucleotides in higher plants [2] because it is the major glucosyl donor for carbohydrates, serving as the direct precursor for synthesis of sucrose. Sucrose is a major product of photosynthesis in green leaves, accounting for much of CO$_2$ fixed during day time. UDP-Glucose’s function in carbohydrate metabolism has not been explained thoroughly. However, it has been reported that UDP-Glucose may play a critical role in plant growth and development, at least in some species and/or during some specific developmental stages. For example, over-expressed bacterial UDP-Glc PPase in transgenic tobacco revealed an increase in growth and subsequently, increased biomass [6]. Furthermore, UDP-Glucose is a direct precursor for cellulose and callose synthesis occurring at the plasmalemma [3].

In addition, UDP-Glucose is involved in the synthesis of carbohydrate moiety of glycolipids, glycoproteins and proteoglycans, among other functions [4],[5]. In mammals, UDP-Glucose is necessary in the formation of glycogen by providing glucose to the progressive lengthening of the (α1→4) glycosidic chain. Glycogen is primarily deposited in the liver and skeletal muscle. Glycogen synthesis in the liver occurs after food intake as a consequence of the increased glucose level and serves as an energy supply.

Kinetic mechanism of UDP-Glc PPase

UDP-Glucose Pyrophosphorylase (EC 2.7.7.9) (UDP-Glc PPase) is one of the
key enzymes of carbohydrate metabolic pathway widely found in prokaryotes and
eukaryotes. The enzyme catalyzes a reversible formation of uridine diphosphate glucose
(UDP-Glc) and inorganic phosphate from Uridine 3-Phosphate (UTP) and glucose 1-
phosphate (Glc-1-P) (Figure 3). There is a single displacement of pyrophosphate from
UTP by glucose-1-phosphate. The catalytic activity of UDP-Glc PPase appears to be
initiated by binding of UTP or UDP-Glucose prior to the binding of glucose-1-P or PP_i.
The enzyme needs magnesium for its maximal activity [1].

**Three dimensional structure of *E. coli’s* UDP-Glc PPase**

Structural analysis of UDP-Glc PPase of *E. coli* revealed that the protein is a
tetramer and it can be seen as a dimer of dimers (Figure 4). Each subunit contains eight
stranded beta sheets. There are two additional layers of beta strands and ten alpha
helices. In each subunit, Pro-24 has a cis conformation. Val-37 and Asn 151 have
dihedral angles outside of Ramachandran plot and the rest of residues appear to be in
allowed regions of the plot. The enzyme’s subunit contains the elongated globular core
because of present α-helices containing Phe 76-Glu 83, Arg 88-Ser 96 and additional two
helices at C-terminus (Lys-269-Arg 282 and Gly 287-Met 298) which form the “tight
dimer” by subunit-subunit interface [24],[25]. Interestingly, it has been noticed that the
UDP-Glc PPase enzyme is structurally similar to glucose-1-phosphate
thimidylyltransferase [25] and UDP-N-acetylglucosamine pyrophosphorylase [38]. In
addition, there is lower but still significant structural similarity of UDP-Glc PPase to
ADP-Glucose Pyrophosphorylase (ADP-Glc PPase) from bacteria (*Agrobacterium
tumefaciens*), and CDP-Glucose Pyrophosphorylase from *Salmonella typhi* [16].
Figure 3. Kinetic mechanism of UDP-Glc PPase.
Figure 4. Crystal Structure of UDP-Glc PPase tetramer (Protein Data Bank code is 2E3D). Subunits 1 and 4 and Subunits 2 and 3 form a “tight” dimer meaning they interact with each other more than Subunits 1 and 2 and Subunits 3 and 4.
Important residues for the substrate binding

There have been unsuccessful attempts to grow crystals of *E. coli* UDP-Glc PPase with the product or its substrate. Since glucose-1-phosphate thymidylytransferase was crystallized with its substrate and the enzyme is structurally similar to UDP-Glc PPase, the previous study done by Thoden group [25] was able to build a model of the UDP-Glc PPase’s active site. The model had shown that Gln-109 serves to anchor the uracil ring through hydrogen bonding. Furthermore, Gly-17 of UDP-Glc PPase interacts with the 2-hydroxyl group of UDP-Glucose. Lys-202 most likely contacts the β-phosphoryl group of the product while Glu-201 seemed to hydrogen bond with the 2’- and 3’-hydroxyl groups of the glucosyl moiety. The same study noticed a distorted loop containing Lys-84, Arg-85, Val-86 and Lys-87 close to the active site. It was hypothesized that in the presence of the substrate or product in UDP-Glc PPase, the enzyme closes down and the distorted region becomes part of the active site [25]. A three dimensional structure of *Corynebacterium glutamicum*’s UDP-Glc PPase with its product, UDP-Glucose, provided a more defined active site of *E. coli*’s UDP-Glc PPase. Residues involved in anchoring the ligand to the active site include Ala-20, Gly-21, Gly-117, Gly-180 and Ala-214 and side chains of Glu-36, Gln-112, Asp-143, Glu-201, and Lys-202. The product’s uracil ring hydrogen bonds to nitrogens of Ala-20 and Gly-117 and the side chain of Gln-112 (homologous to *E. coli* GalU Gln-109 residue (Figure 5)). The ribose’s 2-hydroxyl group hydrogen bonds to Glu-36, water and Gly-21’s nitrogen. Two magnesium ions are observed in the crystal structure to be coordinated to the UDP-Glucose of *Corynebacterium glutamicum*. The glycosyl group’s 4’hydroxyl interacts with the active
site via nitrogen of Gly-180 and the carbonyl oxygen of Ala-214. An α- and β-
phosphoryl oxygen, three waters, and the side chain of Asp-142 ligate the first
magnesium, whereas the second ion is coordinated by the α-phosphoryl oxygen and five
waters [24].

*In vivo regulation*

There are several levels of regulation that can be envisioned for UDP-Glc PPase. They involve regulation at the gene expression level (eventually having an effect on
UDP-Glc PPase activity/protein content), but also post-translational regulation (e.g.
proteins phosphorylation). Other regulating effects include protein interactions and direct
inhibitory/activating effects of metabolites at the active site of the enzyme. In plants,
abiotic stresses are important factors that affect UDP-Glc PPase’s gene expression. For
example, UDP-Glc PPase from *Arabidopsis* was highly up-regulated by cold treatment at
both mRNA and protein level [19]. Studies on barley’s UDP-Glc PPase have shown
oligomerization as a regulatory process that affects protein function/activity. For instance,
it has been reported that subtle changes in an immediate environment such as buffer or
protein dilution influence oligomerization of UDP-Glc PPase [2]. In yeast, the
localization and function of UDP-Glc PPase were found to be affected by PAS kinase-
dependent serine phosphorylation in the N-terminae domain [20]. The in vivo *O-
glycosylation was reported for mammalian UDP-Glc PPase [21], however, the
significance of this modification is not clear at this moment. Bacterial UDP-Glc PPase
(GalU) interacts with a GalF protein, which modulates its activity in vivo, especially
during stress conditions [22].
**Evolution of UDP-Glc PPase**

All plant UDP-Glc PPases form a single monophyletic group, suggesting a single ancestral gene. Plant UDP-Glc PPases also have relatively high identity (39%-51%) with UDP-Glc PPases from the slime mold [12, 5], animals [13] and yeast [14]. However, eukaryotic UDP-Glc PPases are significantly divergent from those of bacterial origin, with very little or no identity at the amino acid sequence level [15], [2]. This may indicate that the genes of eukaryotic UDP-Glc PPase branched off at the very early stage of evolution, or that they have evolved independently. On the other hand, one study postulated that the Sucrose Synthase (SuSy) family in plants is homologous with bacterial UDP-Glc PPase. [2].

**Tissue and subcellular localization**

UDP-Glc PPase is required for all tissues of all living organisms and it is considered to be mainly localized in the cytosol. However, in some plants such as in rice cells had revealed the presence of UDP-Glc PPase in Golgi to some extent. For example, the fractionation of rice and tobacco cells yielded some UDP-Glc PPase activity in the microsomes (Golgi bodies) [4]. In barley, relatively high UDP-Glc PPase activities were found in a membrane fraction [23].

**GalF protein**

GalF is the protein product of the galF gene that encodes UDP-Glc Pyrophosphorylase. There is no report published about the crystal structure of the GalF protein. However, it has been reported that the GalF protein possibly interacts with GalU to control the production of UDP-Glucose. The galF gene’s product could belong to
family of bacterial UDP-Glc PPases because the gene is highly conserved among microorganisms including Haemophilus influenzae [27] and Actobacter xylinum [28].

Protein-protein interactions in vivo between GalU and GalF proteins using the yeast two hybrid method revealed that the GalF protein possibly interacts physically with GalU and regulates the biochemical and physical properties of UDP-Glc PPase. Specifically, the GalF protein causes reduction in the rate of phosphorylisis and provides a higher thermal resistance for UDP-Glc PPase by increasing in UDP-Glucose production [22]. It has been proposed that the GalF protein of *E. coli* is a non-catalytic subunit of UDP-Glucose Pyrophosphorylase which only physically contacts GalU to regulate the production of UDP-Glucose. However, a different study reported that the galF gene (previously called galE) in *Salmonella typhimurium* might contain some activity. But, the galF allele cannot produce large amount of UDP-Glucose as the galU gene. As a result of that, various mechanisms of interaction between these two proteins were hypothesized. One hypothesis assumed that the galF gene codes for the product, UDP-Glucose, which modifies the polypeptide determined by the galU gene by combining with it. As a result, the GalF protein cannot produce an active enzyme by itself [26].
CHAPTER TWO

GOAL OF THIS PROJECT

Overall goal of the project

Based on the UDP-Glucose Pyrophosphorylase (GalU) model with Corynebacterium glutamicum’s UDP-Glucose and magnesium ion, we hypothesized that Glutamic acid-201, Glutamine-109, Lysine-202, Arginine-21, Lysine-31, and Aspartic acid 265 are residues that play a critical role in the UDP-Glc PPase (GalU) either in catalysis or binding of substrates.

The second part of this project focused on finding whether the GalF enzyme is catalytic or not. Also, the alignment of eukaryotic galU and galF amino acid sequences showed that Threonine-20 and Arginine-21 side chains are missing in GalF. Therefore, we hypothesized that by mutating these two residues in the galF gene, GalF activity will be resurrected (if wild type is inactive) or mutations will increase the enzyme’s activity (if wild type is active).

Part I: Biochemical exploration of UDP-Glc Pyrophosphorylase (GalU) in E.coli

The first part of this project investigates specific amino acids of UDP-Glc PPase (GalU) that play an important role in the catalytic function. GalU’s monomer contains 302 amino acids and has a molecular weight of 32 kDa. In this study, we tested two different expression systems, pET28c and pMCSG9, and purified GalU enzymes. We wanted to see which enzyme would yield a higher degree of purity. Furthermore, to see
whether either purified wild type was significantly affected, we compared their activities as well as apparent binding affinities to magnesium ions. GalU had been mutated previously to study potential residues that might be necessary for catalysis. The mutations already studied were GalU_{A16V}, GalU_{L248P} and GalU_{P14S} which had shown lack of synthesis of *E. coli*’s capsular polysaccharide. Therefore, these three residues were essential for the enzyme’s activity [29].

The *E. coli* UDP-Glc PPase’s crystal structure was reported without the substrate [25]. Only recently the structure of *Corynebacterium glutamicum* was resolved with UDP-Glucose and two magnesium ions in the active site [24]. Therefore, we performed a sequence alignment of *C. glutamicum* UDP-Glc PPase and *E. coli* UDP-Glc PPase (Figure 5) to see which amino acids are conserved and potentially catalytic for the *E. coli* GalU enzyme. In addition, we built the model of UDP-Glc PPase with the substrate and one magnesium ion of *C. glutamicum* UDP-Glc PPase (Figure 6). The model includes one magnesium ion which is found in the same position among other eukaryotic UDP-Glc PPases. To see whether the mutagenesis will change the activity and/or substrate binding of *E. coli*’s UDP-Glc PPase, Glutamic acid- 201, Glutamine-109, Lysine- 202, Arginine-21, Lysine-31 and Aspartic acid 265, were chosen to study in this project.
### Figure 5. Sequence alignment of UDP-Glc PPase (GalU) from *E. coli* with UDP-Glc PPase (GalU) from *C. glutamicum*.

The highlighted residues were studied in this thesis.
Figure 6. Model of UDP-Glc PPase (GalU) with UDP-Glc and magnesium ion from *C. glutamicum* UDP-Glc PPase.
Part II: Investigation of putative UDP-Glucose Pyrophosphorylase (GalF) in *E.coli*

Genes of galU and galF are similar in terms of molecular weight (approximately 32 kD) and amino sequence lengths (galU 302 amino acids and galF 294 amino acids). Furthermore, the comparison of their amino acids sequences showed that they are 56.6% identical (Figure 7). In the previous study done on *E.coli* GalF, it has been proposed that GalU encodes a catalytically active subunit while the GalF enzyme could serve a regulatory role [22]. However, the study done on *S. tryphimurium* GalF revealed that it may produce a small amount of the product, UDP-Glucose. Thus, in this project we attempted to purify *E.coli* GalF to determine GalF’s activity. If GalF is not active or has a very low activity compared to GalU then there is a possibility that GalF’s and GalU’s common ancestor was an active subunit and GalF evolved to serve a regulatory role. We tested two different expression systems, pET24a and pMCSG9, and purified GalF proteins to compare their activities. In addition, we compared their apparent binding affinities to magnesium ions.

In this thesis, we also mutated GalF based on the conserved regions of GalU proteins from bacteria and the homology model of GalF (using GalU crystal structure as a template). We hypothesized that Threonine-20 and Arginine-21 side chains of GalU are possible critical residues that the GalF protein lost in evolution which contributed to it becoming non-catalytic or less-catalytic.
Figure 7. Sequence alignment of *E.coli*’s UDP-Glc PPase from galU and galF genes
CHAPTER THREE
RESULTS AND DISCUSSION

Part I: Expression, purification and characterization of GalU

*E. coli* UDP-Glc PPase (GalU) was purified as indicated in materials and methods. Based on the amino acid sequence of GalU, its molecular weight is around 32 kDa. However, purified GalU on an SDS-PAGE showed it to be around 39 kDa (Figure 8 and 9). A similar weight was observed in the previous study [15].

GalU’s higher molecular weight value could be due to the fact that it did not bind to SDS (sodium dodecyl sulfate) efficiently; therefore the electrophoretic mobility was slower, yielding at higher apparent molecular weight than expected. Hossian group [15] predicted that the inefficient binding of GalU to SDS could be due to a large number of acidic residues present in the protein. However, the GalF protein purified in this study has a similar number of acidic residues (Asp and Glu) as the GalU protein and appeared at the proper molecular weight on the SDS-PAGE (Figure 17). Therefore, there must be another factor contributing to GalU’s higher molecular weight on the SDS-PAGE.
Figure 8. SDS-PAGE of GalU expressed using pET28c vector and purified using 10 ml DEAE Sepharose column. Lanes A3-A11 show the fractions collected from DEAE Sepharose column. Fractions A6 and A7 showed the greatest enzyme activity and they were precipitated with ammonium sulfate at 70 % saturation.
Figure 9. SDS-PAGE of GalU expressed using pMCSG9 vector and purified using immobilized metal ion affinity chromatography (IMAC). A) Lane 1: prestained molecular mass marker; lane 2: crude extract; lane 3 indicates the flow through; lane 4 shows the GalU enzyme (39 kDa) cleaved from the His<sub>6</sub>-MBP tag (46 kDa). B) Lane 1: purified GalU; lane 2: ColorPlus prestained protein marker. The SDS-PAGE was prepared as described in materials and methods.
Kinetic analysis

The SDS-PAGE containing all enzyme fractions of GalU (pET28c) from the column showed the intense protein expression in A6-A9 fractions (Figure 8). However, the highest activity was observed in A6 and A7 fractions. For this reason, these two were combined and used for further kinetic analysis. The comparison of wild type activities between the combined A6-A7 fractions and GalU (pMCSG9) revealed a higher $V_{\text{max}}$ value or specific activity for GalU (pET28c) (Figure 10). The higher activity of fractions may be due to differences in stability of different expression systems or differences in oligomeric states of the enzyme. The binding affinity comparison of the enzyme for substrates (UDP-Glc or PPi) between GalU (pET28c) and GalU (pMCSG9) did not show any significant differences (Figure 10).

Since the comparison of two SDS-PAGEs (Figure 8 and 9) showed that GalU (pMCSG9) yielded purer GalU enzyme (Figure 9) and ensured that no endogenous enzyme was contaminating the sample, further experiments were performed using the pMCSG9 expression vector.

Magnesium curve

*C. glutamicum* UDP-Glc PPase (GalU)’s crystal structure had two magnesium ions present, but we included only one of them in our model [24]. The magnesium ion that we incorporated in our model is in the same position as in other eukaryotic UDP-Glc PPase’s crystal structures. Our model of GalU with the substrate and magnesium ion from *C. glutamicum* UDP-Glc PPase (GalU) (Figure 6) showed magnesium ion to be around 2 Angstroms away from either $P_i$ of UDP-Glc. Our studies indicated that
magnesium ions were necessary for UDP-Glc PPase’s maximum activity. GalU enzymes (pET28c and pMCSG9) showed no activity in absence of magnesium ions. The highest activity for both GalU enzymes was observed upon minimum addition of 2 mM of magnesium ions. Higher concentrations of magnesium did not increase the activity of either UDP-Glc PPases. Furthermore, the enzymes’ apparent binding affinities for magnesium ions were not significantly different from each other (Figure 11).
Table 1. Specific activity and $K_m$ comparisons of GalU enzymes.

<table>
<thead>
<tr>
<th>Vector</th>
<th>Specific Activity (U/mg)</th>
<th>$K_m$ (UDP-Glc), mM</th>
<th>$K_m$ (PP_i), mM</th>
</tr>
</thead>
<tbody>
<tr>
<td>GalU pET28c</td>
<td>118</td>
<td>0.23 ± 0.02</td>
<td>0.1 ± 0.01</td>
</tr>
<tr>
<td>GalU pMCSG9</td>
<td>47</td>
<td>0.19 ± 0.01</td>
<td>0.10 ± 0.01</td>
</tr>
</tbody>
</table>

**Figure 10. Specific activity and $K_m$ comparisons of GalU enzymes.** The specific activity was measured as described in materials and methods.

Table 2. $S_{0.5}$ (Mg$^{2+}$) comparison of GalU and GalF enzymes.

<table>
<thead>
<tr>
<th>Vector</th>
<th>$S_{0.5}$ (Mg$^{2+}$), mM</th>
</tr>
</thead>
<tbody>
<tr>
<td>GalU pET28c</td>
<td>0.56 ± 0.03</td>
</tr>
<tr>
<td>GalU pMCSG9</td>
<td>0.65 ± 0.02</td>
</tr>
<tr>
<td>GalF pMCSG9</td>
<td>1.7 ± 0.3</td>
</tr>
</tbody>
</table>

**Figure 11. $S_{0.5}$ (Mg$^{2+}$) comparison of GalU and GalF enzymes.** The graphs with different concentrations of magnesium and the corresponding binding affinities of UDP-Glc PPases are shown in the appendix section.
Kinetic analysis of GalU mutants

Based on the previously reported list of hypothetical residues involved in the activity of UDP-Glc PPase [24], the structural similarity of UDP-Glc PPase to ADP-Glc PPase and CDP-Glc PPase [16], and our model (Figure 12), six potential catalytic residues were studied. Glutamic acid-201, Glutamine-109, Lysine-202, Arginine-21, Lysine-31, Aspartic acid-265 were all converted to Alanine.

Lysine-202 was chosen because a homologous Lysine-195 was found to be important for the substrate binding in ADP-Glc PPase. Specifically, Lys-195 interacts with β-phosphate of the product, ADP-Glucose, and makes a salt bridge with phosphate of the substrate, glucose-1-phosphate. The study done on Lys-195 in ADP-Glc PPase revealed that this residue affects the binding affinity for glucose-1-phosphate [41].

Another study tested ADP-GlcPPase’s Glutamic acid-194 residue which corresponds to Glutamic acid-201 in E. coli UDP-Glc PPase. The mutation, E194A, revealed that this residue is involved in the glucose-1-phosphate binding site. In addition, this mutation had shown a 24-fold decrease in $V_{\text{max}}$ [16]. E. coli ‘s UDP-Glc PPase mutant D265A was chosen based on the comparison of ADP-Glc PPase and CDP-Glc PPase active sites. It had been shown that CDP-Glc PPase’s active site includes the residue D236 whereas ADP-Glc PPase requires D276 for activity. D276A in ADP-Glc PPase revealed a decrease of 1000-fold in $V_{\text{max}}$ compared to wild type, and a 3.4 fold increase $S_{0.5}$ value for the substrate, ATP [16]. Based on the hypothetical active site of UDP-Glc PPase and the comparison to ADP-Glc PPase and CDP-Glc PPase active sites, we hypothesized that the D265 residue might be critical in activity of UDP-Glc PPase.
GalU’s Arginine-21 is one of the residues conserved among other UDP-Glc PPases. Our GalU enzyme model with UDP-Glc and magnesium ion from *C. glutamicum* UDP-Glc PPase showed the possibility of the Arginine-21 residue’s involvement in catalysis because it is pointing towards PP$_i$ part of UDP-Glucose (Figure 12). In addition, the same model showed Lysine-31 to be potentially involved in activity or substrate binding (Figure 12). Based on the protein/model product in one of the studies, it has been hypothesized that Glutamine-109 might be a residue necessary for the substrate binding [25]. Our model showed Glutamine-109 interacting with the Uracil part of UDP-Glucose, therefore this residue might be involved in UDP-Glucose binding (Figure 12).
Figure 12. Models of GalU with UDP-Glc and magnesium ion from *C. glutamicum* UDP-Glc PPase (GalU) illustrating six residues mutated in this study.
The comparison of specific activities of UDP-Glc PPase constructs revealed that GalU-R21A is a critical catalytic residue. There was a four order of magnitude decrease in activity observed by GalU-R21A compared to wild type (Figure 13). Specifically, GalU-R21A showed 0.003 U/mg activity compared to wild type 47 U/mg. The variation of PPi (Figure 15) and UDP-Glucose concentrations (Figure 14) did not reveal significant affects on the binding affinity of GalU-R21A for substrates.

The GalU-D265A and GalU-K31A constructs showed to play an important role in altering catalysis, yet both residues are not as critical for the maximal activity of UDP-Glc PPase as Arginine-21 (Figure 13). Specifically, GalU-D265A and GalU-K31A showed a decrease in specific activity by two and half orders of magnitude and two orders of magnitude respectively. Further results showed that the GalU-K31A enzyme had no significant affect on the PPi or UDP-Glucose binding affinity (Figure 14 & 15). The binding affinities (K_m) for UDP-Glucose in GalU-K31A and wild type were almost the same revealing 0.18 ± 0.01 mM and 0.19 ± 0.01 mM respectively. Also, Aspartic Acid-265 did not have an affect on the binding affinity for UDP-Glucose (Figure 14). However, the same residue was affecting the binding affinity for PPi (Figure 15) showing 1.8 ± 0.3 mM compared to GalU’s K_m of 0.10 ± 0.01 mM. This result suggests some indirect participation of Aspartic Acid-265 in the substrate binding.

The GalU-E201A, GalU-K202A, GalU-Q109A constructs showed the greatest change in affinity for UDP-Glucose binding compared to wild type and other mutants (Figure 14). GalU-E201A had a K_m of 1.0 ± 0.2 mM, GalU-K202A had 1.1 ± 0.3 mM binding affinity, and GalU-Q109A had 0.83 ± 0.06 mM compared to GalU that had K_m of
Specific activities for all three mutants were not significantly changed. Out of all three mutants, GalU-E201A had the highest effect on enzyme’s activity with a decrease of about two orders of magnitude (Figure 13). GalU’s E201A ($K_m$ of $0.94 \pm 0.16$) and GalU-D265A ($K_m$ of $1.8 \pm 0.3$) mutants showed to affect the binding of pyrophosphate compared to wild type ($K_m$ of $0.10 \pm 0.01$), which indicates that these two residues play a role in the binding of substrates.

**Conclusions**

The catalytic characterization of GalU revealed a decrease in UDP-Glucose Pyrophosphorylase activity of all mutants. Out of all mutated residues, Arginine-21 was the most critical catalytic amino acid since the mutation showed the highest drop in activity compared to wild type. Glutamine-109 was important for the binding with Uracil, demonstrating its specificity for this type of enzyme. Lysine-202 and Glutamic Acid-201 had recognition for the sugar, UDP-glucose, but they were not important for the enzyme’s specificity. Aspartic Acid-201 indirectly participated in the pyrophosphate binding.

Magnesium ions were critical for UDP-Glc PPase’s activity. Furthermore, approximately 2 mM was the minimal concentration of magnesium necessary for the enzyme’s maximal activity.
Figure 13. Specific activity ($V_{\text{max}}$) comparison of UDP-Glc PPase (GalU) constructs. $V_{\text{max}}$ for GalU was 47 ± 3.0 U/mg; GalU$_{D265A}$ 0.29 ± 0.02 U/mg; GalU$_{K31A}$ 0.10 ± 0.02 U/mg; GalU$_{R21A}$ 0.003 ± 0.0009; GalU$_{E201A}$ 0.43 ± 0.07 U/mg; GalU$_{K202A}$ 6.9 ± 0.2 U/mg; GalU$_{Q109A}$ 19 ± 2 U/mg. The specific activities were measured as described in materials and methods.
Figure 14. \( K_m \) comparison of GalU constructs while varying UDP-Glucose concentrations. The exact \( K_m \) (mM) for each construct were as follows: GalU 0.19 ± 0.01; GalU-D265A 0.09 ± 0.05; GalU-K31A 0.18 ± 0.02; GalU-R21A 0.11 ± 0.04; GalU-E201A 1.0 ± 0.2; GalU-K202A 1.1 ± 0.3; GalU-Q109A 0.83 ± 0.06. The graphs with different concentrations of UDP-Glucose and the corresponding specific activities of UDP-Glc PPase constructs are shown in the appendix A section.
Figure 15. $K_m$ comparison of GalU constructs while varying PP$_i$ concentrations. $K_m$ (mM) for GalU was 0.10 ± 0.01; 2: GalU$_{D265A}$ 1.8 ± 0.3; GalU$_{K31A}$ 0.21 ± 0.01; GalU$_{R21A}$ 0.09 ± 0.01; GalU$_{E201A}$ 0.94 ± 0.16; GalU$_{K202A}$ 0.11 ± 0.02; GalU$_{Q109A}$ 0.08 ± 0.01. The graphs with different concentrations of PPI and the corresponding specific activities of the UDP-Glc PPase constructs are shown in the appendix A section.
Part II: Expression, purification and characterization of GalF

The expression of GalF proteins (expressed in both pET24a and pMCSG9) from the galF gene appeared to be around 30 kDa just as predicted by amino acid sequence (Figure16 &17).

Kinetic analysis

The measurement of specific activities of GalF proteins (pET24a and pMCSG9) indicated that GalF is active. Specifically, GalF (pET24a) revealed 0.02 U/mg (using 0.1 mM PP$_i$) activity and GalF (pMCSG9) had 0.12 U/mg (0.1 mM PP$_i$) and 0.24 ± 0.06 U/mg (0.5 mM PP$_i$). The difference in activity may be due to differences in stability using different expression systems. Despite the fact that there was a significant difference in activity between GalF proteins, these results suggest that the ancestry of GalF was a catalytic subunit that became reduced in activity as it evolved.

Further experiments were performed using pMCSG9 vector.

Magnesium curve

Our studies indicated that magnesium ions were necessary for the maximum activity of GalF. The comparison of apparent binding affinities between GalF and GalU enzyme showed that the GalF enzyme has lower binding affinity for magnesium ions (Figure 11). As a result, GalF needs higher concentration of magnesium for its maximal activity. Approximately 5 mM of magnesium was necessary to detect the highest activity of GalF.
Figure 16. SDS-PAGE of GalF expressed using pET24a vector and purified using 10 ml DEAE Sepharose column. Lane 1 shows the prestained protein marker; Lane 2 & 3 shows GalF enzyme around 30 kDa.
Figure 17. SDS-PAGE of GalF expressed using pMCSG9 vector and purified using immobilized metal ion affinity chromatography (IMAC). Lane 1 represents the prestained protein marker; Lane 2 is a crude extract; Lane 3: the flow through; Lane 4 indicates GalF enzyme cleaved from His6-MBP tag; Lane 5 shows pure UDP-Glc PPase (GalF).
Kinetic comparisons of GalU, GalF and GalF mutant

Based on analysis of the active site of GalU reported in the previous study [24] and conserved residues among bacterial GalU proteins that are missing in GalF (Figure 18), the hypothetical catalytic residues were chosen to be studied in this project. Since the three dimensional structure of the GalF protein has not been reported yet, we modeled this enzyme based on the GalU’s crystal structure (Figure 19). Also, our model includes the substrate and magnesium ion using \textit{C. glutamicum} UDP-Glc PPase (GalU) as a template. We hypothesized that mutations of Histidine-20 into Threonine and Methionine-21 into Arginine, could potentially resurrect the activity of GalF (if wild type is inactive) or increase the activity of GalF (if wild type is active).

The specific activities and binding affinities for substrates (PP\textsubscript{i} and UDP-Glucose) between GalF and GalF- M15T, H16R were not changed significantly (Figure 20, 21 & 22). These results revealed that the two mutations on GalF had no effect on the enzyme’s activity and it did not affect the binding affinity for either substrate (UDP-Glc or PPi). Threonine and Arginine may not be the only residues that cause low activity of GalF.

The specific activities of GalU and GalF (or GalF- M15T, H16R) showed a 100-fold decrease for GalF. Specifically, GalF’s \( V_{\text{max}} \) was 0.24 \pm 0.06 U/mg while GalU had the specific activity of 47 \pm 3.0 U/mg (Figure 20). Further kinetic analysis revealed that the GalF protein affected binding affinities for UDP-Glucose and PP\textsubscript{i} compared to the GalU enzyme (Figure 21, 22). For example, the binding affinity of PP\textsubscript{i} for GalU was 0.10 \pm 0.07 mM compared to GalF that revealed 0.04 \pm 0.07 mM (Figure 22).
Figure 18. UDP-Glc PPase sequence comparison. The highlighted are the conserved residues of GalU proteins that are missing in GalF enzyme.
Figure 19. Homology model of GalF using *E. coli* GalU as a template and UDP-Glc, magnesium ion from *C. glutamicum* GalU. The model shows two residues that were mutated in this study. Specifically, Methionine was mutated to Threonine whereas Histidine was replaced by Arginine.
Conclusions

It has been reported that the galF gene does not encode an active UDP-Glc PPase, however, our study revealed that GalF is an active enzyme. Yet, the protein’s specific activity was 100-fold lower compared to GalU. This result suggests that the ancestry of GalF is a catalytic subunit and that it became reduced in activity as it evolved.

The saturation curve revealed that GalF needs magnesium ions in order to be active. The comparison of binding affinities between GalF and GalU showed that GalF requires a higher concentration of magnesium in order for its maximal activity to be detected. The two mutations performed on the galF gene did not affect GalF’s activity and they did not have an effect on the binding affinity of UDP-Glc or PPi.
Figure 20. \(V_{\text{max}}\) comparison of UDP-Glc PPases. \(V_{\text{max}}\) for GalU was 47 ± 3.0 U/mg; GalF 0.24 ± 0.06 U/mg; GalF-M15T, H16R 0.21 ± 0.09 U/mg. The specific activities were measured as described in materials and methods.
Figure 21. \( K_m \) comparison of UDP-Glc PPases using different UDP-Glucose concentrations. The specific \( K_m \) of GalU was 0.19 ± 0.01, GalF was 0.78 ± 0.11 and GalF\(_{M15T,H16R}\) was 0.86 ± 0.08. The graphs with different concentrations of UDP-Glucose and the corresponding specific activities of UDP-Glc PPases are shown in the appendix A.
Figure 22. $K_m$ comparison of UDP-Glc PPases using different PP\textsubscript{i} concentrations. GalU’s $K_m$ was 0.10 ± 0.01; Km of GalF was 0.04 ± 0.01; GalF\textsubscript{M15H16R} 0.06 ± 0.01. The graphs with different concentrations of PPI and the corresponding specific activities of UDP-Glc PPases are shown in the appendix A.
CHAPTER FOUR
MATERIALS AND METHODS

Structure prediction and homology modeling

The homology modeling was performed with the program Modeller8v2. GalU and GalF models included the product UDP-Glucose and magnesium ion form C. glutamicum UDP-Glc PPase (PDB ID: 2PA4). The crystal structure of GalU (PDB ID: 2E3D) was used as a template for the GalF model. The alignment was performed using Bioedit (http://www.mbio.ncsu.edu/BioEdit/bioedit.html). The secondary structure prediction was performed with the PSI-PRED method [45] available on the PSI-PRED server (bioinfo.cs.ucl.ac.uk/psipred/) [46]. The pdb files were viewed on DeepView/Swiss-PDBViewer 3.7. Verification of models was performed using the program called Verify3D (nihserver.mbi.ucla.edu/Verify_3D/) [43]. VMD, Swiss-PdbViewer and POV-ray programs (http://www.povray.org/) were used to picture the models. The sequence alignments and Verify3D graphs can be found in the appendix B.

Cloning of galU and galF genes from E. coli genome

The cloning process was performed by Misty Khun. XL1-Blue cells were grown in 5 ml Luria-Bertani (LB) Media overnight at 37 °C. After 12-15 hours, the cells were purified using the Promega’s Wizard Genomic DNA Purification Kit to obtain E. coli’s
genomic DNA. Then, the DNA was digested with Hind III and Xba I restriction sites in order to obtain the pieces of the E. coli’s genome. These restrictions sites were chosen because they did not cut through galU and galF genes. Next, TAG PCR was performed containing the mixture of 1 µl of digested genomic DNA, 25 µl GO-TAG Green Master Mix 2x, 2.5 µl of each primer and 19 µl of nanopure-di-water. The mixture was placed in the thermocycler (initial denaturation of 30 seconds at 98˚C, 30 cycles of 98 °C for 30 seconds, 50 °C for 20 seconds, 72 °C for 1 minute and 5 minute extension at 72 °C) and then run on the agarose gel electrophoresis 80 V for 60 minutes. The correct size band for the galU gene was 906 bp while for the galF gene was 894 bp. The galU and galF bands were extracted, purified, ligated into a Strataclone vector. The transformation into E .coli XL1-Blue cells (Stratagene) and further steps were exactly followed as described under Site-directed mutagenesis section.

**Primers for cloning of galU gene from E.coli genome**

Forward 5’-ATG GCT GCC ATT AAT ACG AAA GTC AAA AAA GCC GTT ATC CCC GTT GCG GG-3’ reverse 5’- TTA CTT CTT AAT GCC CAT CTC TTC TTC AAG CCA GGC TTT AAA TTC CGT GC-3’

**Primers for cloning of galF gene from E.coli genome**

Forward: 5-ATG ACG AAT TTA AAA GCA GTT ATT CCT GTA GCG GGT CTC GGG ATG CAT AT-3’ and reverse: 5’-TTA TTC GCT TAA CAG CTT CTC AAT ACC TTT ACG GAA CTT CGC CCC TTC TT-3’
Site-directed mutagenesis

The sequences of galU and galF E. coli were obtained from the GenBank (http://www.ncbi.nlm.nih.gov/) for visualization. The primers for galU and galF genes containing mutations were designed using the program called BioEdit (http://www.mbio.ncsu.edu/BioEdit/bioedit.html). Site-directed mutagenesis was performed using Sambrook’s laboratory manual with some modifications [37]. Phusion PCR was performed to obtain the PCR product (mutated gene). The PCR mixture contained 32.5 µl of water, 10 µl of 5x Phusion HF buffer, 1 µl of 10 mM dNTPs, 2.5 µl of primers, 0.5 µl of DNA template (3mg/µl) and 0.5 µl of Phusion DNA polymerase. The PCR mixtures were run in the thermocycler (initial denaturation of 30 seconds at 98°C, 30 cycles of 98°C for 30 seconds, 50°C for 20 seconds, 72°C for 1 minute and 5 minute extension at 72°C) then the PCR products were run on agarose gel electrophoresis. The correct bands were extracted, purified and another phusion PCR was performed using these bands as the template DNAs with the appropriate flanking primers. Next, the final PCR product was ligated into a Strataclone vector and transformed into E. coli XL1-Blue cells (Stratagene) using the StaraClone Blunt PCR Cloning kit. Transformants were plated on carbenicillin plates in the presence of X-gal and they were left in the incubator at 37°C overnight for blue-white colony screening [37]. The presence of white colony represents a successful ligation of a vector into a desired gene. The mechanism is based on genetic engineering of the lac operon in the E. coli laboratory with a complemented subunit from the cloning vector. The vector encodes α subunit of LacZ protein with an internal multiple cloning site and the chromosome of
*E. coli* (a host cell) encodes β-galactosidase. The DNA of interest is inserted within the lacZ gene which disrupts the production of β-galactosidase and appears as a white colony. White, bacterial colonies were picked, purified and colony PCR was performed [37]. Agarose gel electrophoresis analysis allowed us to verify whether the vector was inserted into a gene (Figure 23). The correct band was extracted, purified and confirmed by DNA sequencing at University of Chicago Cancer Research Center, DNA Sequencing Facility. The galU and galF genes were sub-cloned into expression vectors pet28c and pet24a (Novagen) respectively and both of them were sub-cloned into pMCSG9 vector [39].
Figure 23. Example of 1.2 % agarose gel electrophoresis showing galF in Strataclone vector. Lane 1: 1 kb ladder; Lane 2: galF band. The correct size band for galF should be around 894 bp which is approximately what is observed on the gel above. The galF band was extracted and purified as described in materials and methods.
Primers used for site-directed mutagenesis

**galU**

Site-directed mutagenesis was performed by PCR overlap extension using the procedure listed under Site Directed Mutagenesis.  galU-R21A forward (5’-TTA GGA ACC GCG ATG TTG CCG-3’) and reverse (5’-CGG CAA CAT CGC GGT TCC TAA-3’), galU-K31A forward (5’-GCC ATC CCG GCG GAG ATG CTG-3’) and reverse (5’ CAG CAT CTC CGC CGG GAT GGC-3’), galU-Q109A forward (5’- CAA GTT CGT CGC GGT CTG GCG-3’), galU-D265A forward (5’-AAG AGC CAT GCG TGC GGT AAT-3’) and reverse (5’-ATT ACC GCA CGC ATG GCT GGT-3’), galU-E201A forward (5’- GGT GTG GTA GCG AAA CCG AAA-3’) and reverse (5’- GGTCGT GTA GCG AAA CCG AAA-3’) and reverse (5’ – TTT CGG TTT CGC TAC CAC ACC-3’), galU-K202A forward (5’-GTG GTA GAA GCG CCG AAA GCG-3’) and reverse (5’- CGC TTT CGG CGC TTC TAC CAC-3’).

**galF**

galF M15T H16R  forward direction (5’- GGT CTC GGG ACC CGT ATG TTG CCT-3’) and reverse (5’-AGG CAA CAT ACG GGT CCC GAG ACC-3’).

**Flanking primers for galU and galF genes**

Forward 5’-TAA TAC GAC TCA CTA TAG GG- 3’ and reverse 5’-GCT ATG TAT TGC TCA GCG G-3’

**pMCSG9 vector**

The pMCSG9 vector [39] was used to purify the proteins of interest.  It contains a his6 tag bound to a maltose binding protein (MBP).  The his6 tag allows the protein of
interest to be purified using the immobilized metal ion affinity chromatography (IMAC) [44]. The maltose binding protein (MBP) improves solubility and there is also a TEV-protease site in order to cleave the protein of interest from the his$_6$-MBP tag (Figure 24). The advantage of using this vector for the protein expression and purification is that it selectively purifies the protein containing his$_6$-MBP tag and therefore eliminates the possibility to have interference with endogenous activity.
Figure 24. Simple model of pMCSG9 vector with the galU gene.
Subcloning genes into pMCSG9 vectors

Genes of galU and galF were amplified by phusion PCR. The primers used for GalU wild type and mutants had the same construct. galU forward 5'- TAC TTC CAA TCC AAT GCC GCA GCA ATG GCT GCC ATT AAT ACG AAA GTC AAA-3’ and reverse 5’- TTA TCCACT TCC AAT GTT ACT TCT TAA TGC CCA TCT CT-3’. GalF and galF-M15T H16R forward 5’-TAC TTC CAA TCC AAT GCC GCA GCA ATG ACG AAT TTA AAA GCA GTT ATT CTT-3’ and reverse 5’- TTA TCC ACT TCC AAT GTT ATT CGC TTA ACA GCT TCT CA-3’. The purified PCR product from agarose gel electrophoresis was cloned into pMCSG9 vector [39], [40]. The gene was cloned into this vector using ligation independent cloning (LIC). The starting mixture contained 1 µl 5mM dCTP, 2 µl T4 polymerase buffer 10x, 1 µl 3mM DTT, 15 µl of PCR gene product, 0.5 µl of water and the reaction was initiated by 0.5 µl of T4 DNA polymerase. The mixture was placed into a thermocycler at 25 °C for 25 minutes and at 75 °C for 20 minutes. Next, the mixture was allowed to cool for 1 hour and 2 µl of pMCSG9 vector was added into a mixture and incubated at room temperature for 10 minutes. 2 µl of the plasmid mix were added into 100 µl of BL21 (DE3) Magic cells and incubated on ice for 10 minutes. The soc medium was added to the mix and left shaking for 1 hour at 37°C. Afterwards, the cells were plated on kanamycin/ampicillin LB agar plates and incubated overnight at 37°C. Next day, some colonies were picked and checked by using the colony PCR procedure.
Colony PCR procedure

One colony from the LB agar plate was chosen by touching the colony with a toothpick and scratched inside the 0.2 ml PCR individual tubes. 20 µl of nuclease free water, 2.5 µl of the appropriate primers for the gene, and 25 µl of 2x Green Master Mix were added into a 0.2 ml PCR tube with a colony in it. Next, the PCR tube with its contents was placed in a thermocycler using a program that starts with denaturation at 95°C for 5 minutes, then 25 cycles of 1 minute at 95 °C, 30 s at 50 °C and 1 minute at 72 °C with the final extension at 72 °C for 5 minutes. Then the mixture was run on 1.2 % agarose gel electrophoresis. The gel was visualized with GIMP2 program and if the band of the gene was correct then its DNA was sequenced.

DNA sequence visualization

The Bioedit program was used to see DNA sequence of the gene. The program was downloaded from http://www.mbio.ncsu.edu/BioEdit/bioedit.html website.

pET vector

The pET vector is used for the expression of recombinant protein in E. coli. The pET vector contains bacteriophage T7 transcription and translation signals, and the expression is induced by providing a source of T7 RNA polymerase in the host cell. The pET vector is good at expressing wild type protein; however it is not the best choice for the protein containing mutation(s). The mutated protein expressed using pET vector maybe mixed with endogenous wild type. This contamination may result in inaccurate mutated enzyme’s activity. Thus, the vector containing tag such as pMCSG9 is a better option for expressing the mutated protein.
**Subcloning genes into pET vectors**

Genes of galU and galF in Strataclone vectors were digested with pET vectors (Novagen). The digestion tube for galU contained buffer #1 (Novagen), NcoI and SacI restriction enzymes obtained from NEW ENGLAND Biolabs Inc, the pET28c vector (http://www.genomex.com/vector_maps/pET28_map.pdf) and galU (Strataclone). The digestion tube for the galF gene contained buffer #4 (Novagen), NdeI and SacI restriction enzymes, the pET24a vector (http://www.genomex.com/vector_maps/pet24a-d.pdf) and galF (Strataclone). The digestion contained 10:1 ratio of the gene to vector. The digestion tubes were placed in the water bath at 37 °C overnight. Next, PCR clean-up was performed using the protocol in Wizard SV Gel and PCR Clean up system (Promega). Next, the quick ligation was performed by pipeting out 10 µl of the PCR clean up system product into a new 1.5 ml microcentrifuge tube. 10 µl of 2X quick ligase buffer (New England Bio Labs Inc) and 1 µl of the quick T4 DNA ligase (New England Bio Labs Inc) were added into 10 µl PCR product. The mixture was briefly centrifuged and incubated on ice for 5 minutes. Then, 2 µl of the mixture was used to transform it into NEB Turbo cells (NEW England Biolabs Inc) and plated on kanamycin plates. Lastly, the colony PCR was performed to check whether the vector was inserted into a gene. The flanking primers were used for the colony PCR for both pETgalU and pETgalF.

**Expression and purification of GalU and GalF from pET vectors**

pETgalU and pETgalF were transformed into *E. coli* BL21 (DE3) cells (Novagen). The single colony were grown at 37 °C in 1 liter of Luria-Bertani medium with kanamycin (50 µg/ml) and shaked until an optical density of 0.6-0.7 was reached at
600 nm. The culture was induced with 1M of IPTG and allowed to grow for an additional 16 hours at 20°C before harvesting and centrifugating (2.5 rpm for 5 minutes). Next, the supernatant was discarded and the pallet was sonicated in buffer A (50m M HEPES, pH 8.0, 5 mM MgCl₂, 0.1 mM EDTA, 10% sucrose). The solution was centrifuged again at 12 rpm for 20 minutes and stored at -80 °C. The purification was performed at 0 to 4°C. The crude extract was applied onto 10 ml DEAE- Sepharose column using Akta FPLC UPC-900 (Amersham Biosciences) equilibrated in buffer A and the fractions were eluted with linear gradient of NaCl. The fractions were collected and the enzyme was monitored by pyrophosphorolysis activity.

Expression of GalU and GalF from pMCSG9 vectors

Single colony of BL21 (DE3) “Magic” cells containing pMCSG9 vector were grown overnight at 37 °C in 5 ml LB medium containing sterile 5 µl of 1000x ampicillin and 5 µl of 1000x kanamycin. On the next day, the 5 ml starting culture was poured onto a 1ml LB medium and 1ml of 1000x ampicillin plus 1ml of 1000x kanamycin were added. The culture was grown at 25°C shaking at 250 rpm until the OD₆₀₀ reached 0.6-0.8. Next, the culture was put on ice and cooled to 16°C. The protein was expressed by an addition of 0.5 mM of IPTG and left shaking at 16°C. After around 16 hours, the cells were harvested by centrifugation 2.5 rpm for 5 minutes in SS34 rotor. The cells were re-suspended in the appropriate volume (5 ml of buffer per 1 gram of cells) of buffer C containing 200 mM NaCl, 20 mM Tris pH 8.0, 10% glycerol and sonicated on ice using the Misonix sonicator 3000 for 4 minutes total (30 seconds on, 45 seconds off). Lastly,
the cells were centrifuged in the SS34 rotor for 20 minutes with speed of 12 rpm after which the supernatant (crude extract) was retained and poured onto a Nickel column.

**Purification of GalU and GalF from pMCSG9 vectors**

The crude extract obtained from last centrifugation was loaded onto a Ni-NTA Aragose resin (Qiagen). The Ni-NTA column with the crude extract was slowly shaken up and down for 1 hr so that the maximum amount of protein attaches to the resin. Then the column was centrifuged at 2000 rpm for 2 min and the flow through was collected which was then put in -80 °C. Next, three column volumes (15 ml) of buffer C were added to the Nickel column in order to wash the unbound protein. The column was then inverted slowly using Barnstead/Thermolyne Labquake Shaker (Rotisserie) for 5 minutes and the wash was collected after centrifugation (2000 rpm for 2 min). The protein was eluted after the addition of combined buffer C and 300 mM of Buffer E (1 M imidazole, 200 mM NaCl, 20 mM Tris pH 8, 10% glycerol). Again, the column was shaken for 5 minutes and centrifuged to obtain the protein (once the protein was eluted, the column needed to be washed with 3 column volumes of mixture of buffer C and 500 mM buffer E and 2 volume columns washes with buffer C). The collected fraction containing the protein was treated with 5 mM EDTA and approximately 4 mg of TEV (tobacco Etch Virus) Protease stirring overnight at 4 °C. Next, the protein was precipitated with ammonium sulfate of 60 % saturation and centrifuged to remove any imidazole. The protein was re-suspended in buffer C (200 mM NaCl, 20 mM Tris pH 8.0, 10% glycerol) and loaded onto a Ni-NTA resin. The column was shaken up and down for 1 hr,
centrifuged and the elute (protein) was collected. Afterwards, the column was washed with buffer C and higher concentration of buffer E as described above.

**Measuring protein concentration**

The concentration of purified protein was measured using a NanoDrop ND-1000 Spectrophotometer with an extinction coefficient of 1 AU/mg$^{-1}$ ml$^{-1}$. The procedure was exactly performed as provided by the manufacturer. The module displays the UV spectrum and measures the protein’s absorbance at 280 nm to calculate the concentration in mg/ml automatically.

**SDS-PAGE preparation**

Electrophoresis was performed using 1.5 mm cassette. The separation gel was made from 3.125 ml of 40% of acrylamide, 4.7 ml of separating gel buffer (1M Tris-HCl, pH 8.8), 125 µl of 10% SDS, 125 µl of catalyst (0.1 g Ammonium Persulfate in 960 µl of water), 2 ml of 50% sucrose, 5 µl of TEMED and 2.55 ml of water. Next, n-butanol was poured onto the top of separating gel to eliminate any bubbles formed. The gel was left for a while until it hardened. Then the gel was flushed with de-ionized-nanopure water to remove the excess of n-butanol. The stocking gel was poured onto the separating gel. The mixture of stocking gel included 0.4 ml of 40% acrylamide, 1.56 ml stocking gel buffer (0.5 M Tris-HCl pH 6.8), 62.2 µl of 10% SDS, 0.25 ml of catalyst (0.1 g Ammonium Persulfate in 960 µl of water), 2.5 µl of TEMED and 4 ml of water. The sample containing 5-10 µl of crude extract or fraction after purification was mixed with 2-Mercaptoethanol, 2x sample buffer (stocking buffer, glycerol, 10% SDS, water, 0.2 % bromophenol blue and 2-Mercaptoethanol) and run on the thermocycler for 10 minutes.
Afterwards, the samples were loaded onto the gel and run for 60 minutes, 200 Volts. The 1X Tris-HEPES-SDS (Invitrogen) buffer was used for electrophoresis. The New England Biolabs prestained protein marker broad range was used to estimate protein’s molecular weight. The gel was stained with Pierce (Rockford, IL, USA) Gel Code Blue Stain Reagent.

**Enzyme assay: Pyrophosphorolysis direction**

The UDP- Glucose Pyrophosphorylase activity was measured in the reverse direction (Pyrophosphorolysis), from UDP-glucose and PP\textsubscript{i} to Glucose-1-P and UTP. The reaction mixture (300 µl) contained 80 mM HEPES at 8.0 pH, 7 mM of MgCl\textsubscript{2}, 2 mM of DTT, 0.5 mM of UDP- glucose, 0.6 mM of NAD\textsuperscript{+}, 10 mM of NaF, 0.01 mM of G1,6BP, 2 U/ml of PGM (Phosphoglucomutase preparation from Rabbit muscle), 2 U/ml of G6PDH (glucose-6-phosphate dehydrogenase form *Leuconostoc mesenter*), 0.2 mg/ml of bovine serum albumin, de-ionized-nanopure water and 10 µl of enzyme. The reaction was initiated by the addition of 0.5 mM of NaPP\textsubscript{i} and placed into the spectrophotometer (BioTek EL 808) (Figure 24).
Figure 25. Spectrophotometric coupled enzyme essay of UDP-Glucose in the pyrophosphorolysis direction. The essay monitors enzyme activity.
Calculation of kinetic constants

The kinetic assays were performed at optimal conditions for all reaction mixtures. The data was plotted for different concentrations of PP$_i$ (mM) on x-axis versus the activity (nmol/min) on y-axis. Then, we used Origin 7.5 program which contains the modified Hill equation for a non-linear curve fit, $V = V_o + (V_{\text{max}} - V_o) \times \frac{s^n}{(k^n + s^n)}$ [42]. $V$ is the velocity while $V_{\text{max}}$ is the maximum velocity at saturation. S symbolizes the substrate and $k$ is the amount of substrate concentration needed to obtain 50% maximum velocity. We varied $V_{\text{max}}$ and $k$ parameters to get the best possible fit. The same process was used for the data with variation of UDP-Glucose (mM).
APPENDIX A:

KINETIC GRAPHS OF UDP-GLC PPASES
GalU (pMCSG9): varying [Mg$^{2+}$]

Conditions: 50 ng of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

Equation: \( y = \text{START} + (\text{END} - \text{START}) \times x^n / (k^n + x^n) \)

\[ R^2 = 0.98966 \]

\begin{align*}
\text{START} & \quad 0 & 0 \\
\text{END} & \quad 1.8819 & 0.02842 \\
k & \quad 0.65024 & 0.02462 \\
n & \quad 2.59794 & 0.23154
\end{align*}
GalU (pMCSG9): varying [UDP-Glc]

Conditions: 25 ng of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

![Graph showing the relationship between UDP-glucose concentration and activity](image)

Equation: \( y = \text{START} + (\text{END} - \text{START}) \times x^n / (k^n + x^n) \)

\( R^2 = 0.99271 \)

- START = 0
- END = 1.24149
- k = 0.19139
- n = 1

Values are given in millimeters per minute (mmols/min).
GalU (pMCSG9): varying [PP_i]

Conditions: 25 ng of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

Equation: \[ y = \text{START} + (\text{END} - \text{START}) \times x^n / (k^n + x^n) \]

\[ R^2 = 0.99345 \]

<table>
<thead>
<tr>
<th>START</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>END</td>
<td>1.04998</td>
<td>0.02703</td>
</tr>
<tr>
<td>k</td>
<td>0.10099</td>
<td>0.00718</td>
</tr>
<tr>
<td>n</td>
<td>1.8</td>
<td>0</td>
</tr>
</tbody>
</table>
**GalU (pET28c): varying [Mg\(^{2+}\)]**

Conditions: 30 ng of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

![Graph showing activity vs. MgCl\(_2\) concentration](image)

**Equation:**

\[
y = \text{START} + (\text{END} - \text{START}) \times x^n / (k^n + x^n)
\]

**R\(^2\) = 0.97692**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value 1</th>
<th>Value 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>START</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>END</td>
<td>2.54289</td>
<td>0.05167</td>
</tr>
<tr>
<td>k</td>
<td>0.55861</td>
<td>0.03137</td>
</tr>
<tr>
<td>n</td>
<td>2.43611</td>
<td>0.30328</td>
</tr>
</tbody>
</table>
GalU (pET28c): varying [UDP-Glc]

Conditions: 7.4 ng of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

Equation: \( y = \text{START} + \frac{(\text{END} - \text{START}) \cdot x^n}{(k^n + x^n)} \)

\( R^2 = 0.99738 \)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>START</td>
<td>0</td>
</tr>
<tr>
<td>END</td>
<td>2.29102</td>
</tr>
<tr>
<td>k</td>
<td>0.22878</td>
</tr>
<tr>
<td>n</td>
<td>1</td>
</tr>
</tbody>
</table>
**GalU (pET28c): varying [PP\textsubscript{i}]**

Conditions: 7.4 ng of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

![Graph showing activity versus PPi concentration](image)

Equation: \( y = \text{START} + (\text{END} - \text{START}) \times x^n \div (k^n + x^n) \)

\( R^2 = 0.99548 \)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value 1</th>
<th>Value 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>START</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>END</td>
<td>1.57189</td>
<td>0.03174</td>
</tr>
<tr>
<td>( k )</td>
<td>0.10048</td>
<td>0.00619</td>
</tr>
<tr>
<td>( n )</td>
<td>1.4</td>
<td>0</td>
</tr>
</tbody>
</table>
GalU D265A: varying [UDP-Glc]

Conditions: 9.4 µg of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

Equation: $y = \text{START} + (\text{END} - \text{START}) \times x^n / (k^n + x^n)$

$R^2 = 0.90523$

<table>
<thead>
<tr>
<th>START</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>END</td>
<td>1.67659</td>
<td>0.1209</td>
</tr>
<tr>
<td>k</td>
<td>0.09162</td>
<td>0.04779</td>
</tr>
<tr>
<td>n</td>
<td>0.4</td>
<td>0</td>
</tr>
</tbody>
</table>
GalU D265A: varying [PPi]

Conditions: 47 µg of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

![Graph showing the relationship between PPI concentration and activity](image)

Equation: \( y = \text{START} + \frac{(\text{END} - \text{START}) \times x^n}{k^n + x^n} \)

\[ R^2 = 0.99755 \]

<table>
<thead>
<tr>
<th>START</th>
<th>END</th>
<th>k</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>19.42007</td>
<td>1.82909</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>2.44702</td>
<td>0.32501</td>
<td>0</td>
</tr>
</tbody>
</table>
GalU K31A: varying [UDP-Glc]

Conditions: 9.4 µg of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

Equation: $y = \text{START} + (\text{END} - \text{START}) \times x^n \div (k^n + x^n)$

$R^2 = 0.97347$

<table>
<thead>
<tr>
<th>START</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>END</td>
<td>0.8706</td>
<td>0.03074</td>
</tr>
<tr>
<td>$k$</td>
<td>0.17859</td>
<td>0.02361</td>
</tr>
<tr>
<td>$n$</td>
<td>$1$</td>
<td>$0$</td>
</tr>
</tbody>
</table>
GalU K31A: varying [PPi]

Conditions: 28 µg of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

Equation: \( y = \text{START} + \frac{(\text{END} - \text{START}) \times x^n}{(k^n + x^n)} \)

\( R^2 = 0.99659 \)

<table>
<thead>
<tr>
<th>START</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>END</td>
<td>3.65974</td>
<td>0.0912</td>
</tr>
<tr>
<td>k</td>
<td>0.20532</td>
<td>0.01227</td>
</tr>
<tr>
<td>n</td>
<td>1.6</td>
<td>0</td>
</tr>
</tbody>
</table>
GalU R21A: varying [UDP-Glc]

Conditions: 12 µg of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods

Equation: \( y = \text{START} + (\text{END} - \text{START}) \times x^n / (k^n + x^n) \)

\( R^2 = 0.82161 \)

\( \begin{aligned}
\text{START} & = 0 & 0 \\
\text{END} & = 0.03205 & 0.00235 \\
k & = 0.11306 & 0.03823 \\
n & = 1.3 & 0
\end{aligned} \)
GalU R21A: varying [PPᵢ]

Conditions: 24 µg of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

![Graph showing enzymatic activity against PPᵢ concentration](image)

Equation: \( y = \text{START} + \frac{(\text{END} - \text{START}) \times x^n}{(k^n + x^n)} \)

\( R^2 = 0.97648 \)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>START</td>
<td>0</td>
</tr>
<tr>
<td>END</td>
<td>0.07584</td>
</tr>
<tr>
<td>k</td>
<td>0.089</td>
</tr>
<tr>
<td>n</td>
<td>1.0</td>
</tr>
</tbody>
</table>
GalU K202A: varying [UDP-Glc]

Conditions: 39 ng of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

Equation: \( y = \text{START} + (\text{END} - \text{START}) \times x^n / (k^n + x^n) \)

\[ R^2 = 0.97609 \]

START 0 0
END 0.52169 0.06211
k 1.12752 0.27321
n 1 0
GalU K202A: varying \([\text{PP}_i]\) 

Conditions: 39 ng of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

Equation: \(y = \text{START} + (\text{END} - \text{START}) \times x^n / (k^n + x^n)\)

\[
\begin{align*}
\text{START} & \quad 0 \quad 0 \\
\text{END} & \quad 0.33377 \quad 0.01384 \\
k & \quad 0.11105 \quad 0.01561 \\
n & \quad 1 \quad 0
\end{align*}
\]

\(R^2 = 0.98586\)
GalU Q109A: varying [UDP-Glc]

Conditions: 12 ng of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

\[
y = \text{START} + \frac{(\text{END} - \text{START}) \times x^n}{(k^n + x^n)}
\]

\[R^2 = 0.99734\]

\[
\begin{array}{c|c|c}
\text{START} & 0 & 0 \\
\text{END} & 0.42744 & 0.01363 \\
k & 0.82958 & 0.06004 \\
n & 1 & 0 \\
\end{array}
\]
GalU Q109A: varying [PP_i]

Conditions: 12 ng of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

Equation: \( y = \text{START} + (\text{END} - \text{START}) \times x^n / (k^n + x^n) \)

\( R^2 = 0.99208 \)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>START</td>
<td>0</td>
</tr>
<tr>
<td>END</td>
<td>0.41432</td>
</tr>
<tr>
<td>k</td>
<td>0.08016</td>
</tr>
<tr>
<td>n</td>
<td>1</td>
</tr>
</tbody>
</table>

\( x \) represents [PP_i] and \( y \) represents activity (nmol/min).
GalU E201A: varying [UDP-Glc]

Conditions: 0.6 µg of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

Equation: \( y = \text{START} + (\text{END} - \text{START}) \times x^n / (k^n + x^n) \)

\( R^2 = 0.98497 \)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>START</td>
<td>0</td>
</tr>
<tr>
<td>END</td>
<td>0.72592</td>
</tr>
<tr>
<td>k</td>
<td>1.01564</td>
</tr>
<tr>
<td>n</td>
<td>1</td>
</tr>
</tbody>
</table>
GalU E201A: varying [PP\textsubscript{i}]

Conditions: 0.6 µg of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

Equation: $y = \text{START} + (\text{END} - \text{START}) \times x^n / (k^n + x^n)$

$R^2 = 0.9947$

| $\text{START}$ | 0 | 0 |
| $\text{END}$ | 1.12592 | 0.1132 |
| $k$ | 0.94202 | 0.16175 |
| $n$ | 1 | 0 |
GalF (pMCSG9): varying \([\text{Mg}^{2+}]\)

Conditions: 4.4 \(\mu\)g of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

![Graph showing enzyme activity vs. MgCl2 concentration]

Equation: \(y = \text{START} + (\text{END} - \text{START}) \times x^n / (k^n + x^n)\)

\(R^2 = 0.97854\)

\[
\begin{array}{c|c|c}
\text{START} & 0 & 0 \\
\text{END} & 1.23893 & 0.12313 \\
k & 1.72281 & 0.30192 \\
n & 1.3944 & 0.2195 \\
\end{array}
\]
GalF (pMCSG9): varying [UDP-Glc]

Conditions: 4.4 µg of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

Equation: \( y = \text{START} + (\text{END} - \text{START}) \times x^n / (k^n + x^n) \)

\( R^2 = 0.99058 \)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value 1</th>
<th>Value 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>START</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>END</td>
<td>2.90082</td>
<td>0.17342</td>
</tr>
<tr>
<td>k</td>
<td>0.77778</td>
<td>0.10799</td>
</tr>
<tr>
<td>n</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
GalF (pMCSG9): varying [PPi]

Conditions: 4.4 µg of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

Equation: \[ y = \text{START} + \frac{(\text{END} - \text{START}) \times x^n}{(k^n + x^n)} \]

\[ R^2 = 0.95832 \]

START 0 0
END 1.08417 0.04363
k 0.04131 0.00705
n 1.1 0
GalF M15T H16R (pMCSG9): varying [UDP-Glc]

Conditions: 3 µg of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

Equation: \( y = \text{START} + (\text{END} - \text{START}) \times x^n / (k^n + x^n) \)

\( R^2 = 0.99508 \)

START 0 0 0
END 1.7037 0.07415
k 0.85466 0.08357
n 1 0
GalF M15T H16R (pMCSG9): varying [PPi]

Conditions: 3 µg of the protein, Temperature: 30 °C, other conditions are the same as described under the enzyme assay section in materials and methods.

Equation: \( y = \text{START} + (\text{END} - \text{START}) \times x^n / (k^n + x^n) \)

\( R^2 = 0.98422 \)

\begin{array}{c|cc}
\text{START} & 0 & 0 \\
\text{END} & 0.71189 & 0.02271 \\
\text{k} & 0.06112 & 0.00775 \\
n & 1 & 0 \\
\end{array}
APPENDIX B:

SUPPLEMENTARY INFORMATION ABOUT PROTEIN MODELS
Sequence alignment for the GalF model

1) galug\[newline\]structureX:galug: 138: A: 264: A:undefined:undefined:-1.00:-1.00; 2) galuu \[newline\]structureX:galuu: 5: A: 298: A:undefined:undefined:-1.00:-1.00; 3) 2pa4a \[newline\]structureX:2pa4a: 12: A: 326: A:undefined:undefined:-1.00:-1.00; 4) GALF \[newline\]sequence:GALF: : : :: : : -1.00:-1.001
Verify3D graph for the GalF model
Sequence alignment for the GalU model

Verify3D graph for the GalU model

- Crystal structure of C. glutamicum GalU
- Crystal structure of E. coli GalU
- GalU model
REFERENCES


VITA

Agnieszka Maria Orlof was born in Tuchów, Poland on August 14, 1985. Agnieszka graduated from Resurrection High School at Chicago in 2004. Next, she obtained a Bachelor’s degree in Biochemistry at Loyola University of Chicago in 2008. Agnieszka pursued a thesis-based Master Degree in Chemistry.
THESIS APPROVAL SHEET

The thesis submitted by Agnieszka M. Orlof has been read and approved by the following committee:

Miguel Ballicora, Ph.D.
Assistant Professor Biochemistry
Loyola University Chicago

Richard Holz, Ph.D., Chair
Professor of Chemistry
Loyola University Chicago

Ken Olsen, Ph.D.
Professor of Biochemistry, Biophysics
Loyola University Chicago

The final copies have been examined by the director of the thesis and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the thesis is now given final approval by the committee with reference to content and form.

The thesis is therefore accepted in partial fulfillment of the requirements for the degree of Master of Science.

__________________      ____________________________________
Date            Director’s Signature