STRUCTURAL AND FUNCTIONAL ASPECTS OF PYRANOSE-FURANOSE MUTASES

By

Jijin Raj Ayanath Kuttiyatveetil

A Thesis Submitted to the

College of Graduate Studies and Research

University of Saskatchewan

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

Department of Chemistry, University of Saskatchewan, Saskatoon, SK, Canada

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my work was completed. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in this thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Chemistry

University of Saskatchewan

Saskatoon, Saskatchewan (S7N 5C9)

Abstract

Pyranose-Furanose mutases are enzymes that catalyze the isomerization of six-membered pyranose and five-membered furanose forms of a nucleotide-based sugar. In this research, the substrate binding site of three different mutases were investigated; UDP-galactopyranose mutase (UGM), GDP-*altro*-heptopyranose mutase (GaHM) and UDP-arabinopyranose mutase (UAM). Both UGM and UAM use a UDP-based sugar as the substrate but require different cofactors, flavin adenine dinucleotide (FAD) and Mn^{2+} respectively, to function. UGM and GaHM use the same cofactor (FAD), but the latter prefers to work with a GDP-based sugar. In this thesis, studies have been conducted on these three mutases using a variety of tools, such as X-ray crystallography, protein modeling, site-directed mutagenesis and kinetic assays, to understand how these enzymes bind their respective substrates.

Among these three mutases, UGM is the best-studied enzyme and is a validated drug target in *Mycobacteria*. Despite this, the structural role of some active site residues in substrate binding is not clearly understood. *Deinococcus radiodurans* UGM (*Dr*UGM) mutants of active site residues Trp184, Arg364, His88, and Asn372 were prepared and evaluated using kinetic and docking studies. The results suggested that these residues are vital to the positioning of UDP-galactopyranose under FAD in a productive conformation, for maximum enzyme efficiency. Inhibition studies, using the inhibitor MS-208, were performed on *Mycobacterium tuberculosis* UGM (*Mt*UGM). Kinetic assays indicated that MS-208 is a mixed-type inhibitor of *Mt*UGM.

In this study, the crystal structures of *Campylobacter jejuni* GaHM (*Cj*GaHM) with a substrate mimic GDP-mannose were solved, allowing for a comparison of GaHM and UGM substrate binding sites. The results highlighted the alterations undergone by *Cj*GaHM to accommodate a GDP-based substrate in the active site.

A preliminary model of UAM was built based on the protein sequence of *Oryza sativa* UAM1 (*Os*UAM1) using the protein structure modeling servers I-TASSER and GalaxyWEB. The models suggested that, unlike the catalytic role played by the FAD cofactor in UGM and GaHM, the role of the Mn^{2+} cofactor in UAM could be to aid the stabilization of the negative charge of the substrate diphosphate. Furthermore, experiments with mutants of *Os*UAM1 have helped identify residues that may bind the metal cofactor.

Acknowledgments

Firstly, I would like to thank my supervisor, Dr. David Sanders, for giving me the opportunity to carry out research in his laboratory and providing an excellent and flexible environment to learn and work. I thank him for the support and guidance, during my Ph.D. program at the University of Saskatchewan. I feel greatly indebted to him for his patience and ever-optimistic attitude, which helped me immensely through tough times.

Secondly, I would like to thank Dr. David Palmer, for his suggestions and timely help in my project work. I have gained greatly from his knowledge and guidance. I am thankful to all other members of my advisory committee, Dr. Dale Ward and Dr. Miroslaw Cygler, whose inputs have enabled my project progress. I am very grateful to the Department of Chemistry for the support during my stay and study.

I have had the good fortune of working with and learning from a number of people during my program. I would like to thank Dr. Karin van Straaten, for her suggestions in conducting research and help in getting familiar with the data collection at CLS and data processing. I want to extend my gratitude to other past members of the Sanders lab, Dr. Sean Dalrymple, Dr. Karunan Partha Sarathy and Dr. Josiah Obiero for teaching me various laboratory techniques and sharing scientific ideas. I want to record my appreciation to all collaborators, Dr. Julien Cotelesage, Department of Geology, University of Saskatchewan, Dr. Todd Lowary, University of Alberta and Dr. Mario Pinto, Simon Frazer University for their contributions towards my research.

I would like to thank all current members of the Sanders lab for creating a great working environment. A special word of thanks to Naheda Sahtout, for her help with my thesis. I also want to take this opportunity to thank all past and present members of the Palmer lab, whom I have had the pleasure of knowing at a personal and professional level. Last but not the least, I thank all my

friends, who have supported me in different circumstances during my stay in Saskatoon. I also thank my family members for their constant encouragement without which I would not be where I am today.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

6d-D-*altro*-Hep*f*	6-deoxy-D-*altro*-heptofurnaose
A. caulinodans	*Azorhizobium caulinodans*
AEBSF	4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride
A. fumigatus	*Aspergillus fumigatus*
*Af*UGM	*A. fumigatus* UGM
A. thaliana	*Arabidopsis thaliana*
*At*RGP1	*A. thaliana* RGP1
*At*RGP2	*A. thaliana* RGP2
*At*RGP3	*A. thaliana* RGP3
ASU	Asymmetric unit
CLS	Canadian Light Source
CPS	Capsular polysaccharide
CD	Circular dichroism
C. jejuni	*Campylobacter jejuni*
*Cj*GAHM	*C. jejuni* GaHM
*Cj*UNGM	*C. jejuni* UNGM
D-Ara*f*	D-arabinofuranose
D-Gal*f*	D-galactofuranose
D-Fuc*F*	D-fucofuaranose
D. radiodurans	*Deinococcus radiodurans*
*Dr*UGM	*D. radiodurans* UGM
dTDP-D-Fuc*f*	Deoxy thymidine-di phosphate-D-fucofuranose
dTDP-D-Fuc*p*	Thymidine-di phosphate-D-fucopyranose
E. coli	*Escherichia coli*
*Ec*UGM	*E. coli* UGM
EDTA	Ethylene diamine tetra acetic acid
EXAFS	Extended X-Ray absorption fine structure
FAD	Flavin adenine dinucleotide
FAD_{ox}	Oxidized FAD

FAD_{red}	Reduced FAD
GaHM	GDP-*altro*heptopyranose mutase
Gal*p*	Galactopyranose
GDP	Guanosine diphosphate
GDP-*altro*-hep*p*	GDP-6-deoxy-D-altro-heptopyranose
G1P	Galactose-1-phosphate
GOLD	Genetic optimization for ligand docking
GMP	Guanosine monophosphate
HPLC	High-performance liquid chromatography
IPTG	Isopropyl-ß-thiogalactoside
I-TASSER	Iterative threading assembly refinement
K. pneumoniae	*Klebsiella pneumonia*
*Kp*UGM	*K. pneumoniae* UGM
L-Ara*f*	L-arabinofuranose
L. major	*Leishmania major*
LB	Luria Bertani
LC-MS	Liquid chromatography-mass spectrometry
LOMETS	Local meta-threading-server
LPS	Lipopolysaccharide layer
M. smegmatis	*Mycobacterium smegmatis*
M. tuberculosis	*Mycobacterium tuberculosis*
MBP	Maltose binding protein
MPD	2-Methyl-2,4-pentanediol
*Mt*UGM	*M. tuberculosis* UGM
NAD	Nicotinamide adenine dinucleotide
NADP	Nicotinamide adenine dinucleotide phosphate
NCS	Non-crystallographic symmetry
P. aerugiosa	*Pseudomonas aeruginosa*
PCR	Polymerase chain reaction
PDB	Protein data bank

PEG	Polyethylene glycol
PIX	Positional isotope exchange
RGI	Rhamnogalacturonan I
RGII	Rhamnogalacturonan II
RGP	Reversibly glycosylated polypeptide
RMSD	Root mean square deviation
T. cruzi	*Trypansoma cruzi*
*Tc*UGM	*T. cruzi* UGM
S. dysenteriae	*Shigella dysenteriae*
S. typhimurium	*Salmonella typhimurium*
SCOP	Structural classification of proteins
SDM	Site-directed mutagenesis
SAD	Single anomalous dispersion
SDS	Sodium dodecyl sulfate
SDS PAGE	SDS polyacrylamide gel electrophoresis
STD-NMR	Saturation transfer difference-nuclear magnetic resonance
TEV	Tobacco etch virus
UAM	UDP-arabinopyranose mutase
UDP	Uridine diphosphate
UDP-Ara*f*	UDP-arabinofuranose
UDP-Ara*p*	UDP-arabinopyranose
UDP-D-Gal*p*NAc	UDP-N-acetyl galactopyranose
UDP-D-Gal*f*NAc	UDP-N-acetyl galactofuranose
UDP-Gal*p*	UDP-galactopyranose
UDP-GlcUA	UDP-glucuronic acid
UDP-Glu*p*	UDP-glucopyranose
UGM	UDP-galactopyranose mutase
UMP	Uridine monophosphate
UNGM	UDP-N-acetyl galactopyranose mutase
UTP	Uridine triphospahate

Chapter 1: Introduction

1.1 Importance of furanoses in cell walls of organisms

Bacteria, fungi, protozoa, and plants have an outer cell wall that protects the vulnerable contents of the cell. Repeating units of carbohydrates, commonly referred to as polysaccharides, form one of the major components of the cell wall. Understanding the various polysaccharides that constitute the cell wall of these organisms, explaining the formation and necessity of these complex polysaccharides and decoding the function of the enzymes that aid in their biosynthesis represents an attractive avenue for research. The monosaccharide sugar units that make up these polysaccharides can exist either as six-membered pyranose sugars or five-membered furanose sugars. Although the pyranose form is more abundant of the two, the furanose form is no less important.

Furanoses have been identified in the cell wall of bacteria, fungi, protozoa, and plants.[1] Gram-negative bacteria have an outer membrane called the lipopolysaccharide (LPS) layer that lies outside the thin peptidoglycan layer. The LPS layer has O-antigens or O-polysaccharides consisting of sugars in the furanose ring form. For example, D-galactofuranose (D-Gal*f*), arguably the most abundant furanose-sugar, has been identified in O-antigens of bacteria such as *Escherichia coli* (*E. coli*), *Klebsiella pneumonia* (*K. pneumoniae*), *Shigella dysenteriae* (*S. dysenteriae*) and *Salmonella typhimurium* (*S. typhimurium*).[2-4] Other furanose sugars are also incorporated in bacterial cell walls. D-arabinofuranose (D-Ara*f*), even though not as common as D-Gal*f*, is found in the cell wall glycoconjugates of *Actinomycetes*, *Pseudomonas aeruginosa* (*P. aerugiosa*), *Azorhizobium caulinodans* (*A. caulinodans*) and in the mycolyl-arabinogalactan layer of *Mycobacterium sp*.[5-7] Furanose residues have also been identified in plant cell wall

glycoconjugates.[8] L-arabinofuranose (L-Ara*f*) residues are found in arabinan and arabinogalactan, which are the major structural components of Rhamnogalacturonan I (RGI) in plant pectin.

Furanoses help to maintain cell wall rigidity. The galactan of *Mycobacterium tuberculosis* (*M. tuberculosis*), believed to contribute to cell wall structure and impermeability, is a good example.[9] Furanoses are crucial to cell growth and survival of bacterial, fungal and protozoan pathogens. Deletion of the *glfA* gene, corresponding to an enzyme essential for galactofuranose biosynthesis in the fungus *Aspergillus fumigatus* (*A. fumigatus*), rendered the cells with a thinner galactofuran layer, less virulent and more susceptible to antifungal agents.

These furanoses are not found in mammalian glycoconjugates.[1] Therefore, understanding the production and cell wall incorporation of furanoses in organisms pathogenic to humans and mammals has its merits. Weakening the cell wall of these harmful organisms has long been considered important in controlling their growth and the spread of diseases. Drug molecules focused on targeting the furanose biosynthetic pathway represents an option for the development of potential chemotherapeutics.[10] Although a number of different furanoses have been identified and studied in the cell walls of these organisms, only those relevant to this thesis work are introduced here.

1.2 D-Galactofuranose

D-Gal*f* residues are vital components that are essential for growth and virulence of a mycobacteria.[9] The galactofuran of the *Mycobacterial* cell wall, which features numerous long chains of D-Gal*f* residues, is a case in point (Figure 1-1).[9] These residues were identified in the LPS O-antigen polysaccharides of bacteria such as *E. coli, K. pneumonia,* and *S. dysenteriae* species. They have also been identified in the cell wall architecture of fungi, such as the *Aspergillus*

species and protozoa, such as *Leishmania major* (*L. major*).[11,12] D-Gal*f* residues have been identified in unique glycoconjugates found in pathogenic bacteria and other organisms pathogenic to humans and mammals.

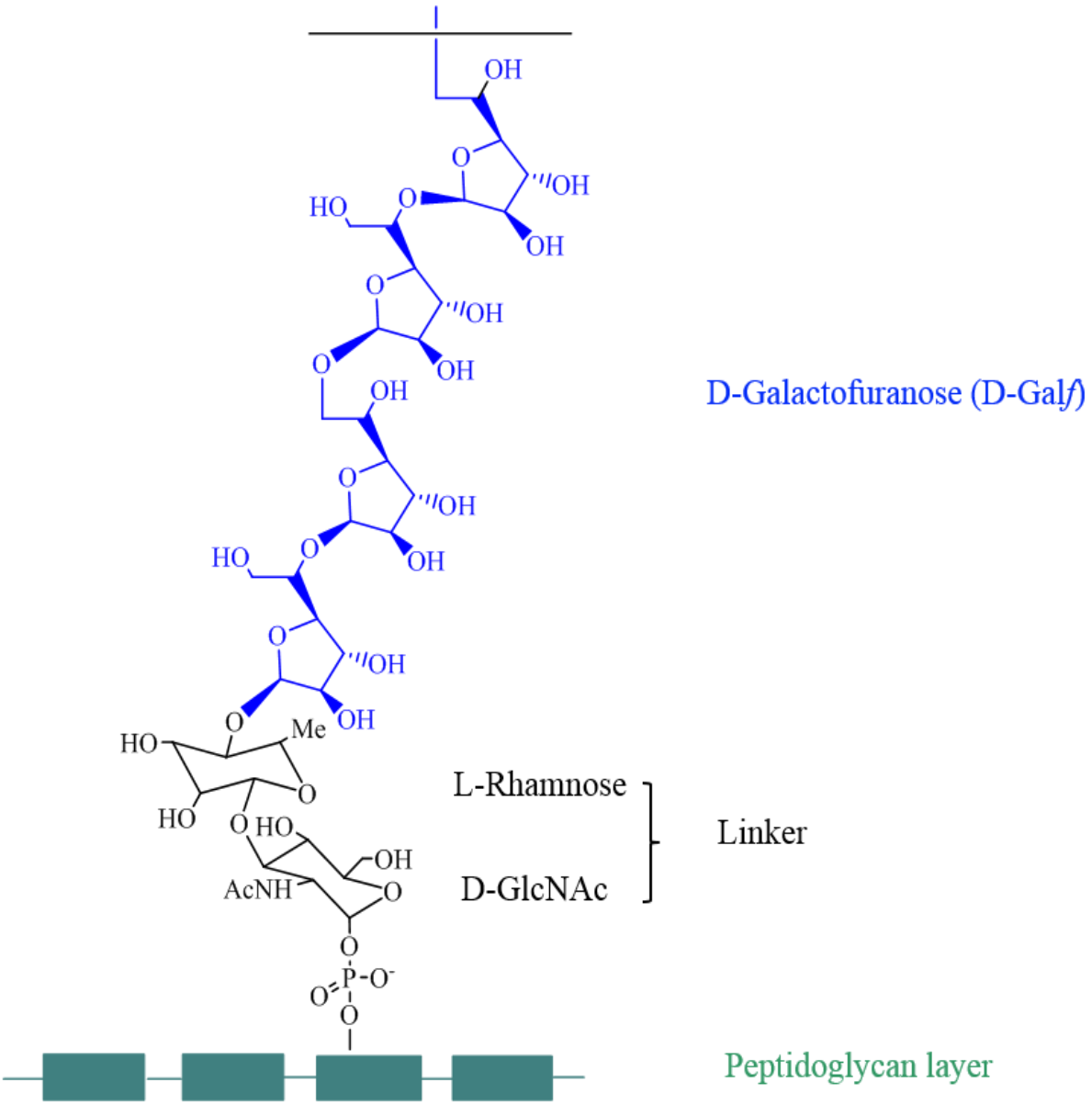


Figure 1-1: A representation of D-Galactofuranose identified in *Mycobacterium* species. D-Gal*f* residues are shown in blue.

1.3 Biosynthesis of D-Galactofuranose

The proposed D-Gal*f* biosynthetic pathway is shown in Figure 1-2.[13] Galactopyranose (Gal*p*) is phosphorylated to galactose-1-phosphate (G1P) by galactokinase, which adds a phosphate group to the C1 of galactose. Then, uridine monophosphate (UMP) is added onto G1P by G1P uridylyltransferase to form uridine diphosphate (UDP) galactopyranose (UDP-Gal*p).*[14] Additionally, UDP-Gal*p* can also be synthesized from UDP-glucopyranose (UDP-Glu*p*) by UDP-galactose-4-epimerase.[15]

The enzymes mentioned above constitute the Leloir pathway. UDP-Gal*f*, the precursor of D-Gal*f* residues, is formed by the mutase enzyme, UDP-galactopyranose mutase (UGM), which interconverts between the six-membered Gal*p* and the five-membered Gal*f*, as shown in Figure 1-2.[2] Finally, galactosyltransferases transfer D-Gal*f* to the cell wall.

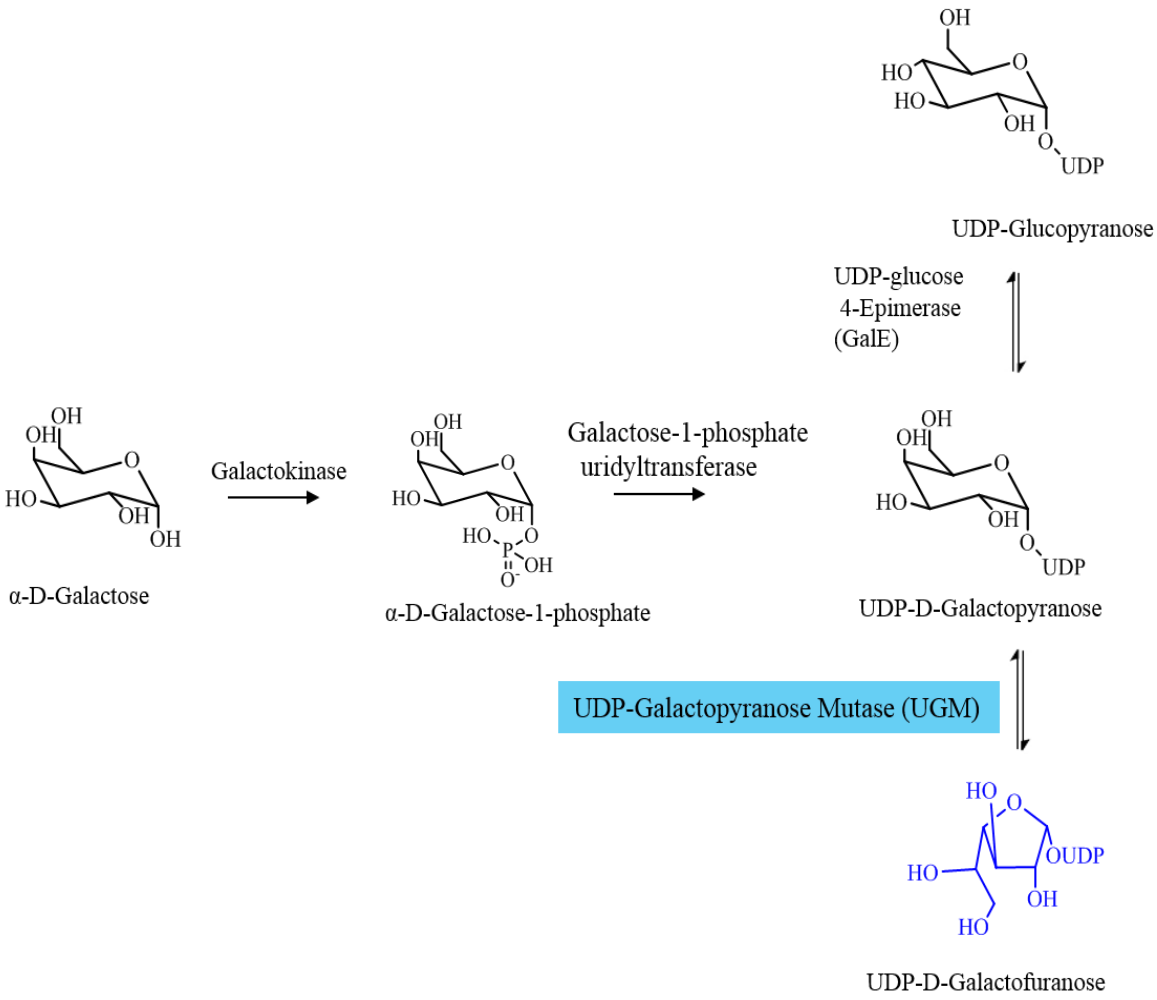


Figure 1-2: Proposed biosynthesis of D-Galactofuranose through the Leloir pathway. UGM and the UDP-Gal*f* are highlighted.

1.4 L-Arabinofuranose

L-arabinofuranose (L-Ara*f*) is an abundant sugar found in arabinogalactan glycoproteins and polysaccharides such as arabinoxylan in hemicelluloses and RG I and II in pectin of plant cell wall.[8] Arabinan domains, one of the structural components of RGI, have linear chains of (1, 5)-linked α-L-Ara*f* residues that are substituted by L-Ara*f* at O-2 or O-3, as shown in Figure 1-3.

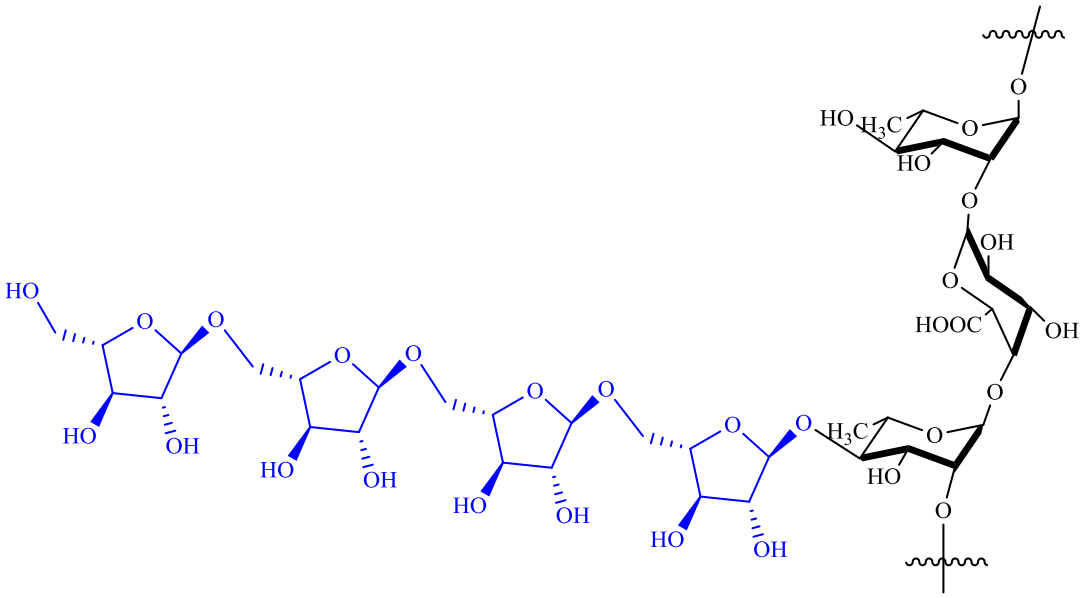


Figure 1-3: A representation of L-arabinofuranose residues in Arabinan of plant Rhamnogalacturonan I.
L-Ara*f* residues have been highlighted in blue.

1.4.1 Biosynthesis of L-Arabinofuranose in plants

UDP-arabinopyranose (UDP-Ara*p*) can be formed either by the action of UDP-xylopyranose 4-epimerase, which catalyzes the epimerization of UDP-xylopyranose to UDP-Ara*p*, or by the action of UDP-Ara*p*-1-phosphate pyrophosphorylase, which transfers UMP to L-Arap-1-phosphate from uridine triphosphate (UTP) (Figure 1-4).[16,17] UDP-Ara*p* is then converted to UDP-arabinofuranose (UDP-Ara*f*) by the enzyme UDP-arabinopyranose mutase (UAM) (as

shown in Figure 1-4).[18] Arabinosyltransferases then act on UDP-Ara*f* and transfer L-Ara*f* to the plant cell wall.

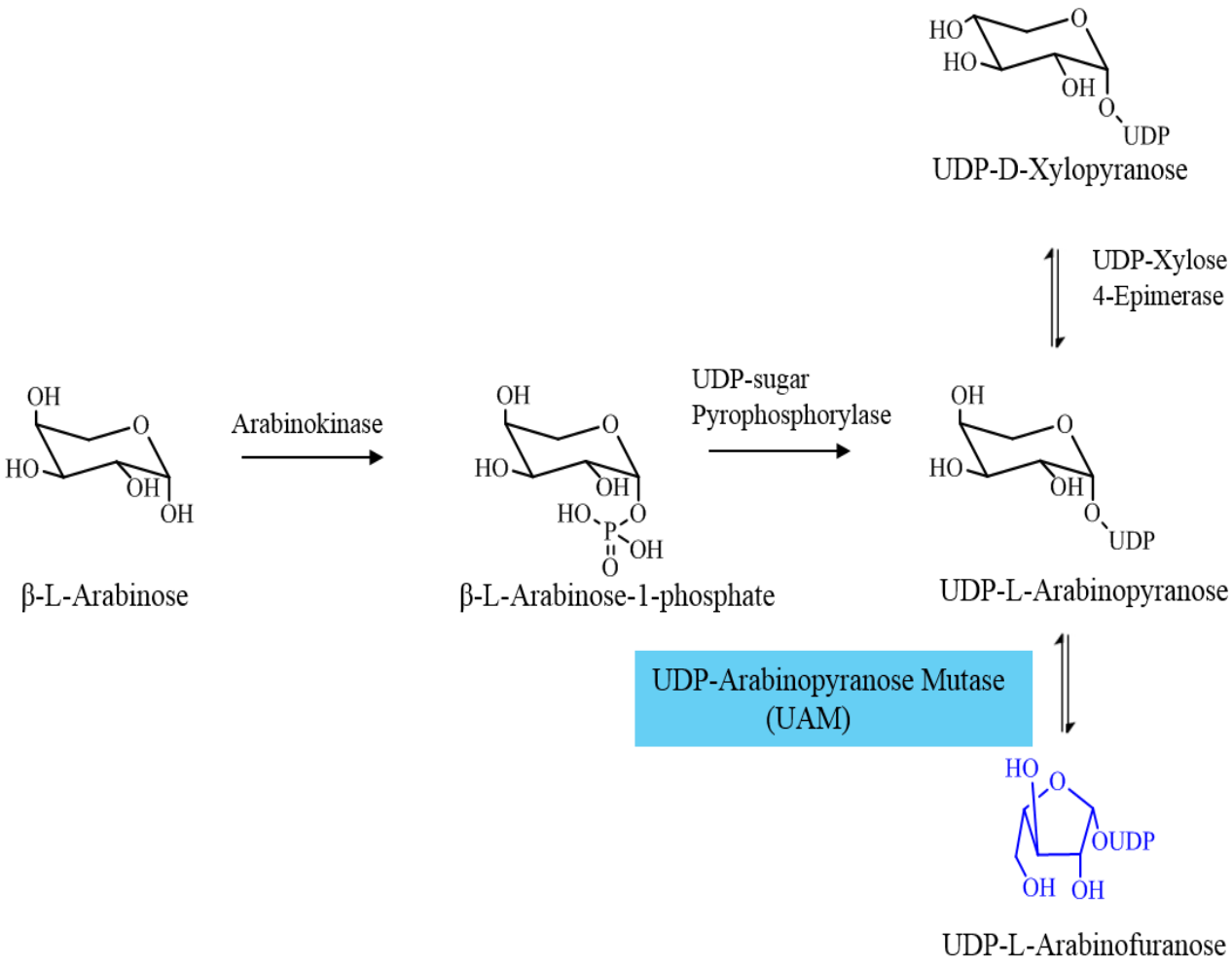


Figure 1-4: Biosynthesis of L-Arabinofuranose in plants.
UAM and UDP-Ara*f* have been highlighted.

1.5 6d-D-*altro*-heptofurnaose

6-deoxy-D-*altro*-heptofurnose (6d-D-*altro*-Hep*f*) is one of three furanose sugars found in the capsular polysaccharide (CPS) and is essential for the virulence of *Campylobacter jejuni* (*C. jejuni*), a bacterium that causes gastroenteritis in humans.[19] The trisaccharide repeat unit of a strain of *C. jejuni*, HS:41 serotype, is comprised of 6d-D-*altro*-Hep*f* residues, L-Ara*f* residues and one

among either 6-deoxy-L-altrofuranose or D-fucofuranose (D-Fuc*F*) occurring as the third furanose sugar (shown in Figure 1-5).[20] Trisaccharides containing 6d-D-*altro*-Hep*f* occur more frequently (75%) than those with D-Fuc*F* (25%). No information is available with regards to the biosynthetic pathway that produces this trisaccharide repeat unit.

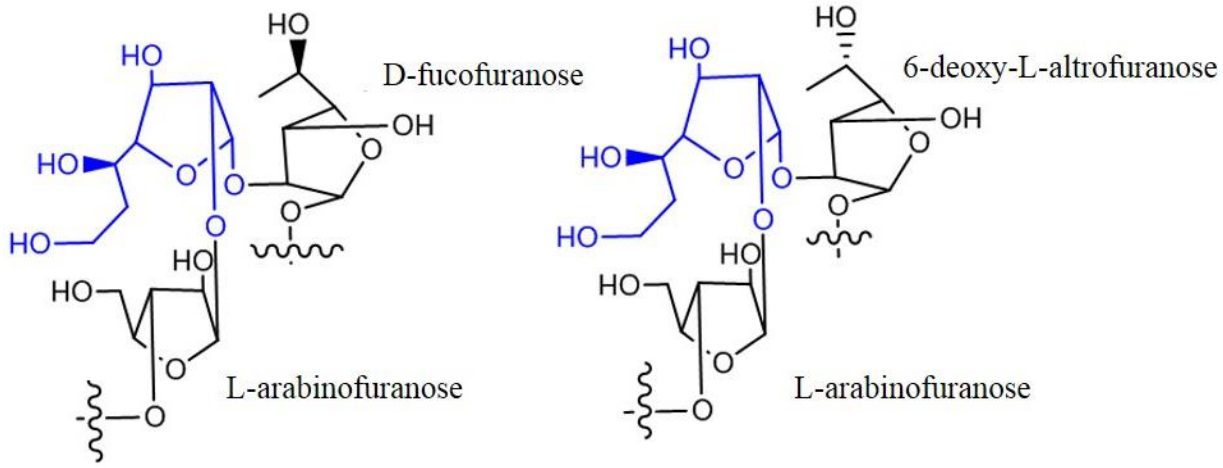


Figure 1-5: The trisaccharide repeat unit isolated from *C. jejuni*, HS: 41 serotype.
Two different trisaccharide repeat units having both 6-deoxy-D-*altro*-Hep*f* (in blue) and L-arabinofuranose have been shown. The difference between these trisaccharide repeat units is the presence of either D-fucofuranose or 6-deoxy-L-altrofuranose.

1.6 Precursors of furanoses

The common theme emerging from the biosynthetic pathways described above is the requirement of sugar nucleotides for the formation and incorporation of furanose sugars in the cell wall. Nucleoside diphosphate sugars act as precursors of the furanoses. In other words, the activated furanose sugar donors are essential for the formation of furanose residues. UDP-D-Gal*f* is thus the precursor of Gal*f* residues.[21] Similarly, UDP-L-Ara*f* and GDP-6d-D-*altro*-Hep*f* are the precursors of L-Ara*f* and 6d-D-*altro*-Hep*f*, respectively.

1.7 Pyranose-furanose mutases

Another theme observed in these pathways is the presence of a pyranose-furanose mutase enzyme which catalyzes the production of the precursors described above. These mutases are named based on the nucleotide-sugar they act on. For example, in the illustration shown for the biosynthesis of D-Gal*f* (Figure 1-3), the mutase that works on interconverting the six and five-membered galactose rings using UDP as the nucleotide is named UDP-galactopyranose mutase (UGM). Pyranose-furanose mutases which are the focus of this thesis will be introduced in this chapter. Apart from UGM, UAM, and GDP-*altro*heptopyranose mutase (GaHM) will be discussed.

1.7.1 UDP-galactopyranose mutase

The *glf* gene encoding for UGM was first identified in *E.coli*.[2] Subsequently, the gene has been identified in *K. pneumoniae*, *M. tuberculosis*, and many other pathogenic bacteria, fungi, and protozoan parasites.[22] UGM is a flavin adenine dinucleotide (FAD) containing enzyme that interconverts between UDP-Gal*p* (six-membered ring) and UDP-Gal*f* (five-membered ring). The equilibrium of this ring contraction reaction favors the six-membered pyranose ring formation in a ratio of 9:1. The reaction scheme is shown below (Figure 1-6).

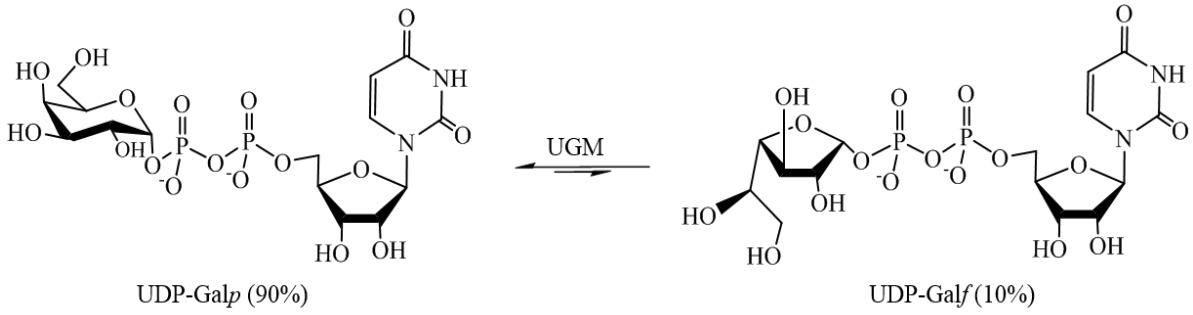


Figure 1-6: Reaction catalyzed by UGM.

FAD is non-covalently bound to the enzyme.[23] The purified UGM fractions were yellow in color and peaks corresponding to the presence of flavin were observed in UV-visible spectra at A_{382} and A_{450}. To provide further evidence for the presence of the FAD cofactor, UGM fractions were thermally denatured, and the peaks corresponding to FAD were observed when analyzed on HPLC.[2] Later, crystal structures were solved to study the binding and interaction of the FAD cofactor with UGM. In general, the molecular weight of UGM was determined to be in the 43 - 45 kDa range. The molecular weight of *E. coli* UGM (*Ec*UGM) determined by mass spectrometric analysis was 43 kDa, and that of *K. pneumoniae* UGM (*Kp*UGM) was ~ 45 kDa. Gel filtration techniques revealed that *Kp*UGM exists as a dimer.[3] The bacterial UGMs identified thus far exist as a homodimer except *Deinococcus radiodurans* (*D. radiodurans*) UGM (*Dr*UGM), which is believed to exist and function as a decamer.[24]

Gene knock-out experiments have demonstrated that the loss of the *glf1* gene rendered *Mycobacterium smegmatis (M. smegmatis)* cells incapable of growth.[9] Significant interest was generated in studying this mutase and its mechanism, due to its important role in bacterial survival. Moreover, the absence of both D-Gal*f* residues and UGM in humans and mammals means that UGM is a potential drug target.

1.7.2 Mechanism of UDP-galactopyranose mutase

Considerable effort has been dedicated to investigating the mechanism of action of UGM. This unique ring contraction reaction is proposed to proceed through the formation of an iminium ion intermediate formed via an S_N2 reaction (Figure 1-7).[25] It was already known from positional isotope exchange (PIX) experiments that the bond between the C1 anomeric carbon of Gal*p* and UDP breaks and reforms during the reaction (Figure 1-7). Additionally, the flavin had to be

reduced for the reaction to occur, as the enzyme remained inactive under oxidizing conditions.[23] Evidence for the formation of the iminium intermediate was obtained by trapping the FAD-galactose adduct using sodium cyanoborohydride and radiolabeled UDP-Gal*p*.[25] Once the iminium ion intermediate is formed, the reaction proceeds by ring closure to form the five-membered furanose and attack of UDP to form UDP-Gal*f*, as shown in Figure 1-7. UGM reconstituted with 5-deaza-FAD (lacking the flavin N5) was unable to demonstrate any cleavage of the anomeric C1 - O (UDP) bond, suggesting that the reaction is unable to proceed without the presence of the N5 of flavin.[26,27] In the crystal structures of UGM, solved with the substrate (to be discussed later), the N5 of reduced FAD (FAD_{red}) is poised for the nucleophilic attack on the C1 anomeric carbon; electron density observed between N5 and C1 provides further evidence in favor of the nucleophilic attack.[24,28]

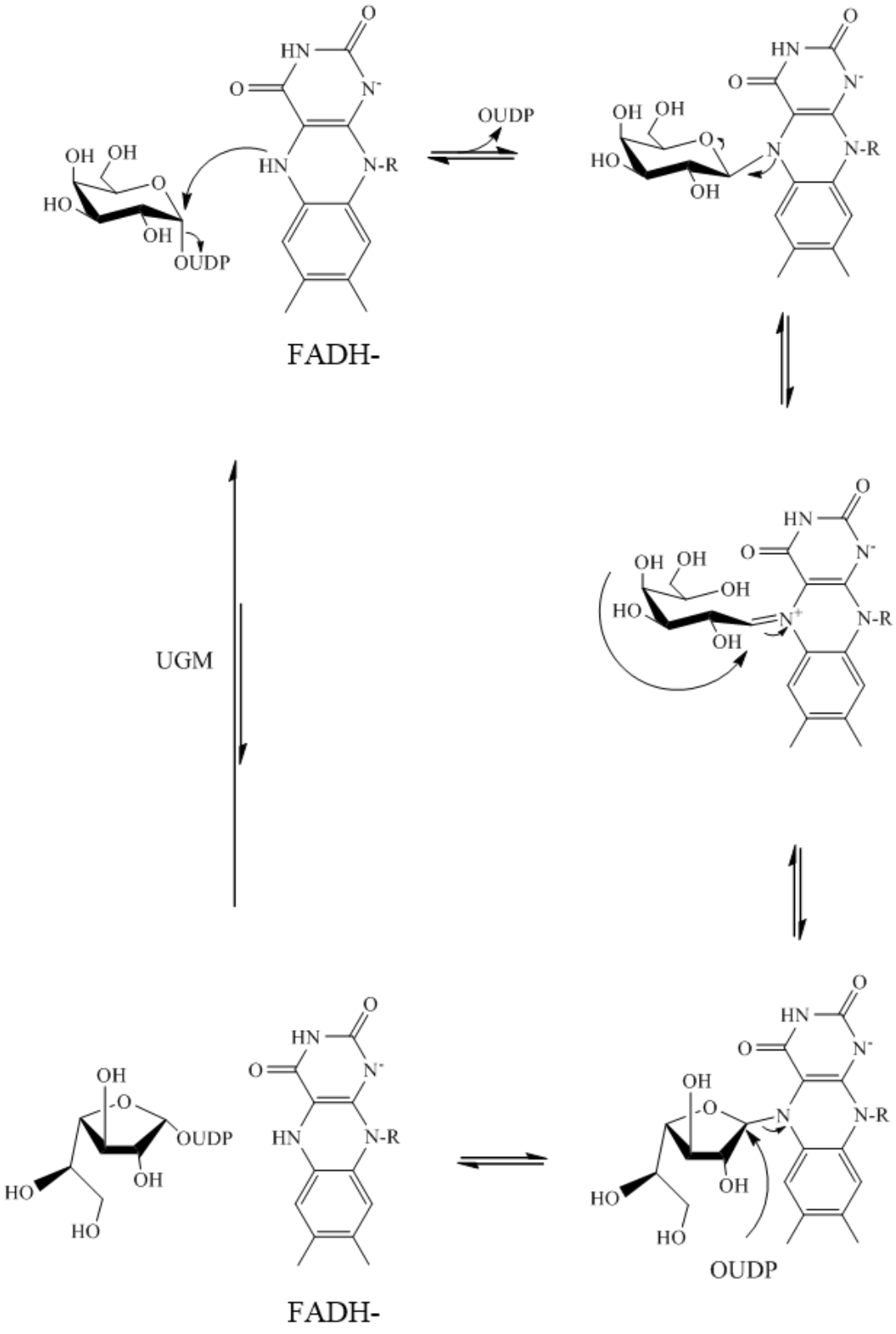


Figure 1-7: S_N2 mechanism proposed for the reaction catalyzed by UGM.

1.7.3 Structure of prokaryotic UDP-galactopyranose mutases

X-ray crystallography was used to gain insights into the structure of UGM. So far, the crystal structure of a variety of prokaryotic and eukaryotic UGMs has been determined, and the structure has been well examined. The first crystal structure determined was that of *Ec*UGM, solved to 2.4 Å, without the substrate (Figure 1-8). Each monomer unit can be divided into three domains, with a flexible loop and a bound molecule of the cofactor FAD.[23]

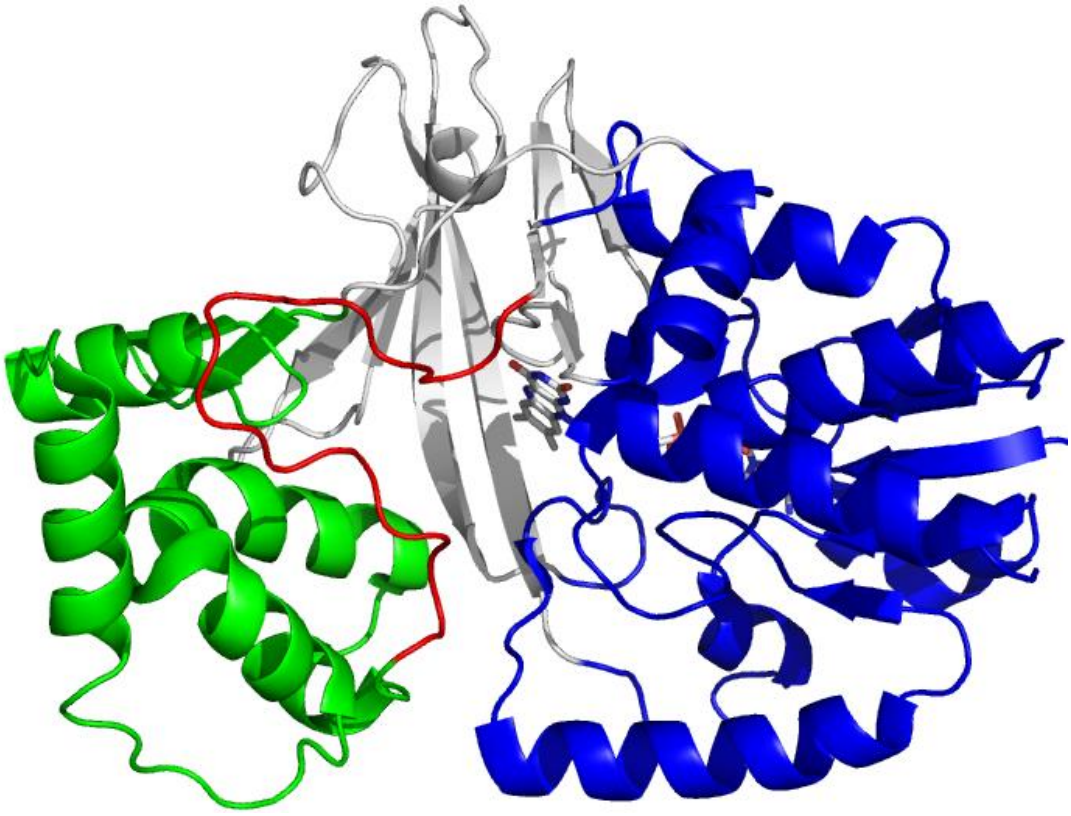


Figure 1-8: Crystal structure of prokaryotic *Ec*UGM
The three domains and the mobile loop are highlighted for the *Ec*UGM crystal structure (PDB id: 1I8T); Domain 1 (blue) Domain 2 (green) and Domain 3 (gray). FAD (white stick model) is shown in Domain 1.

Domain 1 has the FAD binding Rossmann fold, made up of the unique βαβ motif. This domain is common among proteins that bind nucleotides such as nicotinamide adenine

dinucleotide (NAD), nicotinamide adenine dinucleotide phosphate (NADP) and FAD.[29,30] A number of conserved residues serve to bind FAD, mainly by hydrogen bonding interactions. The isoalloxazine ring of FAD is located at the top of a large cleft, below Domain 1, and is oriented to face domain 2. The α-helical Domain 2 lies below the cleft, opposite to Domain 1. This domain has five helices and a long, highly flexible loop. Domain 3 has six β-strands forming an antiparallel β-sheet, which connects Domains 1 and 2 and seals one end of the cleft formed in the front. The substrate was proposed to bind in the cleft region. Two identical monomer units of the enzyme form contact through some non-conserved residues in Domains 2 and the enzyme exists as a homodimer.

The crystal structure of *Kp*UGM and *M. tuberculosis* (*Mt*UGM) were solved later and showed similar overall structures and domain organization, despite their moderate sequence identity to *Ec*UGM, as shown in Figure 1-9.[31] The structures of *Kp*UGM solved with FAD in the oxidized form (FAD_{ox}) and FAD in the reduced form (FAD_{red}) were also similar, except for the small yet significant changes in the FAD puckering. In the FAD_{ox} structures, the isoalloxazine ring of FAD was planar compared to the bent conformation (butterfly-shaped) observed in the FAD_{red} structures; the N5 of FAD is sp^2 and sp^3 hybridized in the FAD_{ox} and FAD_{red} structures respectively.[24,31]

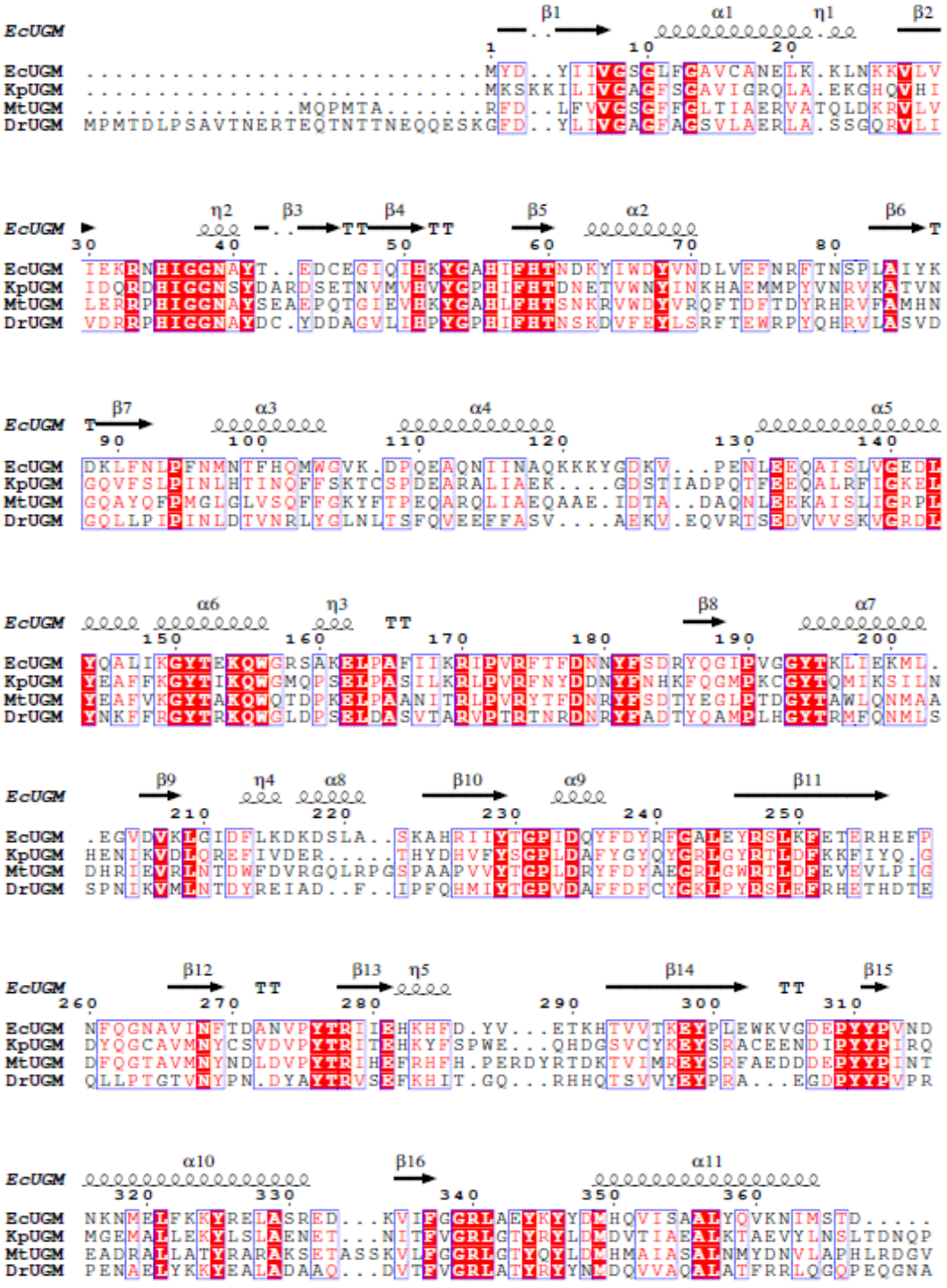


Figure 1-9: Sequence identity of UGM from different bacterial species.
The sequence alignment was performed using ESPript (Version 3.0). The conserved residues are in red blocks (white letters).

*Dr*UGM, *Mt*UGM, and *Kp*UGM were co-crystallized with the substrate UDP-Gal*p*.[24,28,32] These crystal structures allowed for the determination of the substrate binding mode of prokaryotic UGMs and the changes undergone by the enzyme when binding the ligands. The substrate binds to the active site, in the cleft below the isoalloxazine ring of FAD. Substrate recognition and interaction are conserved among UGMs, and each domain contributes residues that recognize the different regions of the substrate. Two residues from the fourth helix of domain 1, a conserved phenylalanine and tyrosine (Tyr), serve to stack the uracil. The substrate diphosphate is held by hydrogen bonding interactions with two conserved arginines. One arginine is contributed by Domain 3 while the other moves in with the mobile loop, which seals the active site completely by moving into a closed conformation in the presence of substrate. This is a common feature seen in all UGMs studied thus far.[33] The substrate sugar (Gal*p*) is held below the FAD through hydrogen bonds made with the Gal*p* hydroxyls.

1.7.4 Structure of eukaryotic UDP-galactopyranose mutases

The three-dimensional crystal structure of eukaryotic UGMs from *A. fumigatus* (*Af*UGM) and *Trypanosoma cruzi* (*T. cruzi*; *Tc*UGM) were solved recently.[34-37] Eukaryotic UGMs have a similar overall fold and domain organization to the prokaryotic UGMs, despite low sequence identity.[37] The three domains described in prokaryotic UGMs also exist in eukaryotic UGMs (Figure 1-10). However, the monomer units of eukaryotic UGMs are larger in size and have regions of insertions in all domains, owing to their longer amino acid sequence (approximately 100 residues longer). Unlike prokaryotic UGMs, that exist as homodimers (with the exception to *Dr*UGM), the two eukaryotic UGMs display different monomeric states. A helix inserted on the C-terminus of domain 1 allows *Af*UGM to exist as a tetramer; *Tc*UGM, which lacks this helix,

exists as a monomer.[33] Perhaps one of the more significant insertions in eukaryotic UGMs is that of a helix in Domain 2, which introduces a second mobile loop in the monomer unit that leads to significant differences between the UGM classes.

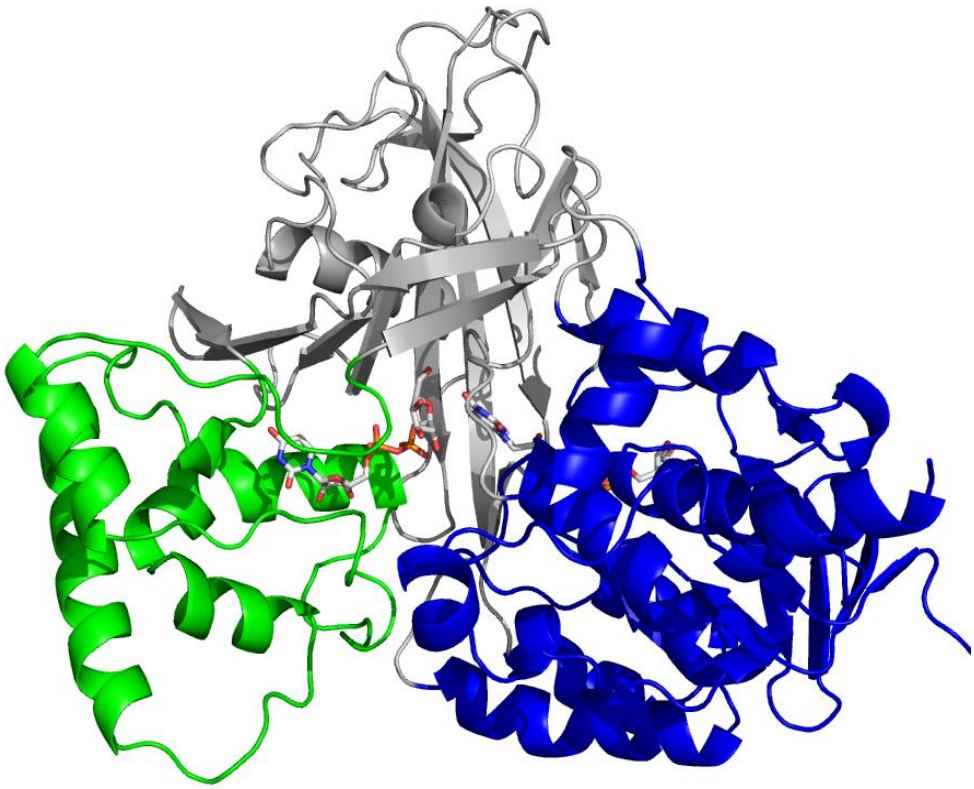


Figure 1-10: Crystal structure of eukaryotic *Af*UGM.
Three domains, Domain 1 (blue) Domain 2 and Domain 3 (gray) are highlighted. The substrate UDP-Gal*f* (white stick model in Domain 2) and FAD (white stick model in Domain 1) are also shown (PDB id: 3UKA).

Most features of substrate binding are common to both classes of UGMs; the substrate binds in the cleft created by the domains, and most of the active site residues are conserved across both UGM classes. However, more positional changes are required for the substrate binding residues of eukaryotic UGMs to accommodate UDP-Gal*p*. The active site of prokaryotic UGMs is more or less set before the substrate comes in and minimal changes are required to bind the substrate in the productive conformation. There are also some differences in the substrate binding

mode. The uridine and the α-phosphate of the substrate UMP, are held at a tilted angle in the active site of eukaryotic UGMs.[33] Other minor differences can also be pointed out. In the uridine-binding region of eukaryotic UGMs, a glutamine residue (Gln107 in *Af*UGM) that forms hydrogen bonds with the uracil ring is located under two tyrosines. In prokaryotic UGMs, the uracil ring is stacked in between aromatic residues, and there is no residue performing a role similar to that of glutamine. These differences arise due to the structural changes in eukaryotic UGMs, but as observed in prokaryotic UGMs, they ultimately aid in the positioning of Gal*p* under the FAD cofactor.

1.7.5 Comparison between prokaryotic and eukaryotic UDP-galactopyranose mutases - mobile loop flexibility

Prokaryotic UGMs have one mobile loop while eukaryotic UGMs have two. The mobile loops are solvent exposed and stay in an open conformation until the substrate enters the active site. In prokaryotic UGMs, upon arrival of the substrate, the mobile loop moves to donate a crucial arginine (~ 7-11 Å shift in Cα positioning) which stabilizes the α-phosphate.[24] In some structures, this residue also hydrogen bonds with the C3 hydroxyl of the sugar, thus forming the 'closed conformation' of the mobile loop. Some residues in the loop, including this arginine, form a short helix which allows for further stabilization of the loop. The helices in the lower region of Domain 2 move in the direction of the active site (by ~ 3 to 8 Å) (Figure 1-11A). This movement and the stabilization of the mobile loop showcase the enzyme's ability to bind and bury the substrate completely in the active site.

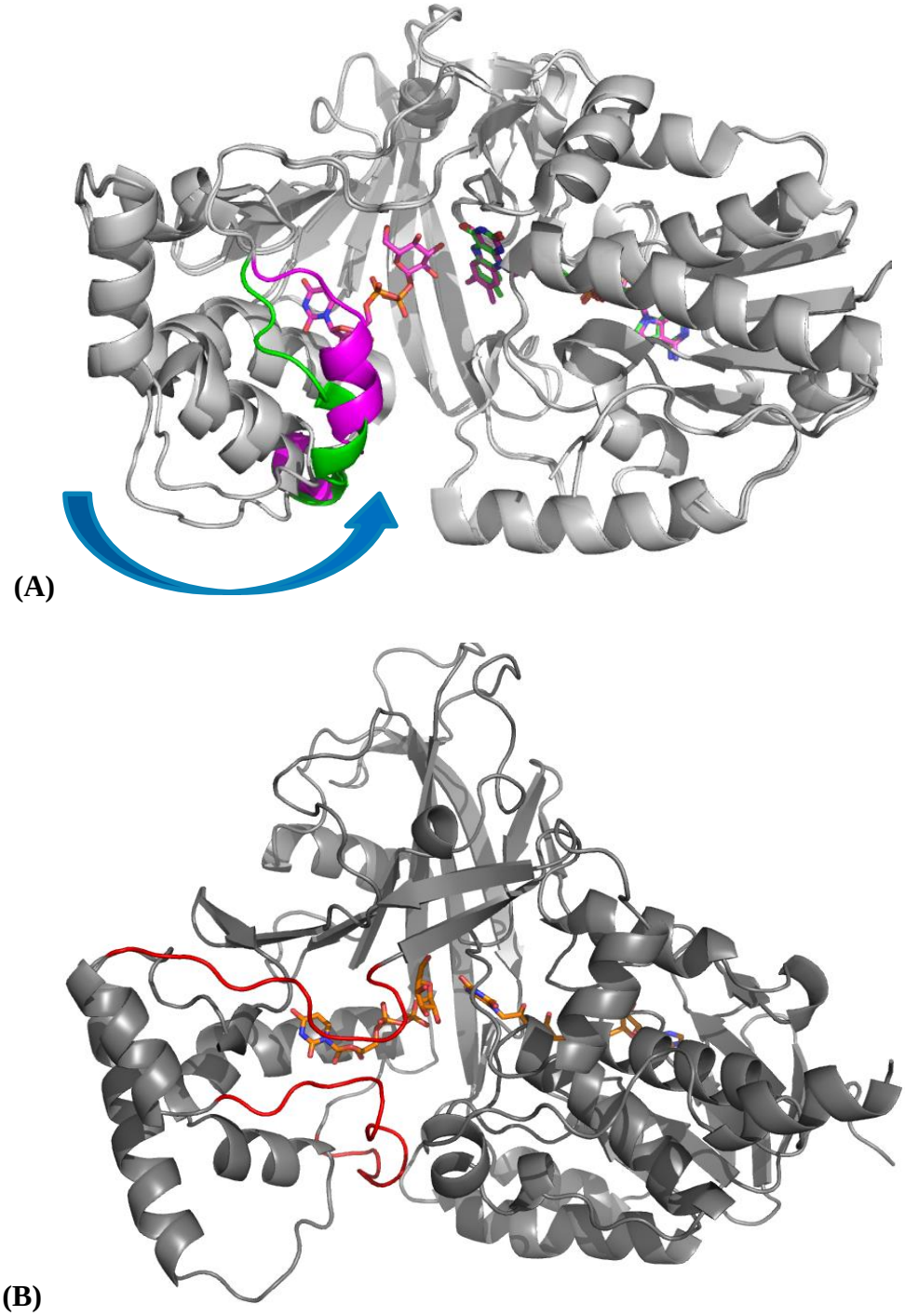


Figure 1-11: A comparison of the loops in both prokaryotic and eukaryotic UGMs.
(A) Prokaryotic UGMs have a single flexible loop shown in open conformation (in green) and closed conformation (in magenta) in the presence of substrate (PDB id: 1I8T). The blue arrow shows how Domain 2 moves to close the active site. (B) Eukaryotic UGMs have two flexible loops, shown in closed conformation since the substrate is present in the active site (PDB id: 3UKA).

Apart from mobile loop 1, eukaryotic UGMs have a second loop (mobile loop 2) located on top of the extra helix of Domain 2, which moves towards the active site in the presence of the substrate (Figure1-11B). In this case, the movement is restricted to the two mobile loops, while only little change occurs in the positioning of the adjacent domain helices. Like the prokaryotic UGMs, loop 1 (Cα of Arg182 moves ~ 11 Å in *Af*UGM) moves to stabilize the substrate α-phosphate. Significant movement (~ 14 Å shift in Cα positioning of Pro206 of *Af*UGM) is also observed in loop 2, which brings in Asn207 (*Af*UGM) into the active site.[35] This asparagine residue hydrogen bonds with the O4 hydroxyl of the Gal*p* sugar. Both loops move to bury the substrate completely in the active site of the enzyme.

1.7.6 Mutation and modeling studies on UDP-galactopyranose mutases

Before crystal structures with ligands were obtained, modeling, molecular dynamics, and docking studies were able to predict the mobile loop movement to form the open and closed conformations and the significance of the arginines that stabilizes the phosphate.[38-40] Over the years, a number of mutants were created using site-directed mutagenesis (SDM); crystal structures of some of these mutants were solved with ligands and kinetic assays were conducted to describe the importance of these critical residues in the active site of UGMs from both classes. The two arginines that stabilize the di-phosphate of the substrate were mutated in both prokaryotic UGMs and eukaryotic UGMs. *Kp*UGM R174A mutant (arginine from mobile loop 1) inactivated the enzyme completely, while *Af*UGM R182A mutant displayed significantly reduced efficiency.[35,40] This shows that the arginine residue is more important in prokaryotic UGMs while eukaryotic UGMs have other residues that help in stabilizing this phosphate. Both *Kp*UGM R280A and *Af*UGM R327A (arginine that stabilizes the ß-phosphate) mutations resulted in no detectable

activity. *Af*UGM R182K and *Af*UGM R327K mutants displayed 10 to 70 fold less efficiency compared to the wild-type enzyme. The crystal structure of *Af*UGM R327K shows that the substrate is still able to bind in the productive conformation.[35] However, the crystal structure of *Af*UGM R327A displays the substrate bound in a non-productive conformation, with the anomeric carbon of the sugar further away from the N5 of the isoalloxazine ring of FAD. Other active site residue mutants of *Ec*UGM and *Kp*UGM were studied to understand the importance of these residues to the active site. In *Kp*UGM, the tyrosines which help in the stabilization of the substrate diphosphate were mutated to phenylalanines. It was observed that the mutation only decreased substrate binding considerably, but did not inactivate the enzyme.[40]

1.7.7 Inhibitors of UDP-galactopyranose mutase

Since UGM is a potential drug target, a significant amount of work has been dedicated to identifying and validating inhibitors for the enzyme. Generally, the approach for developing inhibitors involves synthesizing substrate analogs or identifying lead compounds through virtual screening of compound libraries. A number of groups have designed and synthesized substrate-like inhibitors of UGM. A variety of sugar-based inhibitors, uridine-based derivatives, fluorinated exo-glycal compounds and fluorine substituted sugar-based compounds, and substrate analogs have been studied.[41-50] The fluorine substituted compounds mostly served the purpose of gaining insight into the mechanism of UGM function and were poor substrates. Most of the inhibitors mentioned above did not show satisfactory inhibition against bacterial UGMs. The few that displayed inhibition against bacterial UGMs (*Ec*UGM, *Kp*UGM, and *Mt*UGM) were in the millimolar to micromolar range.[50] Due to their polarity, low cell permeability, and cytotoxicity, further improvement of these compounds is required. Compounds obtained from screening

libraries, containing the 5-arylidene-2-thioxo-4-thiazolidinone core (Figure 1-12), aminothiazole core and the more recently discovered triazolothiadiazine series have shown moderate inhibition of UGM.[51-53] Although these compounds were found to thwart bacterial growth, some compounds such as the aminothiazole core compounds were found to be toxic to human cells.[54]

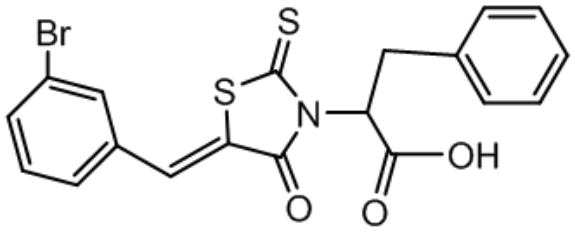


Figure 1-12: 5-arylidene-2-thioxo-4-thiazolidinone core based UGM inhibitor.

1.8 Active sites of prokaryotic and eukaryotic UDP-galactopyranose mutase

Since understanding the substrate binding of pyranose-furanose mutases is the theme of this thesis work, it is important to discuss what is previously known about substrate binding in these mutases. The majority of our knowledge arises from UGM, this part of the introductory chapter will strive to provide a comparison of substrate binding modes of UGM, based on crystal structures solved from prokaryotic and eukaryotic organisms. A comparison of the conserved active site residues of prokaryotic UGMs, such as *Ec*UGM, *Kp*UGM, and *Mt*UGM, and eukaryotic *Af* UGM are shown in Table 1-1.

Table 1-1: Comparison of active site residues of various UGMs.

***Ec*UGM**	***Kp*UGM**	***Mt*UGM**	***Af*UGM**
H56	H60	H65	H63
H59	H63	H68	F66
N80	N84	H89	R91
L147	F151	F157	F158
I148	F152	V158	M159
Y151	Y155	Y161	Y162
T152	T156	T162	N163
W156	W160	W166	W167
R170	R174	R180	R182
Y181	Y185	Y191	P206
F182	F186	F192	N207
N268	N270	N282	Y317
R278	R280	R292	R327
E298	E301	E315	E373
Y311	Y314	Y328	Y419
R340	R343	R360	R447
Y346	Y349	Y366	Y453
D348	D351	D368	N457
M349	M352	M369	Q458

*Kp*UGM and *Mt*UGM have both unliganded crystal structures and those that were complexed with UDP-Gal*p*, for prokaryotes UGMs.[28,31,32] A comparison of these structures has indicated the changes in the α-helical domain (Domain 2) and those undergone by some of the active site residues in order to bind and bury the substrate within the active site. Likewise, in the case of eukaryotic UGMs, crystal structures of *Af*UGM have been solved with and without the substrate in the active site.[35]

A comparison of the liganded and unliganded crystal structures can help understand the changes that happen to the active site, when the substrate approaches the active site of UGM. Domain 2 re-orients to accommodate the uridine portion of the substrate and this triggers the mobile loop to form the closed conformation and bury the substrate within the active site. The five helices of this helical domain, α4, α5, and α6 and a small loop connecting helices α4 and α5 function as a flexible hinge. The crystal structure of *Mt*UGM with UDP-Gal*p* is a good example.[32] The Domain 2 helices have moved towards the UDP of the substrate while this movement is not observed in the unliganded structure, as shown in Figure 1-13. Additionally, a new helix (helix α8) is formed by residues of the mobile loop.[32]

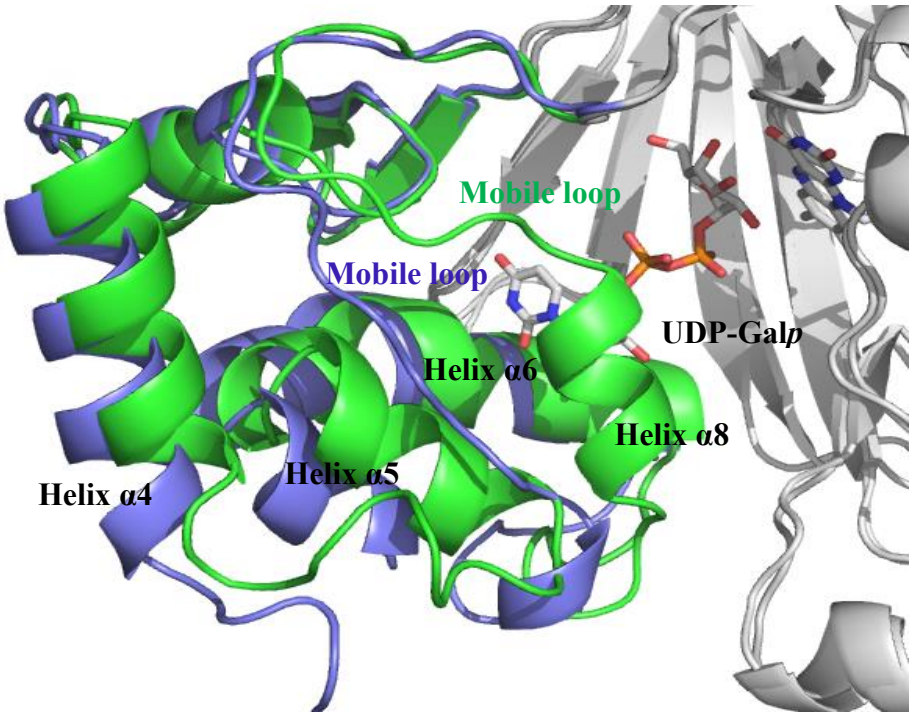
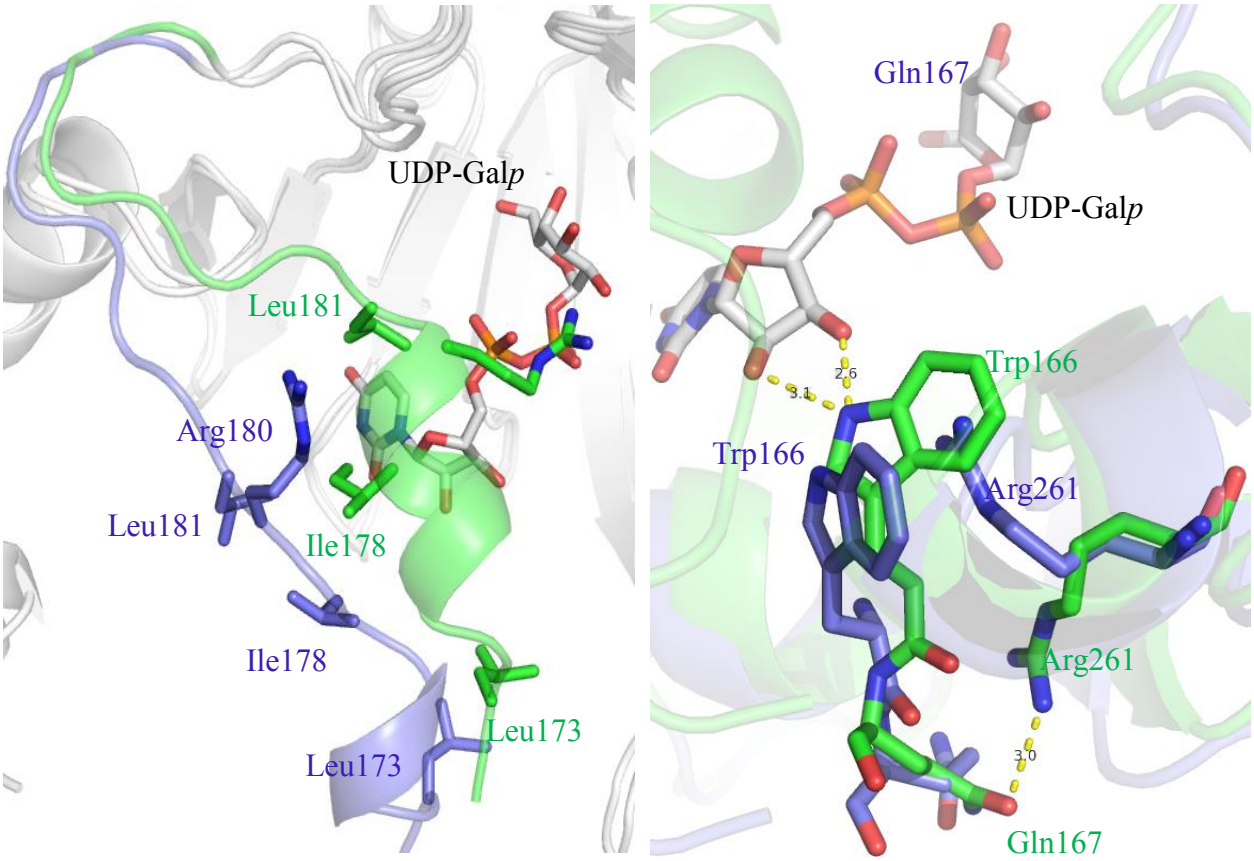


Figure 1-13: An overlay of liganded and unliganded crystal structures of *Mt*UGM.
*Mt*UGM with UDP-Gal*p* (green) and without UDP-Gal*p* (purple) are shown, highlighting the important changes occurring in the α-helical domain, on the arrival of the substrate (PDB id: 1VOJ & 4RPG).

The movement of the mobile loop towards the substrate is expected to close the otherwise open substrate binding site (~ 67 % closure as estimated from *Mt*UGM structure). Apart from bringing in Arginine, which stabilizes the α-phosphate, the mobile loop also brings in hydrophobic residues, which form a hydrophobic pocket that helps in positioning the substrate uridine, as shown in Figure 1-14A.[32]



(A) **(B)**

Figure 1-14: Movement of the mobile loop in the presence and absence of UDP-Galp.
(A) The changes in position of active site residues Leu181, Arg180, Ile178 and Leu173, when the mobile loop moves towards active site in the presence (green) and absence of substrate (purple) (B) The changes in positioning and interactions of residues hypothesized to trigger closure of the loop in the presence (green) and absence (purple) of UDP-Gal*p* (PDB id: 1VOJ & 4RPG).

Apart from the domain movement, the presence of the substrate in the UGM active site also induces local changes in the orientation of residues close by. A tryptophan residue in helix α6

rotates and forms hydrogen bonds with the substrate ribose. This movement breaks its cation-π interaction with an adjacent arginine causing it to rotate back by 180° and form hydrogen bonds with residues nearby. The break in cation-π interaction between Trp166 and Arg261 that occurs during *Mt*UGM substrate binding is shown in Figure 1-14B. The conserved nature of these residues and their positioning in semi-closed and unliganded structures in prokaryotic UGMs lends support to the hypothesis that the movement of these residues could potentially trigger the closure of the active site.

The movement of the α-helices towards the substrate also means the residues have moved further closer to the uridine moiety. Additional changes in the orientation of residues aid in positioning and binding of the uridine moiety. Among these residues, only a tyrosine and asparagine are highly conserved in all prokaryotic UGMs. As shown in Figure 1-15, in *Mt*UGM, Tyr161 rotates ~ 45º and forms cation-π interactions with the uracil and the side chain of Asn282 flips ~ 180º to form hydrogen bonds with the O4 of the uracil ring. A number of different nonconserved residues such as phenylalanine, valine, isoleucine and leucine also aid to position the uracil ring in a hydrophobic pocket in prokaryotic UGMs.

The mobile loop brings an arginine that stabilizes the α-phosphate. In most cases the guanidium moiety of this residue can enable the stabilization of the β-phosphate. Additionally, there are three tyrosines residues which are in position to interact with the substrate diphosphate. Earlier these tyrosines were hypothesized to interact with the sugar hydroxyls, but these evidences show their role in stabilizing the phosphates. These residues are highly conserved in all the prokaryotic UGMs. In Figure 1-16, Tyr366 which moves ~ 5-6 Å and rotates ~ 90º to stabilize the β-phosphate, while also forming a cation-π interaction with Arg180, to maintain the closed conformation of the mobile loop in *Mt*UGM, is shown.

Among the prokaryotic UGMs, there are small differences in the active site residues and their interactions with the sugar portion of the substrate. In *Mt*UGM and *Dr*UGM, the mobile loop arginine forms hydrogen bonds with the 2-OH group of the substrate sugar (Gal*p*), while this interaction is not observed in *Kp*UGM.[24,28,32] The Gal*p* 3-OH is often stabilized by interactions with a conserved histidine residue in bacterial UGMs, such as *Mt*UGM His68. Another histidine residue that interacts with Gal*p* 6-OH in *Mt*UGM is replaced by an asparagine in *Kp*UGM. The incoming substrate triggers the movement of a nearby six-residue loop, towards the adenine moiety of FAD. This loop movement causes another conserved arginine, to re-orient towards the active site and form water-mediated hydrogen bonds with Gal*p* 2-OH (Figure 1-17).

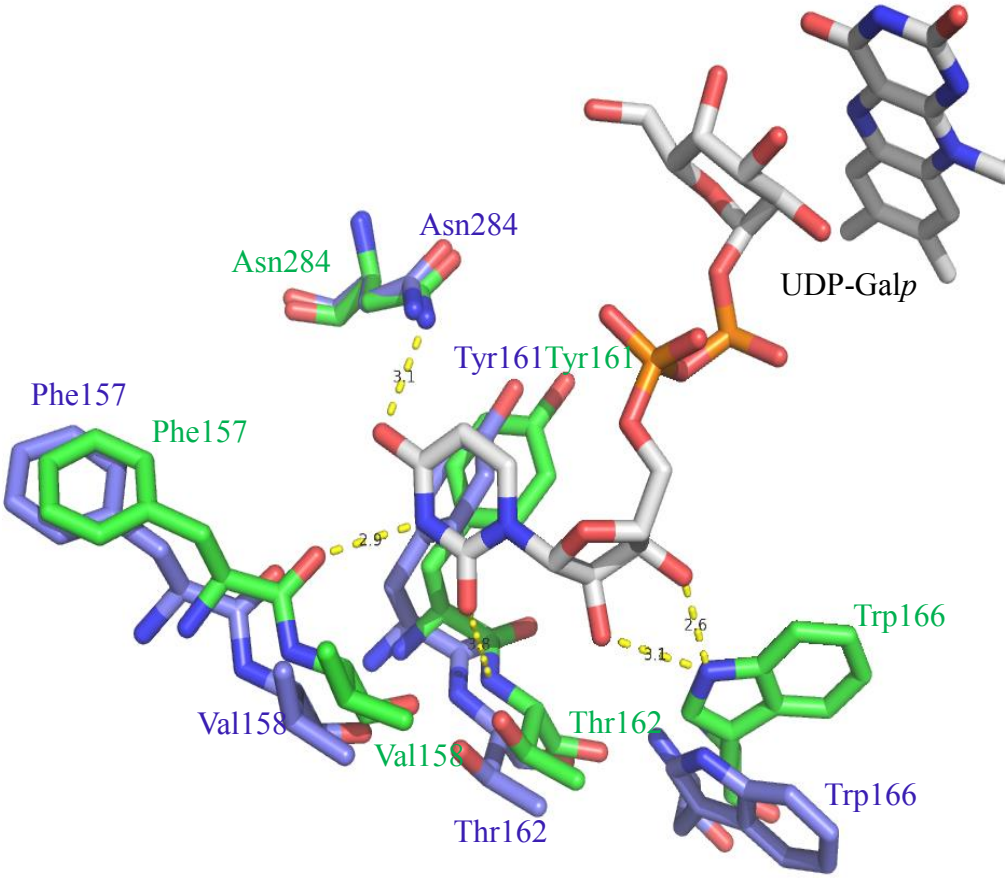


Figure 1-15: The uridine-binding region of *Mt*UGM.
The changes in positions of residues in the absence (purple) and the presence of UDP-Gal*p* (green) are highlighted (PDB id: 4RPG).

Additionally, the O4 of FAD hydrogen bonds with the C4 hydroxyl of Gal*p*. This interaction aids in substrate selectivity by UGM; the enzyme does not recognize UDP-Glu*p* with the C4 hydroxyl in an equatorial position.[28] The substrate bound structures have also generated vital information about the non-productive and productive modes of substrate binding when the FAD is in the oxidized and reduced forms, respectively. In the productive binding mode, the N5 of the isoalloxazine ring of FAD is located within a range of 2.9 to 4.0 Å, from the anomeric carbon of the sugar (Gal*p*), as shown in Table 1-2. The productive mode is observed in the substrate bound FAD_{red} structures.[24] In the non-productive mode, seen in FAD_{ox} structures of various UGMs, this distance is greater than that noted in the substrate bound FAD_{red} structures.

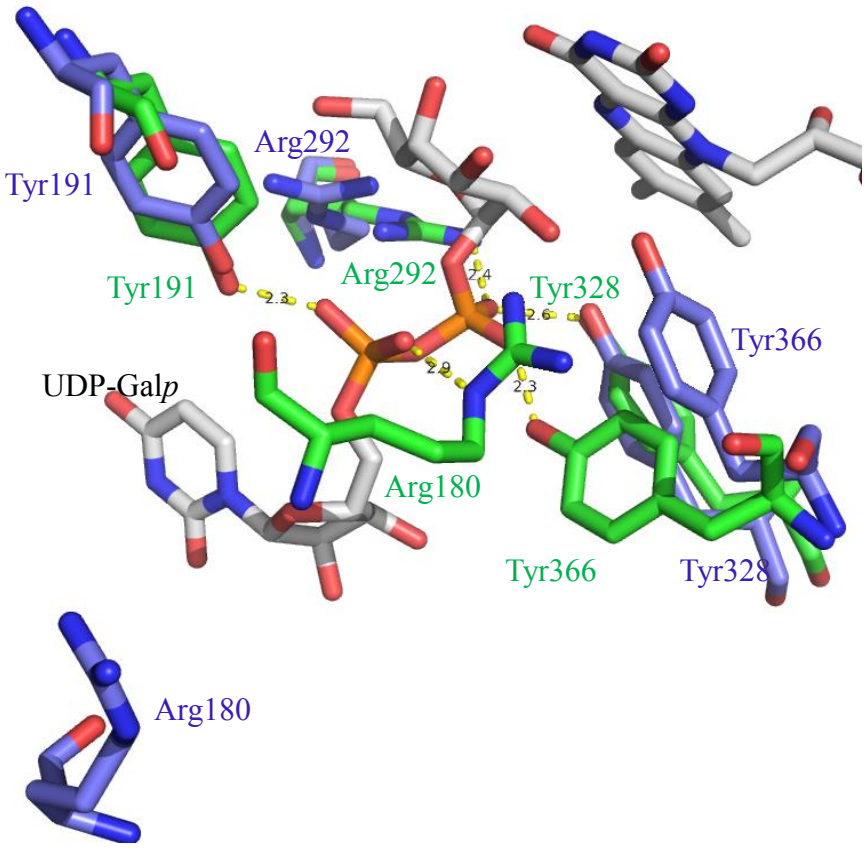


Figure 1-16: The phosphate-binding region of *Mt*UGM.
The changes in the positioning of residues in the absence (purple) and the presence of UDP-Gal*p* (green) are highlighted (PDB id: 4RPG).

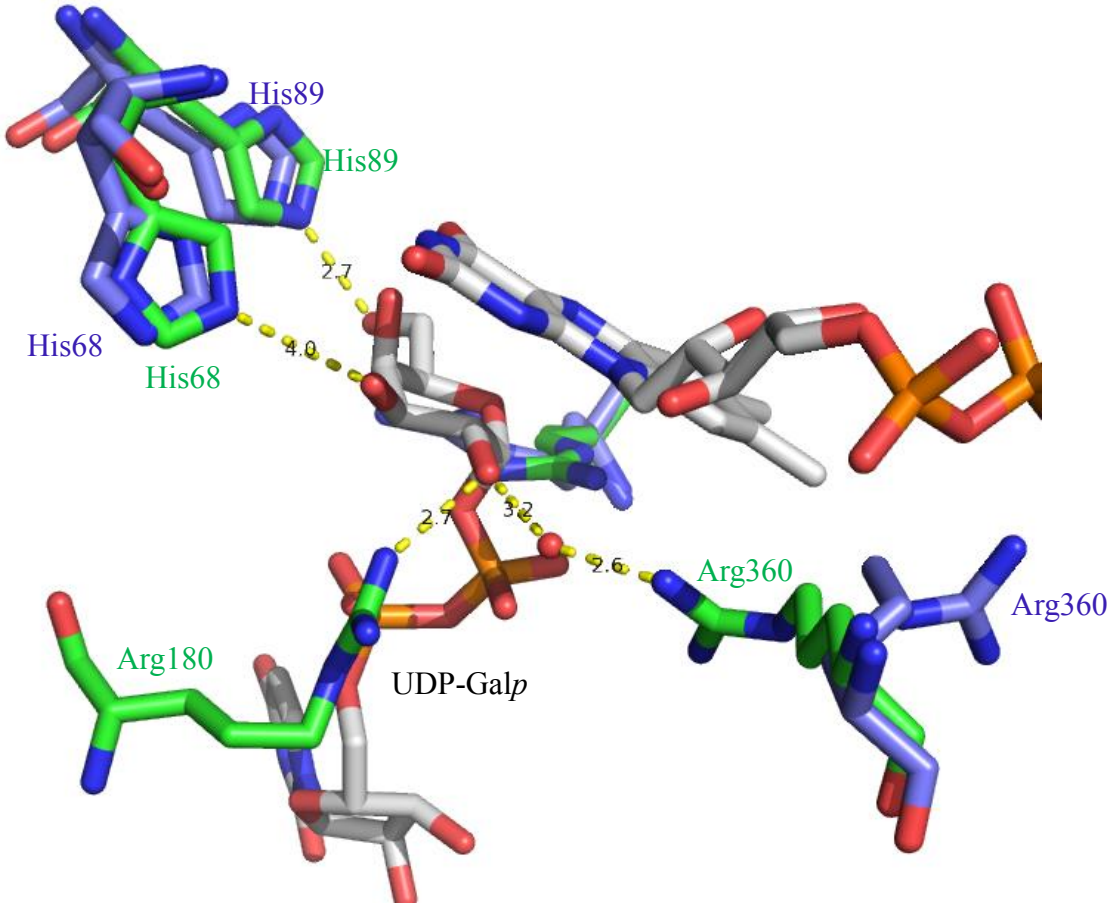


Figure 1-17: The sugar-binding region of *Mt*UGM.
The changes in the positioning of residues in the absence (purple) and the presence of UDP-Gal*p* (green) are highlighted (PDB id: 4RPG).

Table 1-2: A comparison of distances between the N5 of the isoalloxazine ring of FAD and the C1 of Galp from oxidized and reduced crystal structure of various UGMs.

	FAD_{ox} N5 - C1 Gal*p* distance (Å)	**FAD_{red} N5 - C1 Gal*p* distance (Å)**
Prokaryotic UGMs		
*Mt*UGM	4.2*	3.9 - 4.0
*Kp*UGM	7.8 - 8.0	3.6*
*Dr*UGM	3.4 - 3.7	2.9 -3.2
Eukaryotic UGMs		
*Af*UGM	5.2*	3.6*

* Values are taken from a single monomer unit of the enzyme.

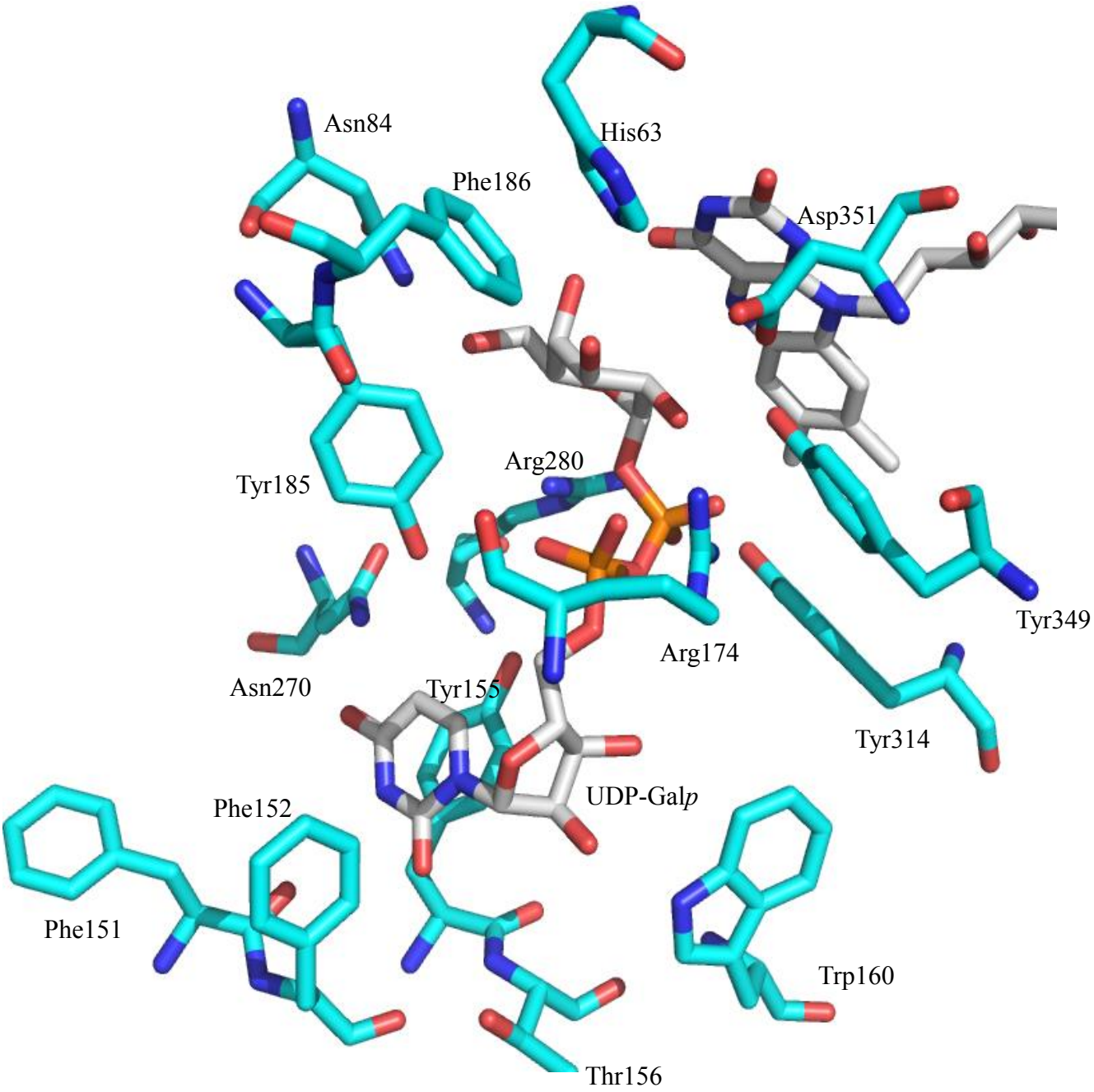


Figure 1-18: The active site residues of *Kp*UGM.
The active site residues (cyan) in the presence of substrate UDP-Gal*p* and FAD are shown as white sticks (PDB id: 3INR).

The substrate binding site of *Kp*UGM, in the presence of UDP-Gal*p*, resembles that of the *Mt*UGM: UDP-Gal*p* active site, shown in Figure 1-18. Although minor differences exist in terms

of the residues that make up this site, it is noted that its overall structure and the substrate binding mode are largely conserved. Furthermore, the changes that occur in the enzyme, to accommodate the substrate in the active site is also very similar across bacterial UGMs as evidenced by crystal structures.[33] Also, the active site of prokaryotic UGMs is comparable to that of eukaryotic UGMs. From Table 1-1, the active site residues are similar if not entirely conserved. However, since eukaryotic UGMs have two mobile loops change to a closed conformation, to bury the substrate in the active site, the functioning of Domain 2 is expected to be different in this case.

Among the eukaryotic UGMs, although crystal structures are available for *Af*UGM and *Tc*UGM, both unliganded and UDP-Gal*p* bound structures are available for only *Af*UGM. The *Tc*UGM crystal structures were solved with UDP in the active site.[36] Although there is not as much re-orientation of the helices in Domain 2, significant changes occur in the conformation of the two flexible loops and the helix that connects them, as shown in Figure 1-19. An analysis of the structures suggest that the movement two flexible loops, brings the residues in place to interact with the uridine moiety of the substrate.[35]

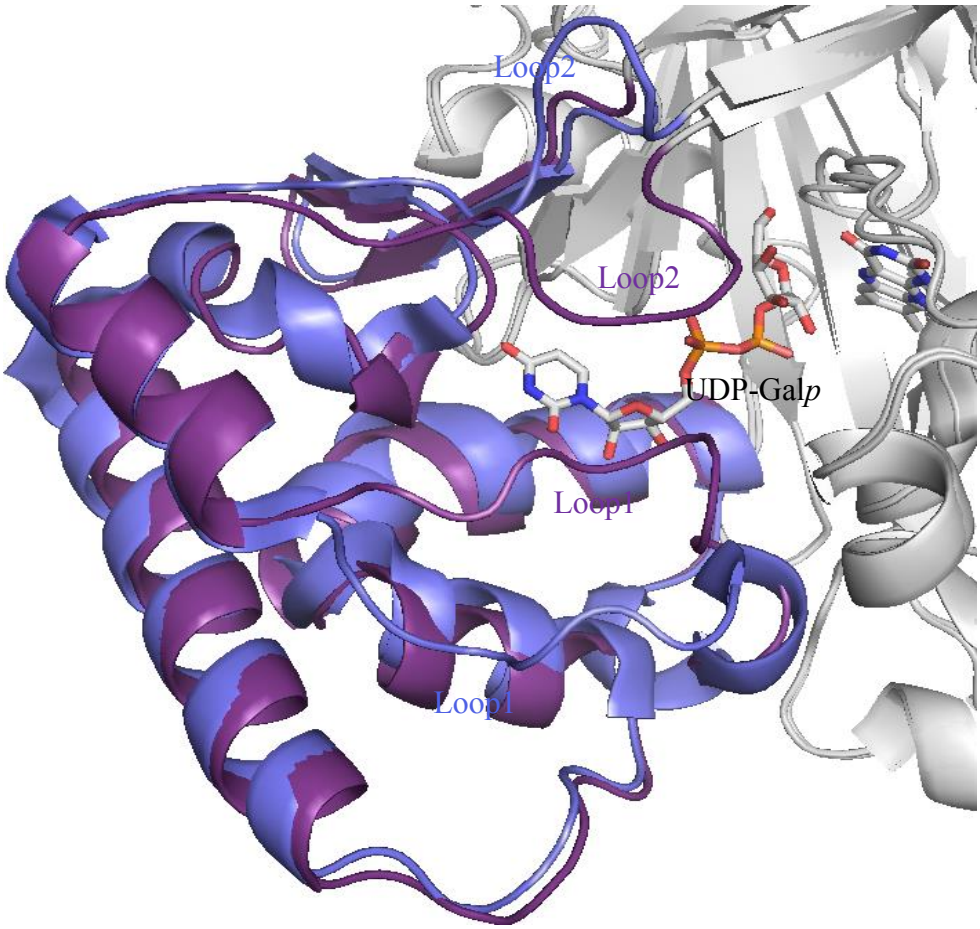


Figure 1-19: Domain 2 of *Af*UGM in open and closed configuration.
The open loop configuration, without the substrate in the active site, and closed loop configuration with the substrate in the active site are shown in purple and burgundy respectively (PDB id: 3UKA & 3UKH).

In the eukaryotic UGM active site, two tyrosine residues stack the uracil, while hydrophobic residues such as proline and phenylalanine aid in positioning it. An additional glutamine residue which moves further into the active site and forms hydrogen bonds with the uracil, is not observed in any prokaryotic UGMs. The closed loop conformation of Loop 2 is stabilized by a highly conserved arginine, which rotates towards Loop 2 and forms hydrogen bonds with the side chains of residues such as tyrosine and some adjacent residues which ends Loop 2, as shown in Figure 1-20.

As already seen with prokaryotic UGMs, Loop 1 in eukaryotic UGMs also contributes hydrophobic residues that aid in positioning the substrate uridine and a negatively charged arginine residue that stabilizes the α-phosphate. This residue also makes additional interactions to stabilize the closed loop conformation, as seen in prokaryotic UGMs. Arg182 of *Af*UGM forms cation-π interaction with Tyr453. In the phosphate binding region, another highly conserved arginine, moves into position to stabilize the β-phosphate.

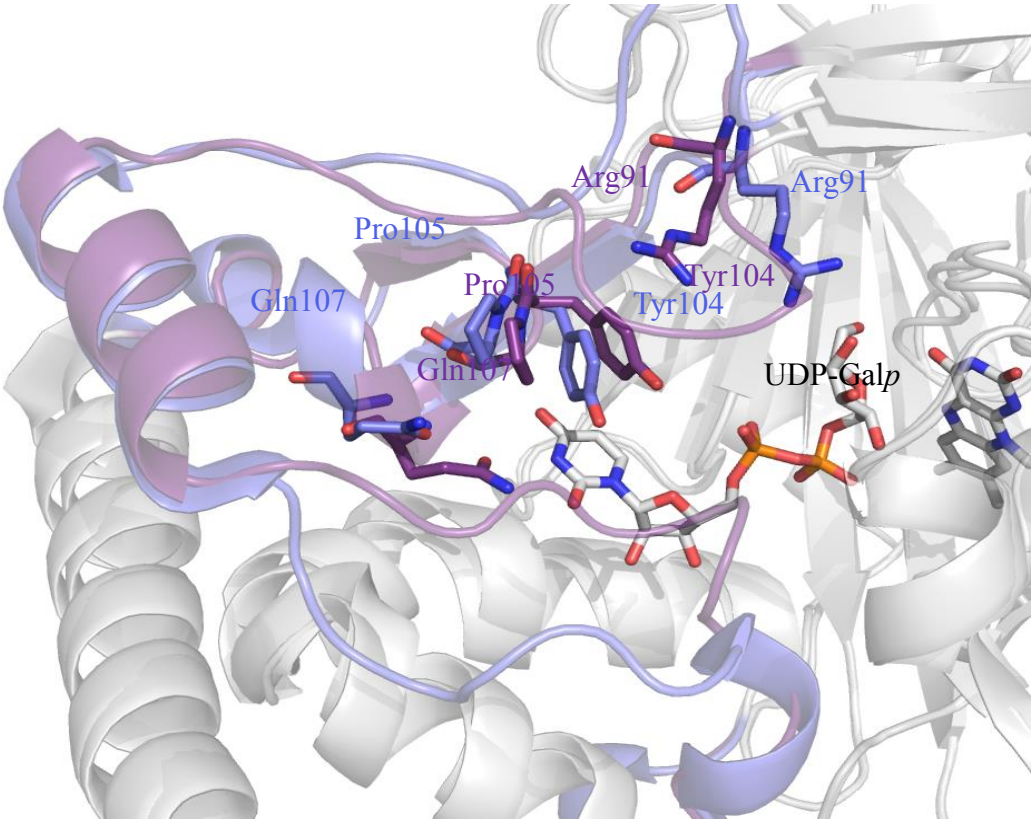


Figure 1-20: The movement of *Af*UGM residues that bind the substrate uracil.
Tyr104, Pro105, and Gln107 that bind the substrate uracil are shown in burgundy, in the presence of substrate. The movement of Arg91 to aid stabilization of Loop 2 is shown in the absence and presence of the substrate in purple and burgundy respectively (PDB id: 3UKA & 3UKH).

The sugar binding region of eukaryotic UGMs are also comparable to prokaryotic UGMs. The sugar is held in the same orientation as seen with the prokaryotic UGMs. In *Af*UGM, residues

such as Asn207, Asn447, and Arg327 hydrogen bond with C3, C4 and C6 hydroxyls of Gal*p*, while Arg447 interacts with the C2 hydroxyl by a water-mediated hydrogen bond. The overall substrate binding site of *Af*UGM is similar to prokaryotic UGMs, as shown in Figure 1-21. It is also noted that the N5 of FAD_{red} is ~3.7 Å away from the Gal*p* C1 carbon while the N5 of FAD_{ox} is ~5.2 Å away from the Gal*p* C1 carbon (Table 1-2).

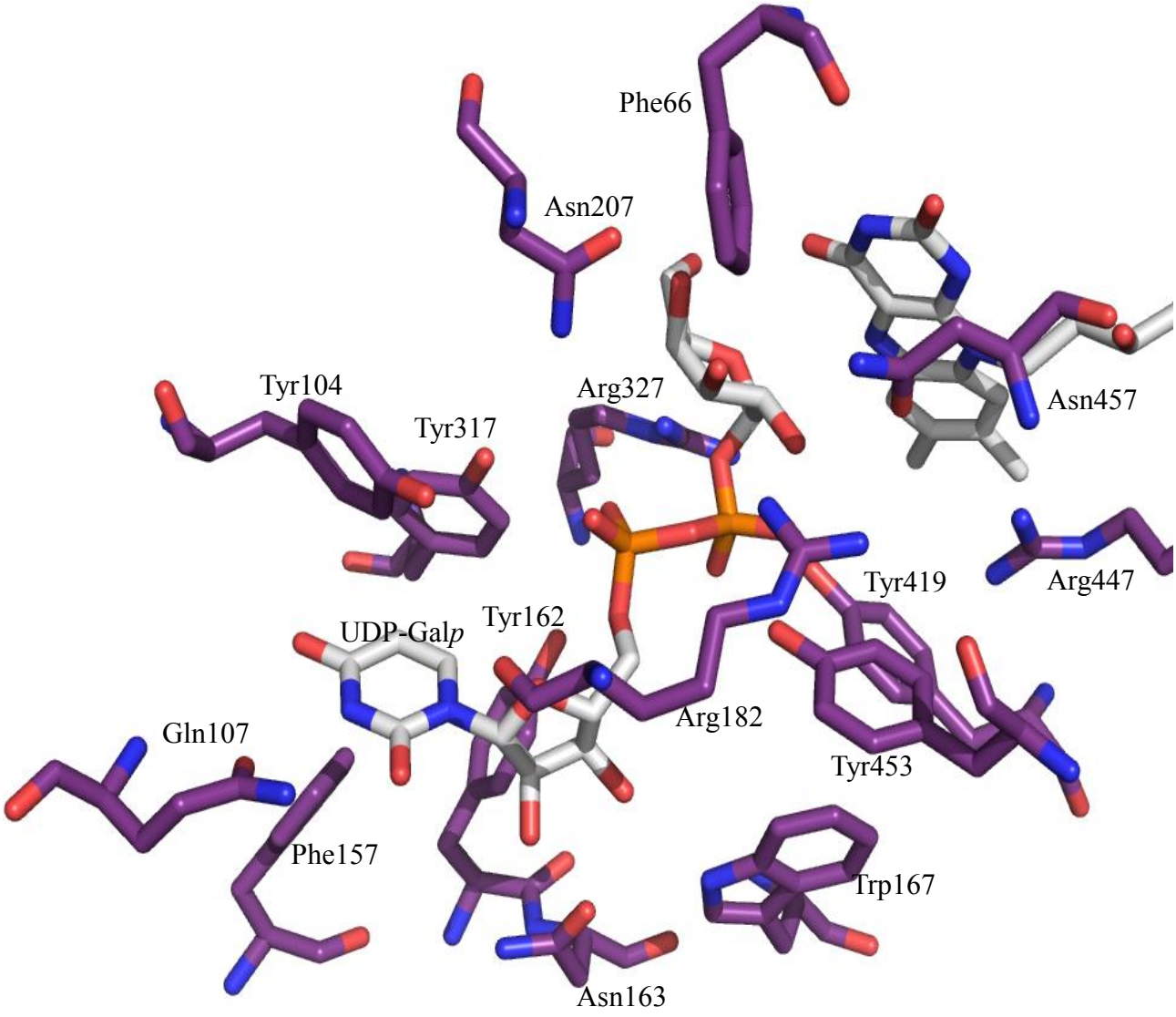


Figure 1-21: A close-up of the active site residues of eukaryotic *Af*UGM.
The active site residues (burgundy) in the presence of substrate UDP-Gal*p* and FAD are shown as white sticks (PDB id: 3UKH).

1.9 Other pyranose-furanose mutases

Until now, only a handful of enzymes has been identified and characterized to perform the pyranose-furanose interconversion reaction. Perhaps the best studied among these enzymes after UGM is UDP-N-acetyl galactopyranose mutase (UNGM), which functions to interconvert between UDP-N-acetyl galactopyranose (UDP-D-Gal*p*NAc) and UDP-N-acetyl galactofuranose (UDP-D-Gal*f*NAc).[55] The enzyme was identified as a gene product of *cj1439c* (a homolog of the *glf* gene), identified in *C. jejuni* strain 11168.[56] Initial characterization of UNGM was performed by Dr. Todd Lowary's group (University of Alberta) and crystallization and structural studies were completed by Dr. Sean Dalrymple and Carla Protsko, previous members of the Sanders group. UNGM is also a flavoenzyme and has a high sequence identity with bacterial UGMs (*Ec*UGM, *Kp*UGM, *Dr*UGM, and *Mt*UGM). Unlike UGM, UNGM is a bifunctional enzyme and can recognize both UDP-D-Gal*p* and UDP-D-Gal*p*NAc as substrates. The crystal structure of UNGM with UDP-D-Gal*p* shows that most of the UNGM active site residues are similar to those from bacterial UGMs, except an arginine residue (Arg59) which interacts with the acetamido group of UDP-D-Gal*p*NAc.[57] GaHM is another flavin-containing pyranose-furanose mutase that has been identified in *C. jejuni*. This enzyme is different from the mutases described so far, as it is hypothesized to require a GDP based substrate, and interconverts between a heptopyranose (7-carbon chain) and a heptofuranose.[58] Though they catalyze similar pyranose-furanose interconversions, these mutases have low sequence identity, as shown in Figure 1-22.

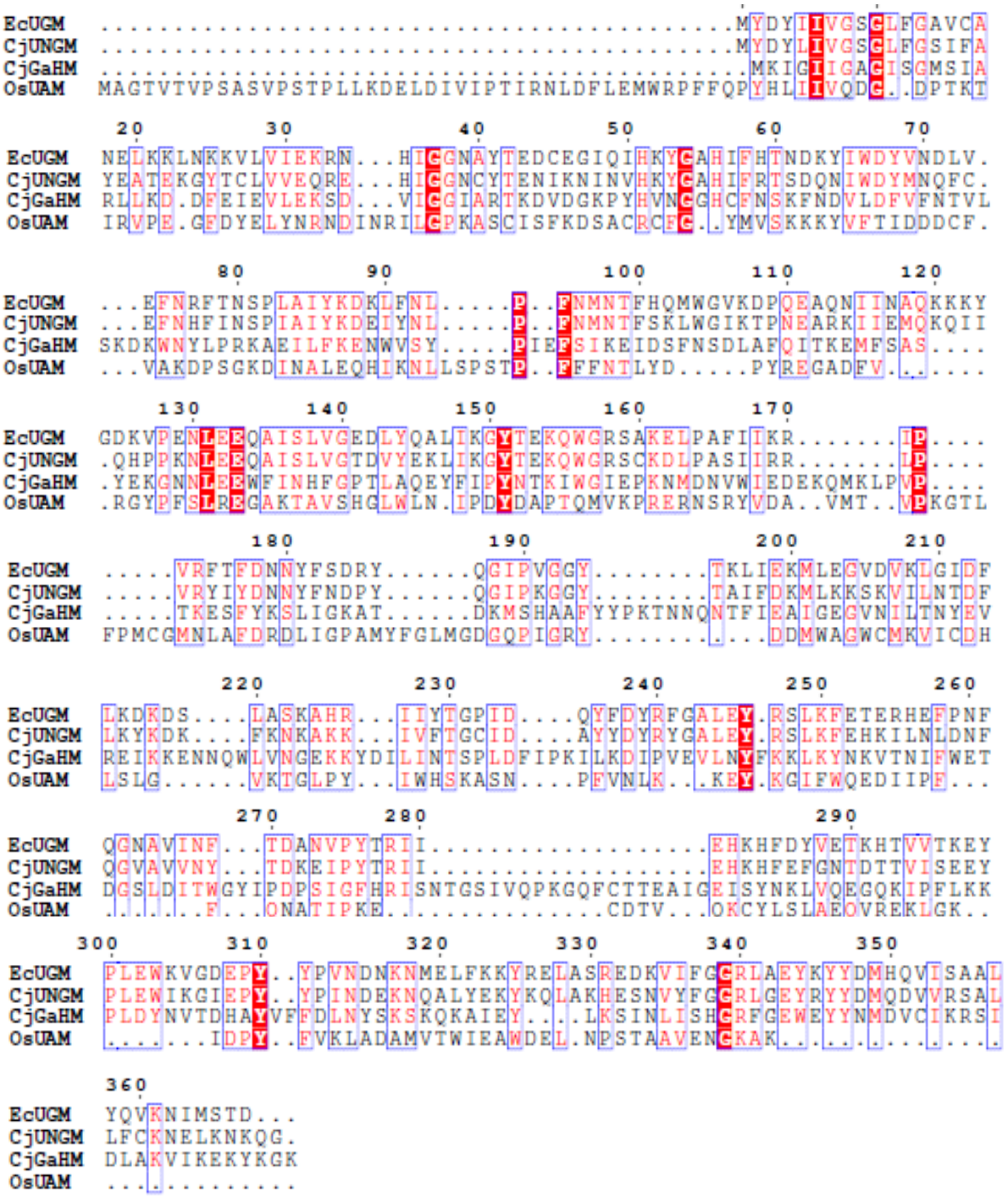


Figure 1-22: Low sequence similarity among the pyranose-furanose mutases.
*Ec*UGM and *C. jejuni* UNGM (*Cj*UNGM) have a higher identity with each other than *C. jejuni* GaHM (*Cj*GAHM) or *Os*UAM. Among all the four sequences *Os*UAM has the least sequence similarity with any of the other enzymes. The sequence alignment was performed using ESPript (Version 3.0). The conserved residues are in red blocks (white letters).

Reports describing the identification and characterization of pyranose-furanose mutases from other organisms have emerged recently. The gene *Fcf2* from *E. coli* O52 encodes for a pyranose-furanose mutase that works in the pathway that synthesizes deoxy thymidine-di phosphate–D-fucofuranose (dTDP-D-Fuc*f*).[59] This enzyme interconverts between thymidine-di phosphate–D-fucopyranose (dTDP-D-Fuc*p*) and dTDP-D-Fuc*f*. Its 40-60% DNA sequence identity with UGMs and the protein sequence alignment has identified the presence of an ADP-binding fold that can bind FAD. Thus, the gene product of *Fcf2* was proposed to be an FAD-requiring mutase that belongs to the UGM family.[59] However, more research has to be conducted to understand substrate binding and how the active site is modified, as compared to other UGMs. An enzyme that is perhaps the most different to any of these mutases is UAM, a metalloenzyme that does not require FAD for performing its function.[18] The next part of this chapter will introduce UAM and GaHM since apart from UGM, this thesis will discuss results pertaining to both these enzymes.

1.9.1 UDP-arabinopyranose mutase

UAM is a pyranose-furanose mutase that functions in the plant cell wall. UAM works in a pathway that leads to the deposition of L-Ara*f* residues in the cell wall. Even though the importance and abundance of L-Ara*f* residues in the plant cell wall is well documented, complete details about the biosynthetic pathway that leads to the deposition of L-Ara*f* residues in the cell wall have only emerged recently.[8,18,60,61]

Among the common nucleotide sugars identified in a plant cell, the biosynthesis of UDP-Ara*p* is well understood. Since only UDP-Ara*p* was previously detected in a plant cell wall, it was considered to be the potential precursor for L-Ara*f* residues.[62] However, experiments conducted to

investigate the theory failed to lend support to this argument. When UDP-Ara*p* was used as the substrate donor for arabinosyltransferases, Ara*p* was transferred onto arabino-oligosaccharides.[61] This meant that plants might require UDP-Ara*f* for biosynthesis of arabinofuranosides. (1, 5) α-L-arabino-oligosaccharides were successfully synthesized *in vitro* by mung bean arabinofuranosyl transferase extracts using UDP-Ara*f* as the substrate.[63] Though this study provided evidence for the production and requirement of UDP-Ara*f* in the plant cell wall, the point of isomerization of Ara*p* to Ara*f* in the biosynthesis pathway was still unknown.

It was demonstrated that plants employ UAM for the conversion of UDP-Ara*p* to UDP-Ara*f* (Figure 1-23). UAM activity was first identified in rice seed extracts, based on their ability to synthesize UDP-Ara*f* from UDP-Ara*p*. The reaction catalyzed by UAM is reversible with the equilibrium favoring the formation of UDP-Ara*p* in a 9: 1 ratio.[18]

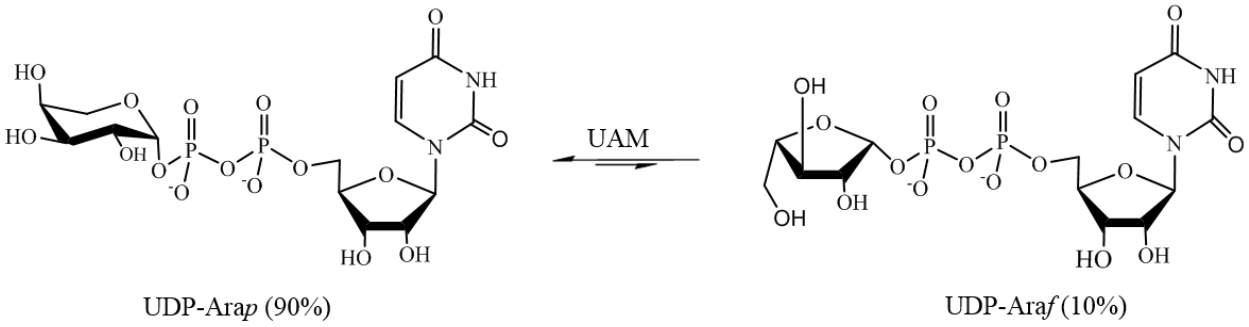


Figure 1-23: A schematic of the reaction catalyzed by UAM.

Although the reaction catalyzed by UAM is similar to the one catalyzed by UGM, these plant mutases neither bind FAD nor is it required for activity.[18] The purified *Oryza sativa* (*O. sativa*, rice) UAM (*Os*UAM) did not have the characteristic flavin UV-absorption peak at 450 nm. Furthermore, it was observed that *Os*UAM required divalent metal ions for activity. Although UAM was active on UDP-Gal*p*, the substrate of UGM, it was only a fraction of the activity compared to the activity obtained with UDP-Ara*p*. UAM was also identified in mung bean

extracts. The presence of UAM encoding genes in a variety of plant species ranging from green algae such as *Chlamydomonas reinhardtii*, mosses such as *Physcomitrella patens* to various dicots and monocots, indicates that UAM is widespread among plants and has a vital role in cell wall biosynthesis.[18]

Plant mutases belong to a small gene family called Reversibly Glycosylated Polypeptides (RGPs), meaning they can be reversibly auto glycosylated by UDP-sugars such as UDP-galactose, UDP-xylose, and UDP-glucose.[64,65] RGPs are highly conserved proteins, implicated in polysaccharide biosynthesis and defense responses and have been identified in the membrane and soluble fractions of plant species such as arabidopsis, pea, cotton, maize, potato, wheat, rice, and tomato.[66] Their identification in only plants (monocots and dicots) thus far shows that RGPs maybe plant-specific proteins. The three UAM genes that were identified from *O. sativa* are also RGPs; out of the three genes only two, *Os*UAM1 and *Os*UAM3, have the arabinofuranose-pyranose interconversion function while the third enzyme, *Os*UAM2 does not display this ability.[18] More recently, five RGP genes have been discovered in *Arabidopsis thaliana* (*A. thaliana*); three of these proteins, RGP1, RGP2 and RGP3 (*At*RGP1, *At*RGP2, *At*RGP3) have UAM function, i.e. catalyze the interconversion between UDP-Ara*f* and UDP-Ara*p*, while the other two, RGP4 and RGP5 (*At*RGP4 and *At*RGP5) do not.[66]

UAM activity is critical for the production and cell wall deposition of L-Ara*f* in plants. Gene knockout studies were used to demonstrate the importance of UAM activity in plant cells.[66] An analysis of cell wall monosaccharides of cell lines that had *AtRGP1* and *AtRGP2* genes knocked out, led to the observation of a 30 % reduction in total leaf cell wall arabinose content.[66] Moreover, in cell lines where the expression of these genes was suppressed, there was a massive reduction (~ 80 %) in the cell wall arabinose content and the UAM activity was reduced to ~ 1 %

of the wild-type.[66] A similar trend was also observed in cells involved in pollen development and seed coat secretion, thus proving the importance of UAM activity to the plant cell wall.

All three rice UAMs (*Os*UAMs) identified are ~ 39 - 41 kDa proteins. *Os*UAM1 and *Os*UAM3 share 88% sequence identity, while *Os*UAM2 shares only ~ 46% identity with *Os*UAM1 and *Os*UAM3.[18] Moreover, when treated with radiolabeled UDP-[^{14}C]-glucose, *Os*UAM1 and *Os*UAM3 were both auto glycosylated, but *Os*UAM2 lacked this ability. Also, the auto glycosylation is deemed reversible, since the [^{14}C]-glucose could be replaced with any among xylose, galactose or arabinose depending on which UDP-sugars reacted with UAM. If *Os*UAM was incubated with any among UDP-Glc, UDP-Xyl, or UDP-Gal for up to 6 hours, before the addition of UDP-Ara*f*, up to 50% reduction in mutase activity was observed.[18]

The molecular weight of UAM, prepared from rice seedling extracts, estimated by Konishi *et al*., (2007) using size-exclusion chromatography was ~ 460 kDa, indicating that its oligomeric state had to be at least a decamer. Rice UAM achieved optimal activity at a temperature of 55 ºC; maximum UDP-Ara*f* formation was observed between pH 7.0 and 7.5 while maximum UDP-Ara*p* formation was seen between pH 5.5 and 6.0. The enzyme had a higher affinity for UDP-Ara*f* (K_m = 55 µM) than *Ec*UGM (K_m = 600 µM).

The *At*RGPs had molecular weights in the range of 38 - 41 kDa.[66] *At*RGPs 1 and 2 overexpressed in growth tissues where cell wall components needed to be synthesized. Their importance in cell wall development was demonstrated when double knock-out mutants failed to develop cell wall polysaccharides and had severe growth retardation. Since *At*RGP1 and *At*RGP2 have the highest sequence identity, it was considered that their function was redundant. In their study, Rautengarten *et al*., (2011) discovered that knocking out either one of these two genes (*rgp1* or *rgp2*), led to significantly less arabinose synthesis. In plant cells, *At*RGP1 and *At*RGP2 are

localized in the cytoplasm and in Golgi compartments where they are required for cell wall polysaccharide biosynthesis. *At*RGP3 was found in cytosolic fractions. The ability to auto glycosylate was also observed in *At*RGPs that possess the mutase interconversion function.

1.9.2 GDP-6d-*altro*-heptopyrnaose Mutase

The gram-negative bacterium, *C. jejuni* is the leading cause of diarrhea, causes intestinal illness, gastroenteritis and is also linked to the neurological disorder Guillian–Barré syndrome and its variant, Miller Fisher syndrome, in humans.[19,67] Though the LPS layer is associated with the many diseases caused by *C. jejuni*, the more recently identified CPS has been considered important for virulence of the bacterium.[68,69] The CPS is composed of a variety of sugars in both pyranose and furanose forms. *C. jejuni* HS: 41 serotype, isolated from patients with Guillian–Barré syndrome, has a trisaccharide repeat unit in its CPS, composed of four different furanose residues. One of the furanoses in this repeating unit is 6d-D-*altro*-Hep*f*.[70] The CPS gene locus of *C. jejuni* HS: 41 serotype has homologs of the *glf* gene (the gene that encodes for UGM) which encodes other mutase enzymes that may synthesize the precursors required for the production of the trisaccharide unit. Three homologs of the *glf* gene, namely *glf1, glf2* and *glf3* were identified in this gene locus.[71]

The gene product of *glf1* is proposed to function as a GaHM by Dr. Todd Lowary's group (University of Alberta). The results of their experiments conclude that recombinant GaHM is an enzyme that synthesizes GDP-6d-D-*altro*-Hep*f*, the precursor of 6d-D-*altro*-Hep*f* (Figure 1-24). Although most mutases studied thus far use a UDP-based substrate for the interconversion reaction, it is believed that GaHM uses a GDP-based substrate and does not bind to a UDP-based substrate. This hypothesis was based on the identification of the *hddC* gene (D, D-heptose-1-

phosphate guanylyltransferase) in the CPS gene locus, known to be responsible for biosynthesis of another GDP-heptose sugar, in *C. jejuni* strain 11168.[71] Since GaHM is the gene product of a *glf* - like gene and a flavoenzyme, it is expected to have characteristics more similar to UGMs and UNGM, than to plant UAMs.

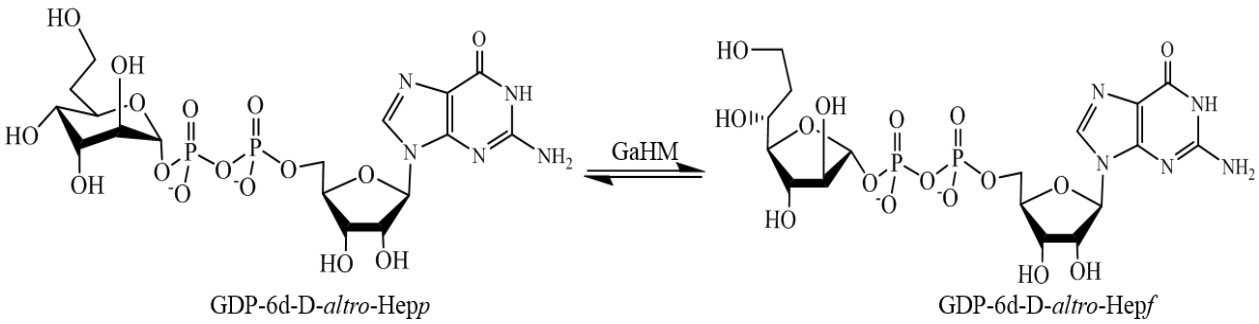


Figure 1-24: A schematic of the reaction catalyzed by GaHM.

1.10 Objectives of the Research

The three pyranose-furanose mutases described in the previous section (UGM, UAM, and GaHM) is the topic of this investigation. The over-arching theme of the research performed in this thesis is to understand the binding of the substrate to the active site of these enzymes and to understand how changes in the active site of the enzymes affect substrate binding. Among the three enzymes, UGM is the most studied while the other two enzymes are not as well understood, perhaps because they have only recently been identified. The structure of UGM (from both prokaryotic and eukaryotic organisms) and the structural basis for substrate recognition are known. However, apart from some of the active site residues, which have obvious importance in recognizing and binding different regions of the substrate, the structural role of other active site residues have not been studied in any of the known UGMs. The importance of these residues in maintaining the structure of the active site cavity in the correct conformation and the manner in

which they ensure the productive binding mode of the substrate are not understood thus far. To accomplish this goal, *Dr*UGM was chosen and six active site point mutants were created using SDM. Experiments were conducted on these mutants to gain input into how these mutations may affect the substrate binding in UGM.

*Os*UAM1 was first identified, purified and characterized by Konishi *et al.* (2007). There is minimal structural and mechanistic information available on UAM. It is a non-FAD-binding enzyme which uses divalent metal ions for catalysis. It was demonstrated that *Os*UAM1 and *At*RGPs were active when assayed in the presence of manganese (Mn^{2+}). UAM's low sequence identity with UGM and metal-ion dependency could mean that its catalytic mechanism and substrate binding mode may be completely different to the well-studied UGM. The aim of this study was to gain an understanding of the active site residues involved in metal co-factor and substrate binding. Divalent metal ion dependency of these plant mutases was also investigated to make a comparative analysis among the chosen UAMs. In the case of both *Os*UAM1 and *At*RGPs, no studies have yet been performed to show if there is a loss of activity in the absence of Mn^{2+} or other divalent ions. No information is available about the binding constants of the metal co-factor and how the metal binds and interacts with any of these enzymes. Other divalent metals were tried in the case of *Os*UAM but not *At*RGPs. Obtaining such information will help to understand better the role the metal plays in enzyme function.

The next pyranose-furanose mutase studied was *Cj*GaHM, which requires the FAD cofactor for activity. A crystal structure of this enzyme, co-crystallized with GDP (solved by previous member of the Sanders lab) was already available [unpublished results]. However, GDP is the nucleotide base for the substrate; the goal here was to obtain structural information with a GDP-sugar and make a comparative analysis with UGM to understand how the active site of

GaHM can accommodate a GDP-based substrate rather than a UDP-based substrate (as seen in UGM and UAM).

Chapter 2: Materials and methods

2.1 Cloning of RGPs from *Arabidopsis Thaliana*

*At*RGP1 (Gene ID: 1523286, Accession no: NP_186872), *At*RGP2 (Gene ID: 15242351, Accession no: NP_197069) and *At*RGP3 (Gene ID: 30680679, Accession no: NP_187502) were identified as UAMs in *A. thaliana*.[66] Over-expression clones of these three proteins were made using Gateway cloning technology.[72] Genes encoding each of the RGPs were synthesized and obtained in a pET29b plasmid. The genes were amplified using the polymerase chain reaction (PCR) using gene-specific primers shown in Table 2-1.

Table 2-1: Primers for cloning of *At*RGPs.

RGP1 Primer N1	5'-GAGAACCTGTACTTCCAGGGTGGTGGTATGGTGGAACCGGC – 3'
RGP1 Primer C	5'-GGGGACCACTTTGTACAAGAAAGCTGGGTTATTAAGCTTTCGTCGGCGG - 3'
RGP2 Primer N1	5'- GAGAACCTGTACTTCCAGGGTGGTGGTATGGTGGAACCGGCG - 3'
RGP2 Primer C	5'- GGGGACCACTTTGTACAAGAAAGCTGGGTTATTAGGCTTTACCGCTGGC-3'
RGP3 Primer N1	5'- GAGAACCTGTACTTCCAGGGTGGTGGTATGGCTCAACTGTAC - 3'
RGP3 Primer C	5'- GGGGACCACTTTGTACAAGAAAGCTGGGTTATTAATTTTTACCCTTCGG -3'
Primer N2	5'- GGGGACAAGTTTGTACAAAAAAGCAGGCTCGGAGAACCTGTACTTCCAG-3'

Primer N1 has a tobacco etch virus protease (TEV protease) recognition site on the 5' end, and Primer C has an *att*B2 recombination site on the 5'-end. PCR reactions were set up using the procedure, shown in the Table 2-2. PCR products obtained with the TEV-protease recognition site and the *att*B2 recombination site on the N and C-termini of the RGP gene, respectively, were gel-purified (Gel extraction kit, Qiagen). A second PCR amplification was performed using this gel-purified product as the template and the generic primer N2, which has a TEV-protease recognition

site at the 3'-end and primer C, specific for each *RGP*. At the end of this PCR, the *AtRGP* genes obtained contained an *att*B1 recombination site and a TEV-protease recognition site at the 5'-end and an *att*B2 recombination site at the 3'-end, as shown in Figure 2-1.

Table 2-2: PCR reaction conditions for cloning of *At*RGPs.

Process	Temperature (°C)	Time	Cycles
Denaturation	95	5 min	1
Annealing	55	30 sec	1
Denaturation	95	45 sec	40
Annealing	55	30 sec	
Elongation	72	2 min	
Elongation	72	10 min	1
Cooling	4	1 hour	

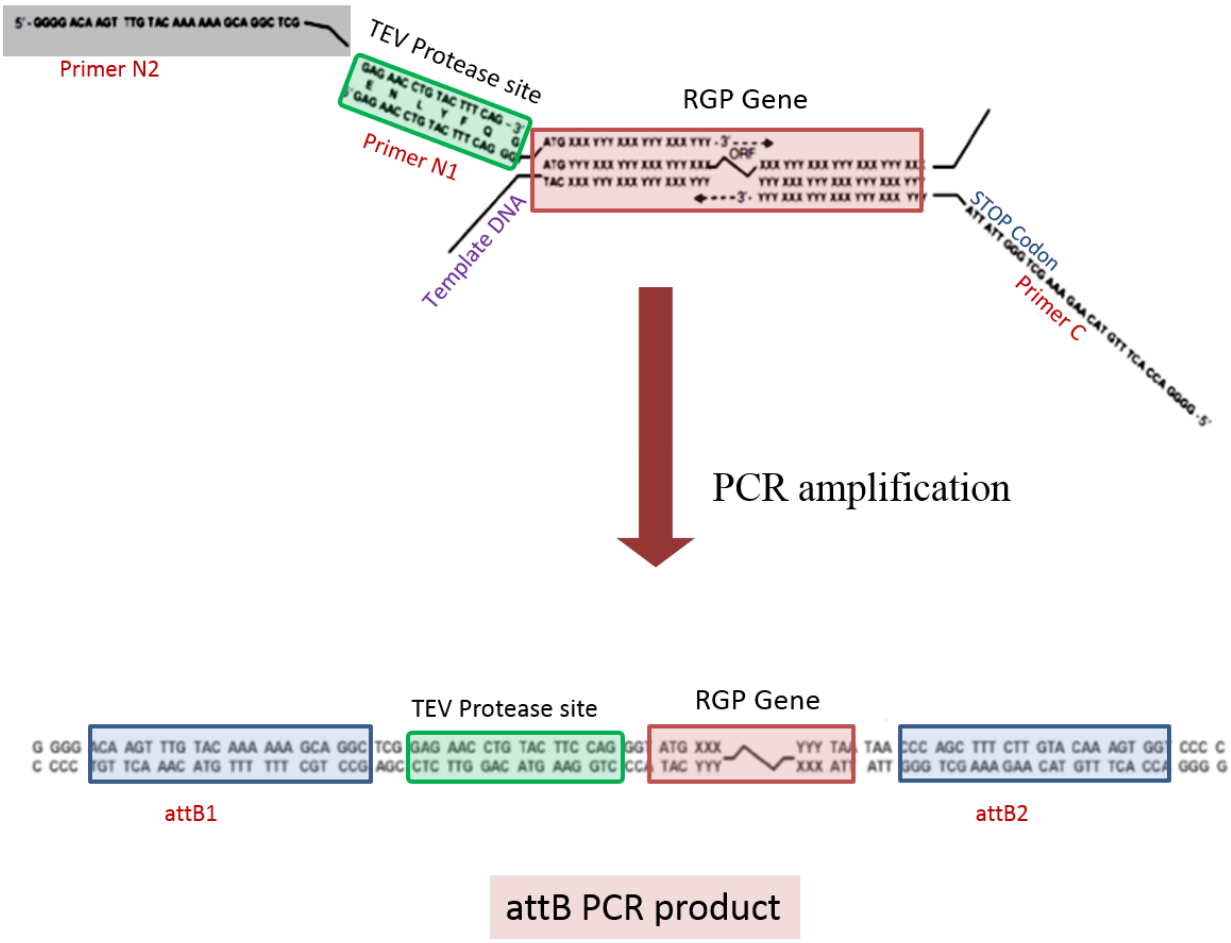


Figure 2-1: A scheme for obtaining PCR amplified RGP gene products.
The RGP gene with Primer N1, Primer N2 and Primer C is shown. The Primer N1 has nucleotide sequence corresponding to TEV protease recognition site on its 5'-end. The Primer N2 has nucleotide sequence corresponding to TEV protease recognition site on its 3'-end. The Primer C has a stop codon on its 3'-end. The RGP gene with the primers are PCR amplified to obtain a product with attB1 and attB2 (attachment) sites on either ends of the RGP gene.

BP and LR reactions were performed sequentially, with this final PCR-product, to clone the *AtRGPs* into the destination vector. The PCR-product was initially cloned into an entry clone vector, pDONR-221. Both were added to a 10 µl BP reaction mixture, in the presence of the enzyme BP Clonase-11 (Invitrogen) and the reaction was allowed to proceed for 3-4 hours. The reaction products were transformed into *E. coli* DH5α cells (Invitrogen) and plated onto LB agar plates with kanamycin and grown overnight at 37 °C. A few colonies were picked and grown overnight, and plasmids were isolated (Qiagen mini-prep kit). Plasmids which contained the desired *RGP* genes, as verified by sequencing, were used for the LR reaction, which transfers the *RGP* genes from the entry clones to a destination vector. A LR reaction (10 µl) was set up by adding the pDONR-221 containing the *RGP* genes into the destination vector, pDEST-HisMBP, in the presence of LR Clonase-11 (Invitrogen). After 3 - 4 hours, the reaction product obtained was transformed into *E. coli* DH5α cells (Invitrogen), plated onto LB agar plates with ampicillin and grown overnight at 37 °C. A few of the grown colonies were picked, their plasmids isolated (Qiagen mini-prep kit) and sequenced (Applied Genomics Center, National Research Council, Saskatoon, Canada). Plasmids with the desired genes were then transformed into *E. coli* BL21 Gold over-expression cells (Stratagene).

2.2 Cloning of *Os*UAM1

The *OsUAM1* gene (Gene ID: 75153246, Accession no: Q8H8T0) was cloned into a pHISTEV vector that encodes for proteins with a 6xHistidine tag on its N-terminal. Additionally, a TEV-protease recognition site is present in between the His-tag and the protein sequence for efficient cleaving of the His-tag. After confirmation of the required sequence, the pHISTEV plasmid with the *OsUAM1* gene was transformed into *E. coli* BL21 Gold cells for over-expression.

2.2.1 Site-directed mutagenesis of *Os*UAM1 mutants

The primers used to create *Os*UAM1 mutants D110A, D111A, D112A and H273A were designed using PrimerX and synthesized by AlphaDNA (AlphaDNA Inc.). PCR was set up using ~50 ng of wild-type *Os*UAM1-plasmid DNA as template, ~ 125 ng of corresponding forward and reverse primers (shown in Table 2-3), dNTPs, KAPA HiFi DNA polymerase (KAPA Biosystems) and the reaction buffer, using the PCR reaction condition shown in Table 2-4.

Table 2-3: Primers for cloning of *Os*UAM1 mutants.

Mutants	Primers
H273A Forward	5'-CGTACATCTGGGCCAGCAAGGCTAG -3'
H273A Reverse	5'-CTAGCCTTGCTGGCCCAGATGTACG-3'
D110A Forward	5'-GTACGTCTTCACCATCGCCGACGACTGCTTCGTTG -3'
D110A Reverse	5'-CAACGAAGCAGTCGTCGGCGATGGTGAAGACGTAC -3'
D111A Forward	5'-CTTCACCATCGACGCCGACTGCTTCGTTGCC -3'
D110A Reverse	5'-GCAACGAAGCAGTCGGCGTCGATGGTGAAGG -3'
D112A Forward	5'-CACCATCGACGACGCCTGCTTCGTTGCC-3'
D110A Reverse	5'-GGCAACGAAGCAGGCGTCGTCGATGGTG-3'

The reaction product was digested with *Dpn1* at 37 °C for one hour, transformed into *E. coli* DH5α cells and plated on LB-agar plates with kanamycin and incubated at 37 °C overnight. A few colonies were selected and inoculated into 5 ml Luria-Bertani (LB) broth with kanamycin and grown overnight. The plasmid DNA was isolated (Qiagen mini-prep kit) and sent for sequencing, to confirm the desired mutations. The plasmid DNA was then transformed into *E. coli* BL21 over-expression cells.

Table 2-4: PCR reaction condition for *Os*UAM1 mutants.

Process	Temperature (°C)	Time	Cycles
Denaturation	95	5 min	1
Annealing	55	30 sec	1
Denaturation Annealing Elongation	95 55 72	45 sec 30 sec 2 min	18
Elongation	72	10 min	1
Cooling	4	1 hour	

2.3 Over-expression and purification of mutases

2.3.1 Over-expression and purification of *Os*UAM1

BL21 Gold cells containing the plasmids encoding for the desired *Os*UAM1 was grown overnight at 37 °C in 100 ml LB media with 100 µg/ml kanamycin. 10 ml of this culture was subcultured into 1L cell culture flasks and left to grow at 37 °C to an OD_{600} of 0.6. 500 µM isopropyl-ß-thiogalactoside (IPTG) was added for protein over-expression, and the cells were allowed to grow overnight at 15 °C. The next day, cells were harvested by centrifugation at 4 °C, 3600 rpm and stored in -80 °C. Frozen cell pellets were re-suspended in lysis buffer, containing 50 mM sodium phosphate pH 8.0, 50 mM NaCl, 10 µg/ml DNase, 20 µg/ml lysozyme, 0.1 % Triton-X and 0.1 % ß-mercaptoethanol, and stirred at 4 °C for 30 minutes. This was followed by sonication for 3 minutes, (15 sec pulse on/off), to break open the cell wall, and centrifugation at 15000 rpm for 30 minutes at 4 °C, to obtain soluble protein in the supernatant and to remove unwanted cell debris.

Purification was performed using nickel affinity chromatography followed by size exclusion chromatography. The supernatant was loaded onto a Protino Ni-IDA column (Macherey-Nagel), previously charged with Ni^{2+} and equilibrated with 50 mM sodium phosphate pH 8.0 and 50 mM NaCl. The column was washed with 50 mM sodium phosphate pH 8.0, 50 mM NaCl containing 30 mM imidazole to remove protein impurities. *Os*UAM1 was then eluted with 50 mM sodium phosphate pH 8.0, 50 mM NaCl and 250 mM imidazole. *Os*UAM1 fractions were dialyzed against 25 mM Bicine pH 8.5 and concentrated to ~ 5 mg/ml before being loaded onto a Gel-filtration column (GE Healthcare), equilibrated with 25 mM Bicine pH 8.5. The enzyme eluted as a big first peak followed by other contaminating proteins. Purified UAM fractions were pooled and dialyzed against 25 mM Bicine pH 8.5 with 1 mM Manganese chloride and concentrated to ~ 6.5 mg/ml, for crystal trials.

A stock UAM concentrated to ~ 1 mg/ml was preferred for kinetic and metal binding assays. For kinetics, the purified protein was treated with a metal chelating agent, 200 mM ethylenediaminetetraacetic acid (EDTA), for 3 hours and dialyzed against 25 mM Bicine pH 8.5 (four exchanges) prior to flash freezing with liquid nitrogen. This was done to remove any metal that might affect the UDP-Ara*f* to UDP-Ara*p* conversion during metal binding assay experiments.

2.3.2 Overexpression and purification of *Os*UAM1 mutants

Four different mutants of *Os*UAM1; D110A, D111A, D112A, and H273A, were all over-expressed and purified as described above and a final stock concentration of ~ 1 mg/ml was prepared for each, to use for kinetic assays.

2.3.3 Over-expression and purification of *At*RGPs

All three *At*RGPs (RGP1, RGP2, and RGP3) were over-expressed and purified following a similar procedure. *RGP* genes cloned into pDESTHis-MBP over-expression vector, expressed with a 6x Histidine-tagged Maltose binding protein (MBP) at the N-terminal. Cells expressing these proteins were grown overnight in 100 ml LB media with 100 µg/ml ampicillin at 37 °C. The next day, 10 ml of the overnight culture was subcultured into 1L of autoclaved LB media containing 100 µg/ml ampicillin. The cells were allowed to grow at 37 °C until an OD_{600} of 0.5 was achieved and then protein over-expression was induced by the addition of 500 µM (final concentration) of IPTG and set to grow overnight at 15 °C. The cells were harvested the next day by centrifugation at 4 °C and 3500 rpm for 20 minutes and stored at -80 °C until further use. Cell lysis was done by resuspending the cells in 50 mM potassium phosphate pH 8, 50 mM NaCl, with 20 µg/ml lysozyme and 10 µg/ml DNase for 30 minutes followed by sonication for 3 minutes, (15 sec pulse on/off). The solution was then clarified of cell debris by centrifugation at 15000 rpm for 30 minutes at 4 °C.

The supernatant was loaded onto a Ni-sepharose column (GE-Healthcare), pre-equilibrated with buffer containing 50 mM potassium phosphate pH 8 and 50 mM NaCl. Bound His_6-MBP-RGPs were eluted out of the column with Elution buffer (50 mM potassium phosphate pH 8, 50 mM NaCl and 250 mM Imidazole), after the impurities were removed by washing the column with 50 mM potassium phosphate pH 8, 50 mM NaCl and 30 mM Imidazole. The His_6-MBP-RGP fractions were pooled together, dialyzed against 50 mM potassium phosphate pH 8, 50 mM NaCl, to remove imidazole, and concentrated to ~ 30 mg/ml, for setting up a tag-cleavage reaction with TEV protease. To cleave the His_6-MBP-tag, 2 mg TEV protease was added to 10 mg purified RGPs (a 1:5 ratio), and the reaction was left to shake at 4 °C overnight. The TEV protease digested

sample was then loaded onto a Ni-sepharose column (GE-Healthcare) pre-equilibrated with 25 mM Bicine pH 8.5. The required RGP proteins (MBP-tag less) were eluted in the wash with the same buffer and the 6xHis-MBP-tag (which is still bound to the column due to the 6xHis tag), was eluted with 25 mM Bicine pH 8.5, 250 mM imidazole. The RGP fractions were pooled and concentrated to ~ 10 mg/ml for crystal trials. For metal-binding assays, a procedure similar to that performed on *Os*UAM1 was adopted; all *At*RGPs were treated with EDTA to remove bound metal, and the enzymes were concentrated to ~ 1 mg/ml in 25 mM Bicine pH 8.5.

2.3.4 Over-expression and purification of *Dr*UGM wild-type and mutants

The gene encoding wild-type *Dr*UGM was cloned into a pHISTEV vector and was transformed into *E.coli* Tuner cells (Novagen, USA).[24] The cells were grown overnight at 37 °C in a 100 ml LB culture with 100 µg/ml kanamycin. 10 ml of this overnight culture was subcultured into 1L culture flasks and grown until an OD_{600} of 0.6 was reached. Protein over-expression was induced with 0.4 mM IPTG, and the cells were cultured at 30 °C for 4-5 hours. The cells were stored in -80 °C after harvesting by centrifugation at 4000 rpm and 4 ºC. The frozen cell pellets were re-suspended in lysis buffer; 100 mM potassium phosphate pH 8.0, DNase and lysozyme at 20 µg/ml, 1 mM 4-(2-Aminoethyl) benzene sulfonyl fluoride hydrochloride (AEBSF), 0.1 % Triton-X and stirred at 4 °C for 30 minutes. The lysed cells were sonicated for 2 min (10 sec on/off pulse) and centrifuged at 15000 rpm for 30 minutes at 4 °C. The supernatant was heat treated for 10 min at 55 °C, to precipitate contaminating proteins. The suspension was then centrifuged again at 15000 rpm for 30 minutes at 4 °C. The sample was dialyzed against 25 mM potassium phosphate pH 8.0. After four buffer exchanges, the dialysate was loaded onto an HQ20 anion exchange column (Applied Biosystems, USA) pre-equilibrated with 25 mM potassium phosphate pH 8.0.

The protein was eluted with 25 mM potassium phosphate pH 8.0 buffer, with a gradient from 0 M and 1 M NaCl. Fractions containing the enzymes were pooled together and dialyzed against 50 mM potassium phosphate pH 8.0. Required amount of solid ammonium sulphate was added to the dialyzed protein to make it up to 30 % (w/v) ammonium sulphate. The sample was loaded onto an HP-20 hydrophobic interaction column (Applied Biosystems, USA) pre-equilibrated with 50 mM potassium phosphate pH 8.0, 30% (w/v) ammonium sulphate. The protein was eluted with 50 mM potassium phosphate pH 8.0 using a gradient of 30 - 0% (w/v) ammonium sulphate. The eluted protein fractions were dialyzed against 50 mM Bis-tris propane pH 8.0 and concentrated to ~5 mg/ml, for the next stage. A third purification step was employed, using a gel-filtration column (GE Healthcare), pre-equilibrated with 50 mM Bis-tris propane pH 8.0. The protein was loaded onto the column, and the enzyme fractions were collected and concentrated to ~ 7.5 mg/ml for use in crystal trials. Gel electrophoresis was used to analyze the purity of the sample. The protein was flash-frozen in liquid nitrogen and stored as 50 µl aliquots at -80 °C. Stocks were also made to ~ 1 mg/ml, for use in kinetic assays. All the *Dr*UGM point mutants were also prepared by following the same procedure.

2.3.5 Over-expression and purification of *Mt*UGM

The gene encoding for *Mt*UGM was cloned into pDEST-HisMBP vector, using Gateway cloning technique that was described earlier.[72] The plasmid was transformed into BL21 Gold over-expression cells and confirmed by sequencing (Applied Genomics Center, National Research Council, Saskatoon, Canada). A single colony of cells was grown overnight in LB media at 37 °C with 100 µg/ml ampicillin. The cells were subcultured into 2 ×1 L LB media cultures, supplemented with 100 µg/ml ampicillin and grown until an OD_{600} of 0.4 - 0.5 was reached. The

cells were, induced with 1 mM IPTG (final concentration) and grown overnight at 15 °C. The cells were harvested by centrifugation at 3500 rpm and 4 °C for 20 minutes. The cell pellets were re-suspended in lysis buffer (25 mM Tris pH 7.5, 150 mM sodium chloride and 20 µg/ml of DNase) for 30 minutes. Sonication (3 min, 15 sec on/off pulse), and centrifugation (15000 rpm, 30 min at 4 °C) were performed on the re-suspended cells.

The supernatant, which was filtered using a 0.2 µM filter, was loaded onto a GE healthcare Dextrin Sepharose column, pre-equilibrated with 25 mM Tris pH 7.5, 150 mM NaCl. The column was washed with the same buffer, to remove all unbound proteins, and the *Mt*UGM with MBP-tag was eluted from the column using 25 mM Tris pH 7.5, 150 mM NaCl and 50 mM maltose. The eluted protein was dialyzed against 25 mM Tris pH 7.5, 150 mM NaCl to remove maltose and concentrated to ~ 5.5 mg/ml. To cleave the MBP-tag, the concentrated sample was then treated with TEV protease in a 1:10 ratio and the reaction was left overnight at 4 °C. The reaction mixture was further diluted 10 times, using 25 mM Tris pH 7.5, 150 mM NaCl and loaded onto a Ni^{2+} Sepharose affinity column (GE-Healthcare), pre-equilibrated with the same buffer. The cleaved *Mt*UGM was collected in the flow through. The MBP and TEV-protease bound to the column were then eluted with 25 mM Tris pH 7.5, 150 mM NaCl and 250 mM imidazole. The purified *Mt*UGM was dialyzed against 25 mM Tris pH 7.5, 150 mM NaCl and concentrated to ~ 1 mg/ml and use in kinetic assays and inhibition studies.

2.4 HPLC-based assay for mutases

2.4.1 Kinetic assays for *Dr*UGM mutants

The kinetic assay for the *Dr*UGM mutants was performed using the following procedure.[51] A known concentration of the protein, obtained from purification experiments, was used to react with a known concentration of substrate in a reaction buffer and the product was analyzed using a High-performance liquid chromatography (HPLC) (Waters HPLC system, with a Waters 510 pump connected to Waters 717 plus Autosampler and Waters 2487 Dual λ Absorbance Detector). The reaction buffer (500 mM sodium phosphate pH 7.0), double-distilled water (Millipore) and the reaction vial were argon-degassed. The reactions were performed by adding the *Dr*UGM mutant enzyme (at the desired concentration) into a vial containing 50 mM sodium phosphate pH 7.0. Final enzyme concentrations in the reaction mixture for each mutant was the minimal concentration required to make a complete saturation curve by varying concentrations of the substrate. The substrate used was UDP-Gal*f*, and the reaction was monitored for the formation of UDP-Gal*p*. The final enzyme concentrations used for the assay are as follows; wild-type - 10 nM, N372D - 50 nM, W184A - 50 nM, W184F - 50 nM, H88F - 50 nM, R364A - 50 nM and R364K - 50 nM. The enzyme added to each reaction mixture was reduced with the addition of 20 mM sodium dithionite (final concentration) and left to reduce for 30 sec. The substrate (UDP-Gal*f*) was added into the vial next to commence the reaction. Typically the substrate concentration was varied between 10 µM to 1.5 mM to obtain a saturation curve. The reaction was quenched by the addition of n-butanol at a reaction time that resulted in 30 - 40 % conversion of UDP-Gal*f* to UDP-Gal*p*. The aqueous layer was collected and analyzed on a CarboPac PA1 (Dionex Inc) column, previously equilibrated with filtered and degassed 200 mM ammonium acetate pH 7.0. The formation of product was monitored on a UV-spectrophotometer at 262 nm. A typical HPLC-run

shows two prominent peaks, one corresponding to the substrate (UDP-Gal*f*) and the other corresponding to the product (UDP-Gal*p*). The area under the peaks was calculated using the Millennium software (Version 4.0) and the % conversion of substrate was calculated as follows,

$$\% \text{ Conversion} = \frac{\text{Area under UDP-Gal}p \text{ peak}}{[\text{Area under UDP-Gal}f \text{ peak} + \text{Area under UDP-Gal}p \text{ peak}]} \times 100\% \quad \textbf{(1)}$$

Using the % conversion value (calculated by Equation 1) obtained at each substrate concentration; the rate of product formation per second was calculated and plotted against substrate concentration, to obtain kinetic curves for each of the mutants. As substrate available was limited, reactions at each different substrate concentration were performed in duplicates and the data was plotted using GraphPad Prism software (GraphPad Software, San Diego, CA).

2.4.2 Kinetic assays for *Mt*UGM

Kinetics experiments for *Mt*UGM wild-type enzyme were conducted in a similar manner to the method described 2.4.1, with minor changes in the reaction condition. All reactions for the assay were performed at 37 °C; a final enzyme concentration of 100 nM and a substrate range between 10 - 150 µM were used, to make a saturation curve.

2.4.3 Inhibition assays for *Mt*UGM by MS-208

For performing inhibition assays on *Mt*UGM, after the enzyme was reduced with the addition of sodium dithionite, the inhibitor MS-208 (stock made at 1mM concentration) was added to the reaction mixture at the required concentration and left to incubate at 37 °C for 1 minute. The reaction was then carried out with the addition of the substrate, quenched with n-butanol and analyzed on HPLC as performed for the other assays (Section 2.4.1). Three different

concentrations of the inhibitor, 60 µM, 120 µM and 200 µM, were used to obtain kinetic curves. The plots of rate vs. substrate concentration were prepared by GraphPad Prism software (GraphPad Software, San Diego, CA); SigmaPlot software (SigmaPlot 12.0) was used for global fitting of the data obtained for each saturation curve; Dixon and Cornish-Bowden plots were drawn using MS Excel, to aid in determining the mode of inhibition.

2.4.4 Kinetic assays for *Os*UAM1 and mutants

The kinetic assays were performed adopting the procedure described earlier for *Dr*UGM mutants, with minor changes.[18] The *Os*UAM1 enzymatic assay was carried out by incubating the enzyme in 20 mM HEPES buffer pH 6.8, 5 mM $MnCl_2$ at 37 °C for 2 minutes. As with *Dr*UGM, the enzyme concentration was chosen so that a complete kinetic curve could be obtained. However, no reducing agents were added to the reaction mixture, as UAM does not require to be reduced for activity. The reaction was initiated with the addition of the desired concentration of substrate, UDP-Ara*f*, and the reaction was left to proceed for the necessary amount of time that gave ~ 30 - 40 % product conversion. The reaction was quenched with the addition of n-butanol, and the aqueous layer was analyzed on a CarboPac PA1 Dionex column, previously equilibrated with 210 mM ammonium acetate pH 6.0, with the flow rate of 1 ml/min. Formation of the product, UDP-Ara*p*, was confirmed by monitoring UV vis-spec at 262 nm. The % conversion was calculated using equation (1) and the rate vs. UDP-Ara*f* concentration was plotted using GraphPad Prism software (GraphPad Software, San Diego, CA).

2.4.5 Metal binding assays for *Os*UAM1 and RGPs

Samples of *Os*UAM1 wild-type, H273A and all three *At*RGPs, pre-treated with 250 mM EDTA, were used for metal binding assays. For all studies performed at varying metal concentrations, a final enzyme concentration of 20 nM and final substrate concentration of 10 µM were used and the reactions were conducted for 30 seconds. To obtain the Mn^{2+} metal curve, concentrations of $MnCl_2$ were varied between 10 µM to 5 mM. Similarly solutions of $ZnCl_2$, $CoCl_2$, $CuCl_2$, $MgCl_2$, and $CaCl_2$ were used to obtain curves for the other divalent metal ions, Zn^{2+}, Co^{2+}, Cu^{2+}, Mg^{2+} and Ca^{2+} respectively. Each of the reactions was performed in duplicates, and the % conversion values were obtained. The maximum value of % conversion obtained at a particular concentration of Mn^{2+} metal was chosen as 100 % relative activity, and all other values were calculated relative to this value for metal-binding assay plots. A similar strategy was followed for all other divalent metals.

2.5 Crystal trials for mutases

2.5.1 *Dr*UGM mutants and wild-type enzyme

Microbatch under oil method was used to perform crystal trials on *Dr*UGM mutants and wild-type enzyme.[73,74] Since wild-type *Dr*UGM and *Dr*UGM N372D were previously co-crystalized with the substrate UDP-Gal*p*, crystal trials were focused on obtaining co-crystals of the five other point mutants with the substrate and the holoenzyme for wild-type *Dr*UGM. The concentration of the point mutants used for all the trials was kept at ~ 7.5 mg/ml. The enzyme was first reduced by the addition of 20 mM sodium dithionite, the reducing agent, and 15 mM final concentration of substrate was added to the sample. To set up crystallization drops, the protein was pipetted into wells in a micro batch plate, and the crystallization solution was added in a 1:1 ratio.

The drops were covered with paraffin oil quickly and set at room temperature. Commercial screen kits (Qiagen), each containing 96 different conditions was used for setting up the micro batch plates. Initial crystal hits were obtained across conditions in many screens including, Classics II suite, PEGs I and II suites and the PACT suite. A similar approach was adopted for trials carried out for the wild-type enzyme without the addition of reducing agent or the substrate. The crystal hits, in this case, were obtained with conditions from the Classics II suite. For optimization of the hits, grid-screens were set-up around these positive conditions. Grid screens were manually prepared by varying the final concentrations of the various components in the crystallization solution. In the case of the holoenzyme, additive screens (Qiagen) were also used for further optimization.

Rod-shaped bright yellow crystals were obtained in a week and were let to grow for two weeks before harvesting and flash-freezing. 20 % PEG 400 with 80 % crystallization solution was used as a cryoprotectant solution for flash-freezing *Dr*UGM crystals in liquid nitrogen.

2.5.2 GaHM crystals with substrate mimics

The set-up for obtaining GaHM crystals was similar to that described for *Dr*UGM.[73,74] GaHM was co-crystallized with two different substrate mimics, GDP-mannose, and GDP-glucose. Microbatch under oil technique was used, and the protein solution was reduced with sodium dithionite (turning the protein colorless), prior to the addition of the substrate mimic, GDP-mannose. The concentrations of GaHM protein, sodium dithionite, and GDP-mannose were ~ 7.5 mg/ml, 20 mM and 15 mM respectively. The drops were set-up at 4 °C, in a 1:1 ratio of protein to crystallization solutions, and covered with paraffin oil. A few conditions in Classics II, PACT and PEG1 screens gave initial hits. A condition from Classics II screen containing, 0.2 M Ammonium

sulphate, 0.1 M Bis-Tris pH 6.5 and 25% (w/v) PEG 3350, produced a cluster of needle-like yellow crystals that appeared within a week. Grid-screens were set around this condition, and further optimization was done by using the additive screens. Crystals were obtained with both the substrate mimics and were left to grow for two to three weeks before flash freezing using liquid nitrogen.

Two different types of cryoprotectant solutions were made for flash freezing GaHM crystals.[74] For FAD_{ox} crystals, 20 % (v/v) of ethylene glycol and 80 % crystallization solution was used, while $FAD_{red,}$ crystals were initially re-reduced with the addition of sodium dithionite (final concentration in 20 mM) along with 15 mM GDP-mannose in a cryoprotectant solution made of 20 % ethylene glycol and 80 % crystallization solution. The crystal was then looped out of the microbatch plate wells and washed thrice in the cryoprotectant solution, to remove mother liquor, and flash-frozen rapidly.

2.5.3 Crystal trials on *Os*UAM1 and *At*RGPs

Microbatch under oil and vapour diffusion techniques were both used for crystal trials of *Os*UAM1 and *At*RGPs. Crystallization experiments were conducted to co-crystallize these enzymes with UDP-Gal*p*, since the actual substrate, UDP-Ara*p*, is not commercially available. All the enzymes were dialyzed against a buffer containing, 25 mM Bicine pH 8.5, 2 mM manganese chloride and therefore, required only the addition of UDP-Gal*p* to a final concentration of 15 mM before setting up the protein drops using crystal screens. Microbatch plates were set up in a 1:1 ratio, of protein solution and crystallization solution, using commercially available screens (Qiagen). A number of grid-screens were set up around positive hits, and additive screens were also used, to see if any of these conditions could be optimized, to further improve the quality of crystals. *Os*UAM1, without the His-tag, was also prepared and used for setting up trials to see if

the absence of the tag would help in obtaining better quality crystals. Trials were performed at 4 °C, 15 °C and room temperature.

Crystal trials, in the case of *At*RGPs, were carried out in a similar fashion to that described for *Os*UAM1. Conditions giving microcrystalline materials and spherulites of *At*RGP1 were further optimized. Grid-screens for optimization were set up using vapour diffusion - Hanging drop technique, where the crystallization drop is set to hang over a reservoir containing the crystallization solution. The drop was prepared by mixing the protein solution and reservoir solution in a 1: 1 ratio set on a thin plate, which is then sealed over the reservoir. The crystalline material that appeared within two days in some of the conditions was looped out, washed with reservoir solution and crushed using a seed bead, to be utilized as macro-seed-stock for further experiments. The seed-stock was further diluted down to different ratios 1: 10, 1:100 and 1:1000 and used to set up more hanging drop plates. The crystallization drop, which now contains protein solution, reservoir solution and seed solution in 1.5:1:0.5 ratio was sealed over the reservoir solution. Other drop ratios were also tried to see if the quality of the crystals could be improved.

2.6 Data collection, processing, and refinement of GaHM crystal structures

GaHM crystals, co-crystallized with substrate mimics $GaHM_{ox}$: GDP-mannose, and $GaHM_{red}$: GDP-mannose were diffracted at the Canadian Light Source (CLS), Saskatoon, Canada. Datasets were collected on the 08B1-1 beamline, equipped with MD2 micro diffractometer and a Rayonix MX300HE X-ray detector. The crystal to diffractometer distance was set at 280 mm and 600 images were collected for each dataset with an oscillation of 0.25° and exposure time of 1 sec. The data processing and scaling were performed using the Autoprocess pipeline.[75,76]

GaHM crystal structures were solved using the Molecular replacement method with MOLREP program within the CCP4 package.[77-79] GaHM models available from crystal structures of GaHM solved with GDP previously solved by Dr. Sean Darlymple [unpublished results], were used for finding a structure solution. All the solutions had 2 copies of the GaHM monomer in an asymmetric unit (ASU).[80]

PHENIX was used to refine the structures, using the output model PDB from the CCP4 package and the corresponding .mtz files for each structure.[81,82] Initially, Rigid-body refinement was performed to refine the position of the monomers. Non-crystallographic symmetry (NCS) restraints were turned on for each refinement run. To remove model bias, simulated annealing was performed using Cartesian dynamics at 5000 K. The output models, after each refinement run, were re-built in COOT.[83] The refinement and rebuilding of each model were performed iteratively until satisfactory progress was made with R-work/R-free. Another round of simulated annealing was done using Cartesian dynamics; this time, at a lower starting temperature of 2500 K. The geometry restraints information of oxidized FAD, reduced FADH and the corresponding ligands were generated with eLBOW available within the PHENIX software.[84] The refinement was re-run after including the PDB coordinates and restraint files for the ligands. The electron density maps were examined after each run to manually fix residues and regions that were inconsistent. Water molecules were added into the refined model using 'Update water' option in PHENIX when the R-free value was lower than 30%. The waters were verified once again before a final round of refinement was completed. All the figures were made using PYMOL (PyMOL version 1.7.4, DeLano Scientific LLC)

2.7 Modeling studies

2.7.1 Modeling of *Dr*UGM mutants

*Dr*UGM mutants were modeled using an online software Rosetta-Backrub.[85,86] For modeling experiments, *Dr*UGM wild-type crystal structure, solved with UDP-Gal*p* (PDB id: 3HDY), with FAD_{red} was the input into the software. The residue to be mutated was chosen from the sequence (by the residue number), and the desired mutation was submitted to the program. Up to twenty different models were made in each run; the model with the best score was chosen and used for further docking studies.

2.7.2 *Os*UAM1 Modeling

GalaxyWEB and Iterative threading assembly refinement (I-TASSER) were used to generate 3D-models of *Os*UAM1.[87,88] Both these programs generate models based on the input protein sequence. Five different models obtained, from both GalaxyWEB and I-TASSER, were manually compared based on the residues predicted to form secondary structure regions using HHpred. This was done to check the consistency of the predicted models. A comparison of all the models was performed using PyMOL software (PyMOL version 1.7.4, DeLano Scientific LLC).

2.7.3 Modeling of loop regions of *Cj*GaHM

The crystal structure of GaHM solved with GDP-mannose has little or no electron density for the flexible loop regions. Since the crystal structures of both *Af*UGM and *Tc*UGM were overall similar to that of *Cj*GaHM, they were used as templates to model the missing GaHM loop regions. The missing loop regions were built into the structure using COOT, based on the *Cj*GaHM sequence.

2.8 Genetic Optimization for Ligand Docking docking for *Dr*UGM mutants

Docking of the substrate (UDP-Gal*p*) into the active site of the *Dr*UGM mutants models (generated from Rosetta-Backrub program) was performed using Genetic Optimization for Ligand Docking (GOLD) software (version 5.2.2).[89-92] For performing the docking runs, protein .pdb files and substrate .mol2 files were used. A 6 Å binding site was defined using the UDP-Gal*p* binding site from the *Dr*UGM: UDP-Gal*p* crystal structure, as the reference ligand. This strategy also generated root-mean-square-deviation (RMSD) numbers that helped determine the average deviation of all the docked poses generated for a certain .pdb file from the ensemble. Each run was allowed to generate 50 docked poses that were scored with ChemPLP fitness function. The output was read through the Hermes visualization software (version 1.6.2). The top scoring poses, for all the .pdb files in the ensemble, were exported as .pdb files for viewing through PyMOL (PyMOL version 1.7.4, DeLano Scientific LLC).

2.9 Circular dichroism on *Os*UAM1 wild-type and mutants

Circular dichroism (CD) spectra were used to determine the approximate secondary structure content of *Os*UAM1 wild-type and all the four mutants (H273A, D110A, D111A, and D112A). The experiments were performed using the Chirascan-plus CD spectrometer (Applied Photophysics), at the Protein Characterization and Crystallization Facility (PCCF), College of Medicine, University of Saskatchewan. At first, the buffer in which the enzymes were present (25 mM Bicine pH 8.5, 1 mM Mn^{2+}) was scanned to make a baseline measurement in the wavelength range that produced CD spectra of high signal to noise ratio (in this case 210 nm - 280 nm). The concentrations of each of the samples were adjusted to maintain a total absorbance under 2.0. All experiments were performed in a quartz cuvette with a path length of 0.5 mm (Hellma, Germany).

All data was collected in triplicate at 20 ºC. The average CD spectrum for each sample was obtained after a baseline correction was performed. The secondary structure content was determined using the data analyzing software, CDNN 2.1.

Chapter 3: UDP-Galactopyranose Mutase

3.1 UGM from *Deinococcus radiodurans*

*Ec*UGM, *Kp*UGM, and *Mt*UGM have been crystallized without the substrate (unliganded), to gain insights into the structural architecture and the UGM - FAD interaction.[23,31] In the Sanders group, UGM from *Dr*UGM and *Mt*UGM were crystallized with the substrate UDP-Gal*p*.[24,32] In this thesis, the role of some of the active site residues in binding the substrate were studied, using *Dr*UGM as a model system. *Dr*UGM has only 37%, 42% and 39% sequence identity with *Ec*UGM, *Kp*UGM, and *Mt*UGM, respectively, but the substrate and FAD binding residues are highly conserved across the species.[24] Crystals of *Dr*UGM, in complex with UDP-Gal*p*, were produced by co-crystallization experiments. The structures of the *Dr*UGM: UDP-Gal*p* complex with oxidized and reduced FAD were solved to 2.40 Å and 2.50 Å, respectively.[24]

The overall structure and active site of UGM was discussed in detail in Chapter 1. The overall structure of *Dr*UGM is no different to the other known bacterial UGM structures, having three domains and a flexible loop (Figure 3-1). In brief, Domain 1 binds the cofactor FAD, Domain 2 has five α-helices and a mobile loop while Domain 3 is made of anti-parallel ß-strands. When the substrate approaches the active site, the loop and Domain 2 moves towards the active site and the loop closes to completely bury the substrate in the active site. Additionally, all three domains have significant roles in interacting with the incoming substrate. The FAD-binding Domain 1 aids in positioning the galactose sugar under the isoalloxazine ring of FAD for the nucleophilic attack. Domain 2 helps to stack the uridine portion of the substrate and ensures that the substrate is in a favorable binding mode while Domain 3 and the loop contribute crucial residues towards stabilizing the substrate diphosphate.

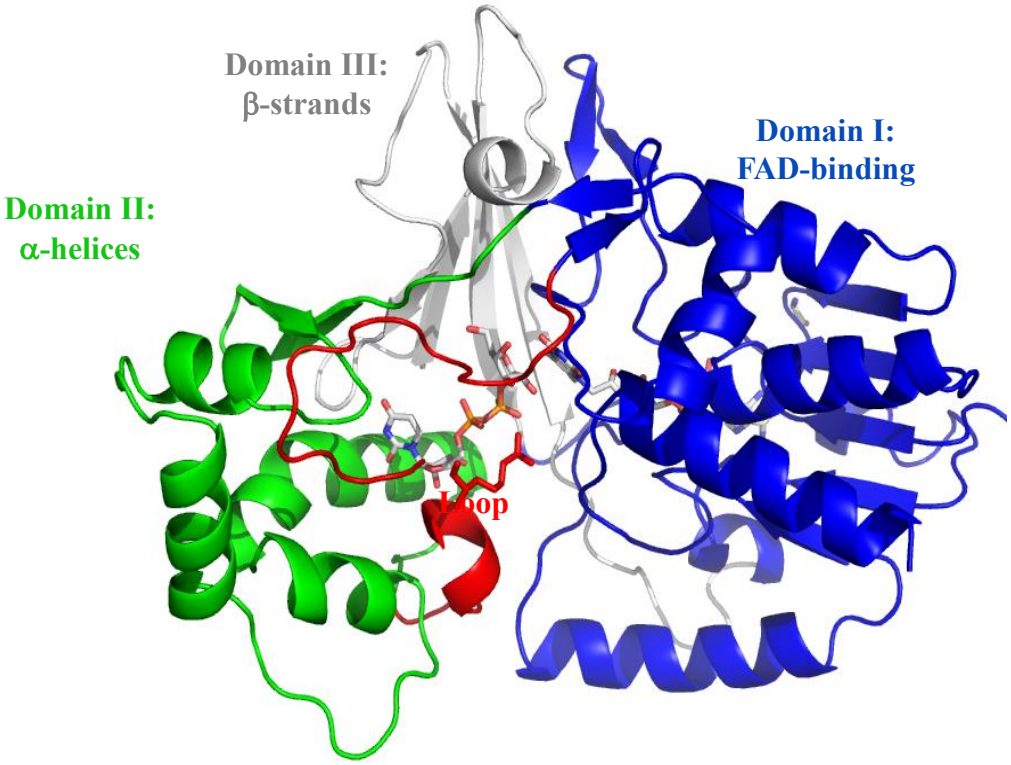


Figure 3-1: The crystal structure of *Dr*UGM.
The domains (Domain 1 - blue, Domain 2- green and Domain 3 - grey) and the mobile loop (red) are highlighted. Domain 1 binds the FAD (white stick); the mobile loop is in the closed conformation, with UDP-Gal*p* (white stick, Domain 2) in the active site.

3.1.1 Active site residues and interactions

In the bound state, UDP-Gal*p* is folded into a U-shaped conformation and is completely buried in the active site. The substrate binding site of *Dr*UGM: UDP-Gal*p* can be divided into three regions; the Uridine-binding region, the Diphosphate-binding region, and the Gal*p*-binding region. The uridine-binding region in both the liganded and unliganded structures from other bacterial species is not much different.[33] Tyr179, Phe176, Tyr180, and Phe175 are important residues involved in recognizing and stacking the uridine. The uracil ring is stacked in between Tyr179, and Phe176; other residues such as Tyr180, Asn296, and Phe175 form a network of hydrogen bonds with O2, O4 and N3 of the uracil ring respectively, as shown in Figure 3-2A.

W184 is involved in hydrogen bonding the C2 and C3 hydroxyl groups of the ribose (Figure 3-2A). Also, there is little difference in the uridine-binding region when the crystal structures with oxidized and reduced FAD were compared.

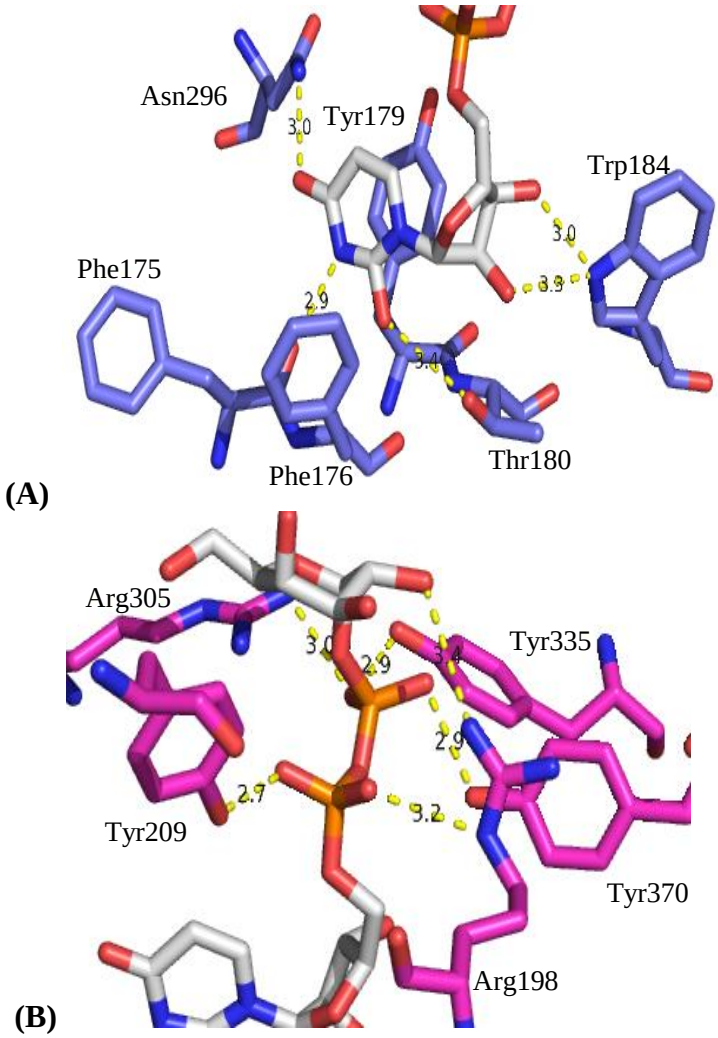


Figure 3-2: Uridine and phosphate binding regions of *Dr*UGM.
(A) Uridine-binding region, displaying the substrate uridine and active site residues in purple (B) Phosphate binding region showing the substrate diphosphates (orange) and active site residues in magenta

Two highly conserved arginine residues, Arg198, and Arg305, play critical roles in stabilizing the substrate diphosphate in the Diphosphate-binding region. Arg198 is in the center of

the mobile loop and is solvent exposed in the absence of the ligand. This residue moves into the active site to stabilize the α-phosphate by forming ionic interactions. Similarly, the β-phosphate is also stabilized by another arginine, Arg305, contributed by Domain 2 (Figure 3-2B). Further stabilization of the diphosphates is achieved by hydrogen bonding interactions of the conserved residues, Tyr209, Tyr370, and Tyr335.

One of the conserved features of the *Dr*UGM structure is the binding of the sugar moiety below the isoalloxazine ring of FAD. The sugar (Gal*p*) is stabilized by His88, His109, Arg364, and Asn372 residues, as shown in Figure 3-3. C3 and C6 hydroxyls of Gal*p* forms hydrogen bonds with His88 and His109 respectively. Arg364 interacts with the C2 hydroxyl through a water-mediated hydrogen bond. Furthermore, the C4 hydroxyl of Gal*p* hydrogen bonds with the O4 of FAD. This interaction also serves as one of the features for substrate recognition; UDP-Glu*p* with the C4 hydroxyl in equatorial position is not recognized by the enzyme.[93,94]

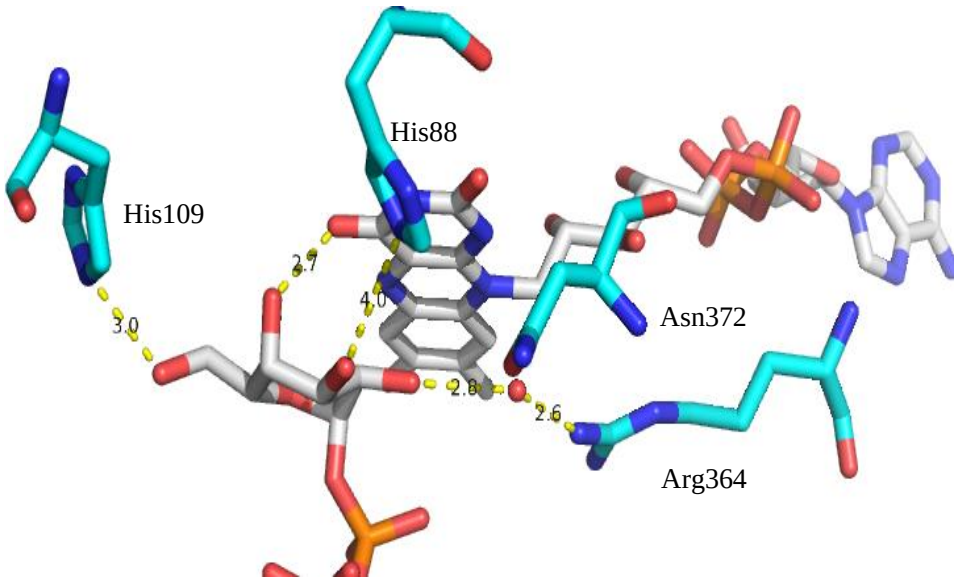


Figure 3-3: Gal*p* binding region of *Dr*UGM.
The active site residues are in (cyan).

Thus, 14 critical active site residues have been identified as interacting with the substrate in the *Dr*UGM: UDP-Gal*p* complex. These residues by enabling the positioning of other residues that interact with the substrate, work to ensure that substrate binding is in the most productive orientation within the active site. Based on the observations from the FAD_{red} crystal structures, the distance between the C1 of Gal*p* and N5 of FAD is in the range of 2.9 Å to 3.2 Å. In the FAD_{ox} crystal structures, the distance between the C1 anomeric carbon of Gal*p* and N5 of FAD is greater than 3.5 Å, thus making the N5 of FAD incapable of carrying out the initial nucleophilic attack on the C1 anomeric carbon.[24] A table (Table 1-2), for comparison of distances between the N5 of the isoalloxazine ring of FAD and the C1 of Gal*p* from oxidized and reduced crystal structure of various bacterial UGMs, is shown in Chapter 1.

3.2 Purification of *Dr*UGM wild-type and point mutants

The *Dr*UGM point mutants were prepared by Dr. Karunan Partha Sarathy, a previous member of the Sanders lab.[94] The *Dr*UGM wild-type and its mutants used in this study were purified using anion-exchange chromatography, hydrophobic chromatography followed by size exclusion chromatography. The sample purity was verified using Sodium dodecyl sulfate – polyacrylamide gel electrophoresis (SDS-PAGE) (Figure 3-4).

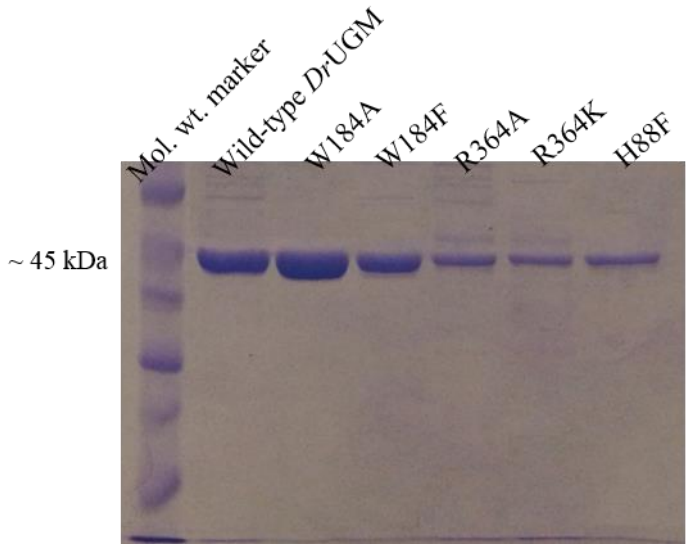


Figure 3-4: SDS-PAGE analysis of concentrated fractions of wild-type *Dr*UGM and point mutants.

3.3 Crystallization of wild-type *Dr*UGM

The structure of unliganded *Dr*UGM wild-type enzyme has not been reported so far. This structure will give insights into the enzyme's structural similarities or differences with the unliganded enzyme from other bacterial species. This will also give us vital information on the structural and conformational changes *Dr*UGM has to undergo to bind the substrate. Unliganded *Dr*UGM was crystallized in 0.1 M Bis-tris propane pH 6.5, 0.2 M potassium bromide, 20 (w/v) % PEG 3350 (Figure 3-5). The crystal condition was further optimized to 0.1 M Bis-tris propane 6.0, 0.2 M potassium bromide, 20 (w/v) % PEG 3350 using grid-screens. The crystals were further improved with an additive, iron (III) hexahydrate, added in a 9:1 ratio to the optimized crystallization condition. The crystals, however, did not yield good quality diffraction data, as they diffracted to ~ 6 Å at the CLS.

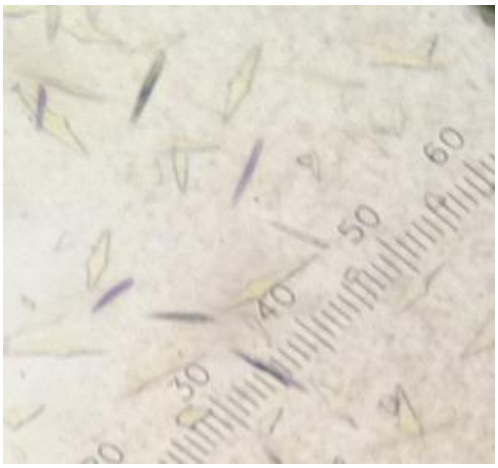


Figure 3-5: Optimized crystal condition for wild-type *Dr*UGM (holoenzyme).

3.4 *Dr*UGM point mutants

Although the roles of most active site residues in substrate interaction and their importance to the *Dr*UGM active site are well understood, the structural roles of a few active site residues, such as Trp184, Arg364, Asn372, and His88, in binding the substrate are not clear (Figure 3-6). Six point mutants of *Dr*UGM were prepared to study their importance to substrate binding. The roles of the residues mentioned above were studied by performing HPLC-based kinetic assays, modeling using Rosetta Backrub and GOLD docking studies. First, results from the efforts to crystallize these mutants with the substrate are discussed.

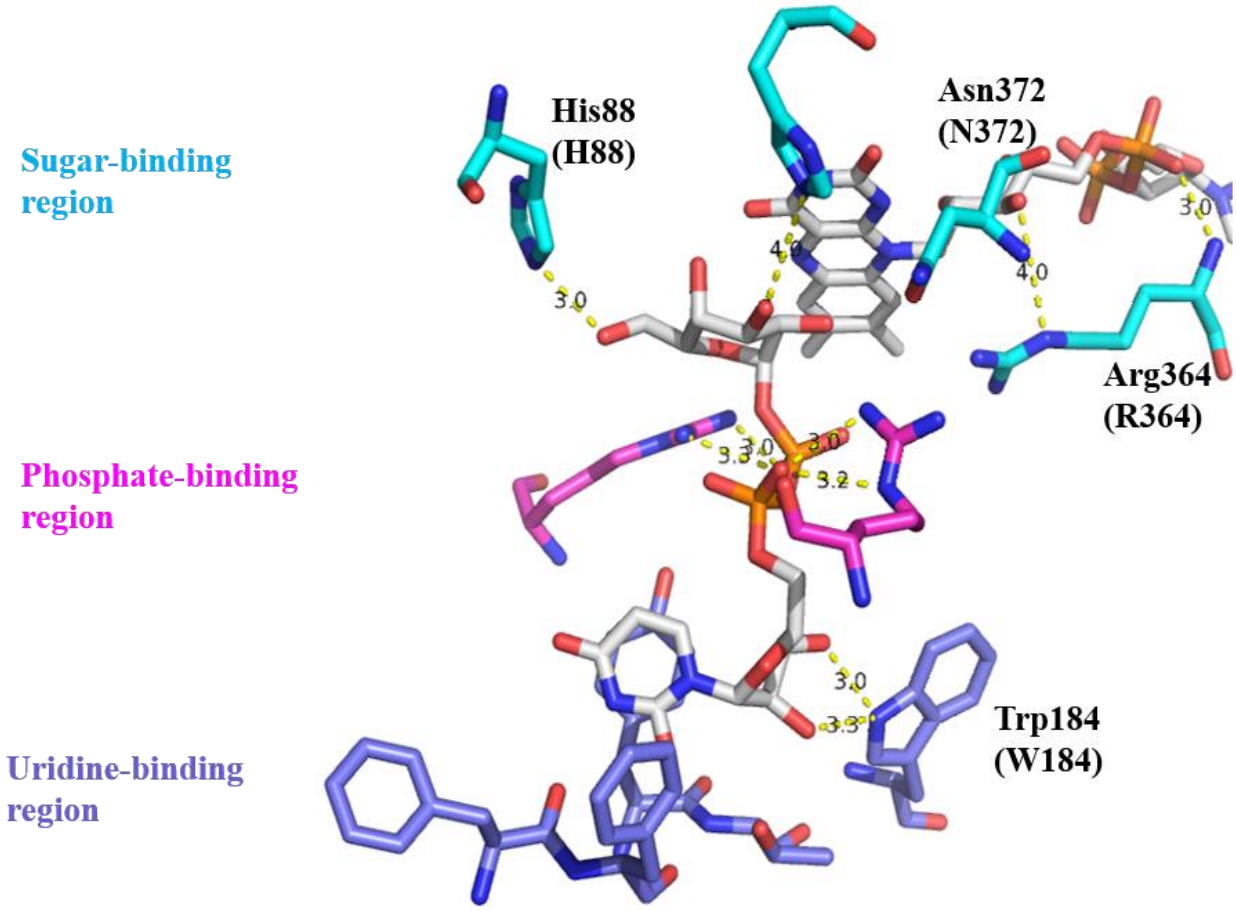


Figure 3-6: The active site residues of *Dr*UGM highlighting point mutants prepared for study.

3.4.1 Crystallization and diffraction of *Dr*UGM point mutants

Crystals of the point mutants co-crystallized with UDP-Gal*p* were obtained in the following conditions. H88F was crystallized in, 0.1 M Sodium acetate pH 4.6, 25% (w/v) PEG 4000; R364A in 0.1 M Sodium acetate, pH 4.6, 15% (w/v) PEG 20000; R364K in 0.1 M sodium acetate pH 4.6, 4% (w/v) PEG 4000; W184F in 0.2M Sodium iodide 0.1M Bis-Tris propane pH 6.5, 20% (w/v) PEG 3350 and W184A 0.1 M MES pH 6.5,30% (w/v) PEG 4000, respectively. The crystals are shown in Figure 3-7.

In general, the *Dr*UGM mutant crystals diffracted poorly; diffraction experiments were conducted on a large number of crystals of each mutant at the CLS. However, the best diffraction datasets collected were in the 4 - 6 Å range, which was not good enough to observe the mutations or changes in substrate binding. Apart from optimizing the crystals by various grid screens or additive screens, a number of post-crystallization treatment techniques were also performed. Allowing the crystals to anneal for 5-10 seconds before diffraction (performed at CLS); exposing them to air for up to 20 minutes before freezing; transferring hanging drops onto crystallization solutions having increasingly higher concentrations of various crystallization ingredients (up to 15 %) over 8 -12 hours.[95-97] However, none of these methods seemed to improve diffraction quality since the crystals developed cracks or damaged easily.

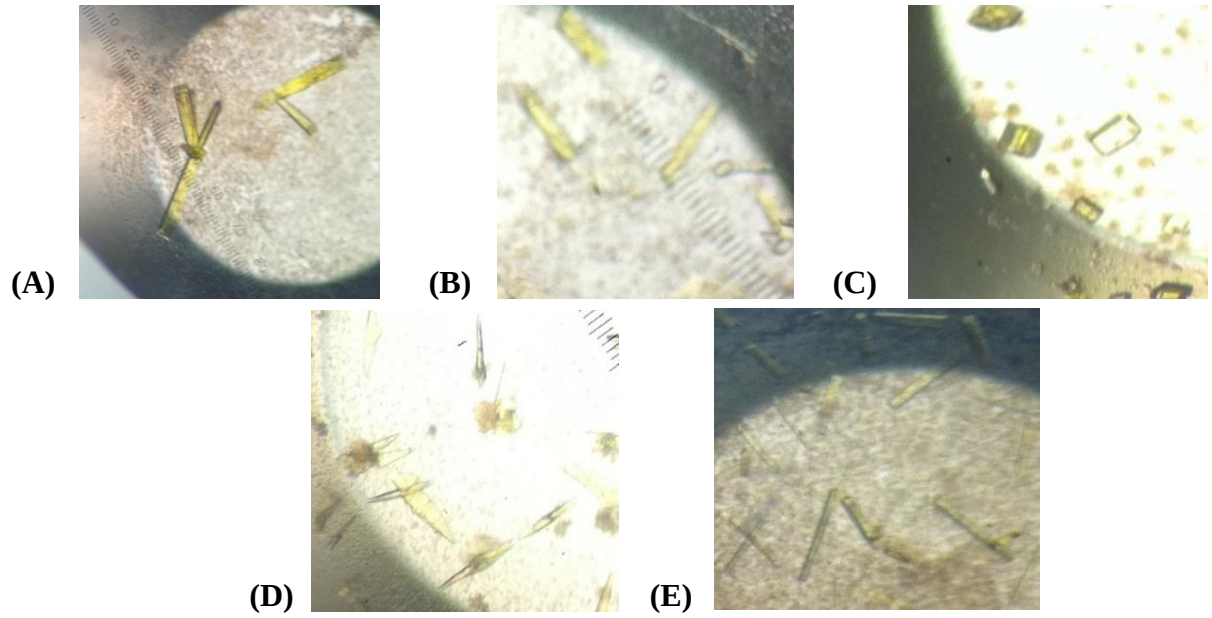


Figure 3-7: Crystal hits for *Dr*UGM point mutants.
(A) H88F, (B) R364A, (C) R364K, (D) W184F and (E) W184A

3.4.2 Kinetic evaluation of mutants

3.4.2.1 W184A and W184F

The Michaelis-Menten curves for W184A and W184F mutants are shown in Figure 3-8. Mutating Trp184 to alanine has a drastic effect on the K_m value of *Dr*UGM. The K_m increases ~ 20-fold compared to the wild-type enzyme, which means that the enzyme now requires more substrate to achieve half-maximal activity. k_{cat}, the turnover number of the enzyme decreases ~ 360 times. The K_m and k_{cat} values for the W184F mutant are better compared to the alanine mutant, although still not as good as the wild-type enzyme. The K_m of W184F is 670 µM, which is a ~ 12 - fold increase in the substrate required to obtain half-maximal activity but the k_{cat} decreases by ~ 15 times. In addition, the specificity constant (k_{cat} / K_m) value is also better for W184F than the W184A. The kinetic constants are shown in Table 3-1.

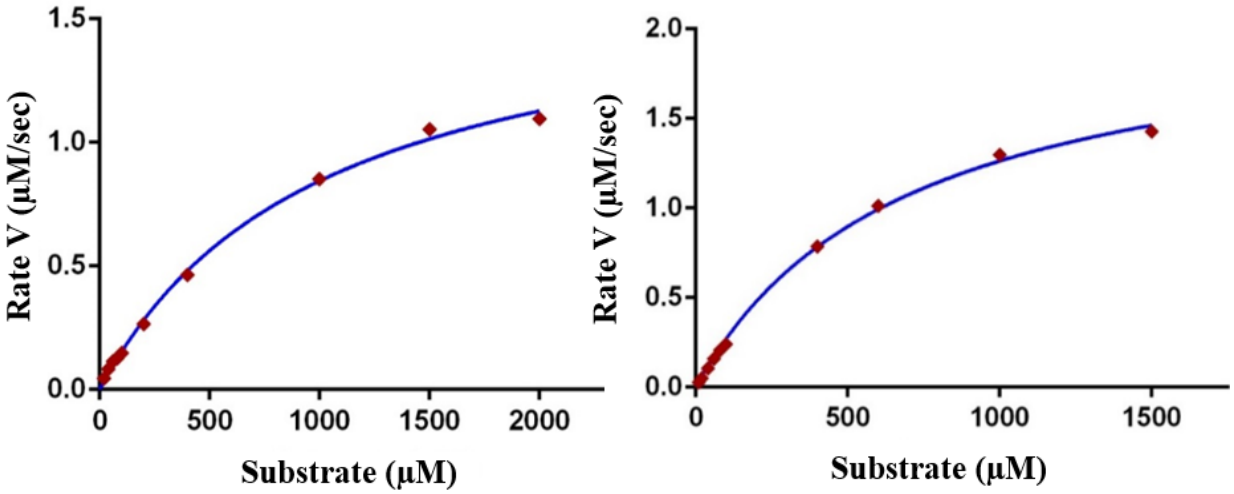
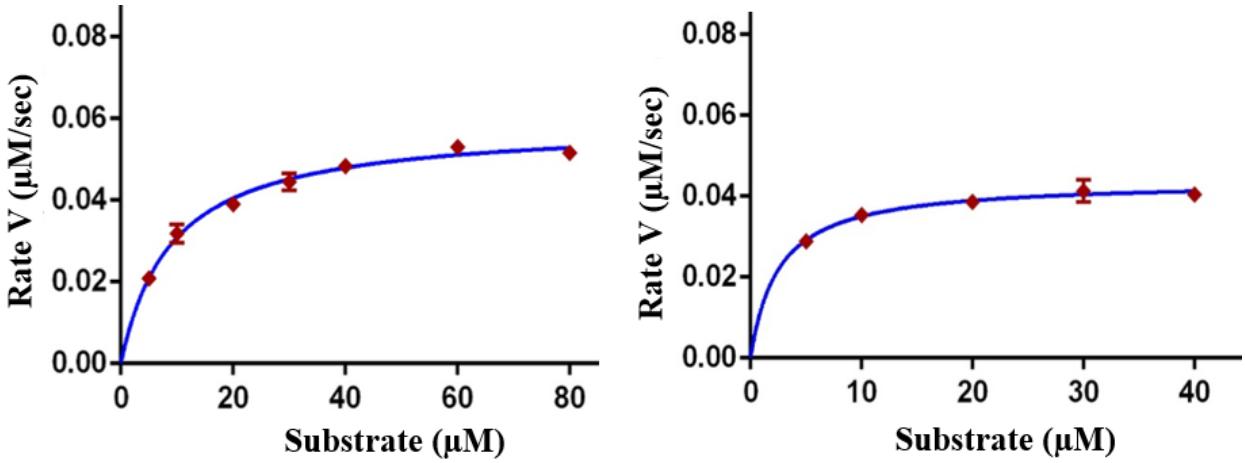


Figure 3-8: Michaelis-Menten curves for (A) W184A and (B) W184F mutants.

3.4.2.2 R364A and R364K

Mutating Arg364 to alanine or lysine decreased the K_m value of the enzyme by 6- fold and 22-fold respectively, as shown in Table 3-1. Both these mutants required lesser amounts of

substrate to get to half-maximum activity, compared to the wild-type but their k_{cat} values were rather insignificant. R364K had only a slightly better k_{cat} than the alanine mutant, even though lysine is a positively charged residue. The Michaelis-Menten curves for R364A and R364K are shown in Figure 3-9.



(A) **(B)**

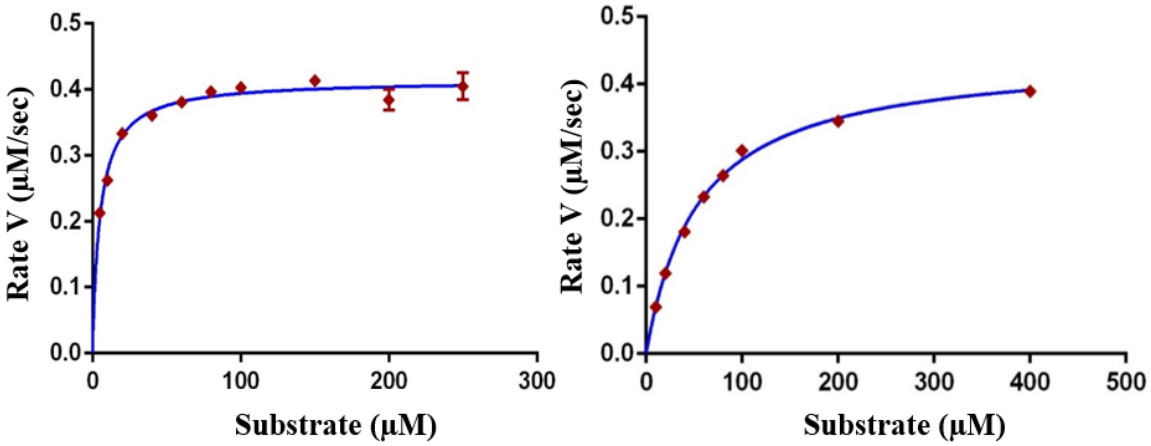
Figure 3-9: Michaelis-Menten curves for (A) R364A and (B) R364K mutants.

3.4.2.3 H88F

In *Dr*UGM, His88 interacts with the C3 hydroxyl of the substrate sugar. However, in eukaryotic UGMs, this histidine is replaced by phenylalanine. The H88F mutation was made to see if there was any observable effect on the active site of bacterial UGMs since phenylalanine cannot contribute to coordinating the hydroxyl of the sugar. The kinetic parameters derived from the plot (Figure 3-10) show that there is a decrease in the K_m by ~ 11-fold, but the k_{cat} is only 4-fold less (Table 3-1). The turnover number is only 4-fold less than the wild type enzyme, making it more efficient as evidenced by the better specificity constant.

3.4.2.4 N372D

Apart from stabilizing the α-phosphate of the substrate, Arg198 also binds the substrate in the active site by hydrogen bonding with Asn372. The N372D mutant was made on the basis that in *Kp*UGM, the residue that interacts with arginine is an aspartic acid. Our kinetic evaluation of this mutant shows that the kinetic parameters are more similar to that of wild-type *Kp*UGM than to wild-type *Dr*UGM, in terms of k_{cat} (Table 3-1).



(A) **(B)**

Figure 3-10: Michaelis-Menten curves for (A) H88F and (B) N372D mutants.

Table 3-1: Kinetic evaluation of *Dr*UGM active site point mutants.

Enzyme	K_m (µM)	k_{cat} (s^{-1})	k_{cat} / K_m ($s^{-1}µM^{-1}$)
*Dr*UGM wild type	55 ± 7	66 ± 2	1.18
W184A	1007 ± 61	0.18 ± 0.01	1.8×10^{-4}
W184F	670 ± 26	4.26 ± 0.08	6.36×10^{-3}
R364A	9 ± 0.3	2.94×10^{-5}	3.26×10^{-6}
R364K	2.5 ± 0.6	$0.17 \pm 5 \times 10^{-3}$	7.6×10^{-2}
H88F	5.1 ± 0.3	16.1 ± 0.2	3.36
N372D	54 ± 7	9 ± 0.4	0.17
*Kp*UGM wild type	45 ± 6	5 ± 0.6	0.11

To understand the structural changes caused by these active site mutants and to observe possible changes that may affect the substrate binding mode, *Dr*UGM mutant models were built, using the modeling program Rosetta-Backrub, and they were compared with the crystal structure of the wild-type enzyme. An outline of the method used by this software to perform modeling has been included in Supplementary Section S.1. In general, the overall structures of these mutants were similar to the wild-type enzyme except for a few residues in close proximity to the mutation.

3.4.3 Results from modeling studies

3.4.3.1 W184A and W184F

The structural overlay of the wild-type enzyme on *Dr*UGM W184A shows that the alanine is ~ 6 - 7 Å away from the uridine-ribose of UDP-Gal*p*, and unlike the tryptophan, it is not in a position to coordinate the C2' and C3' hydroxyls of the ribose (Figure 3-11A). In *Dr*UGM W184F, the phenylalanine is also unable to establish contacts with the ribose (Figure 3-11B). However, phenylalanine, being a much more bulky group than alanine, contributes in the π-stacking of the uridine-ribose. In both these models, Gln183 rotates to help in coordinating the C2' ribose hydroxyl. *Dr*UGM W184A shows more changes in the backbone of residues, around the point mutation, compared to the *Dr*UGM W184F enzyme.

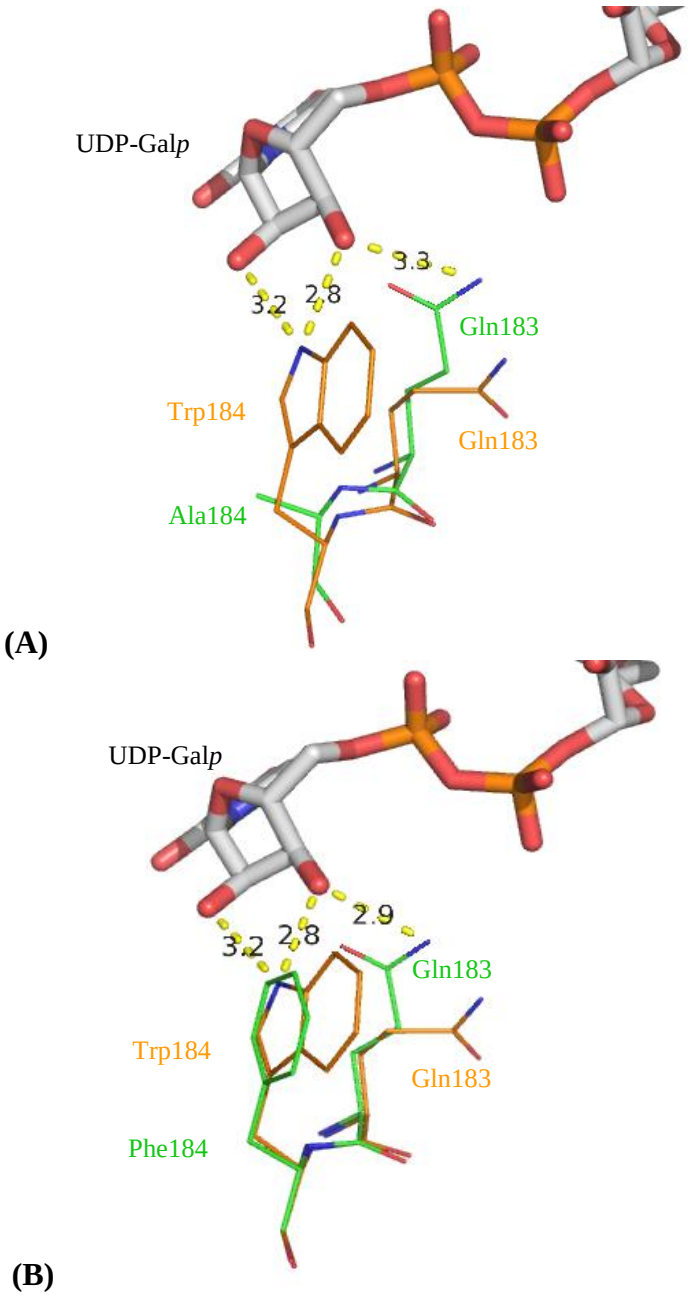


Figure 3-11: Models of W184A and W184F as generated by Rosetta Backrub.
The wild-type residues are in orange, and the mutant residues for (A) W184A and (B) W184F are in green.

3.4.3.2 R364A and R364K

Arg364 is in a position to coordinate the cofactor FAD and also make water-mediated contacts with the C2 hydroxyl of the galactose ring of the substrate. Mutating this arginine to alanine removes these interactions (Figure 3-12A). Based on the models generated, the mutation may also change the positioning of a few residues in the active site, such as Tyr370 and the critical phosphate stabilizing Arg198. These small changes in the active site could affect the substrate binding mode. The model generated for R364K is also similar to R364A in that the above-mentioned residues are positioned differently. However, in this case, Lys364 can either interact with FAD or form a water-coordinated hydrogen bond with the C2 substrate hydroxyl due to its flexible side chain, since unlike Arg364 it cannot simultaneously make both interactions (Figure 3-12B).

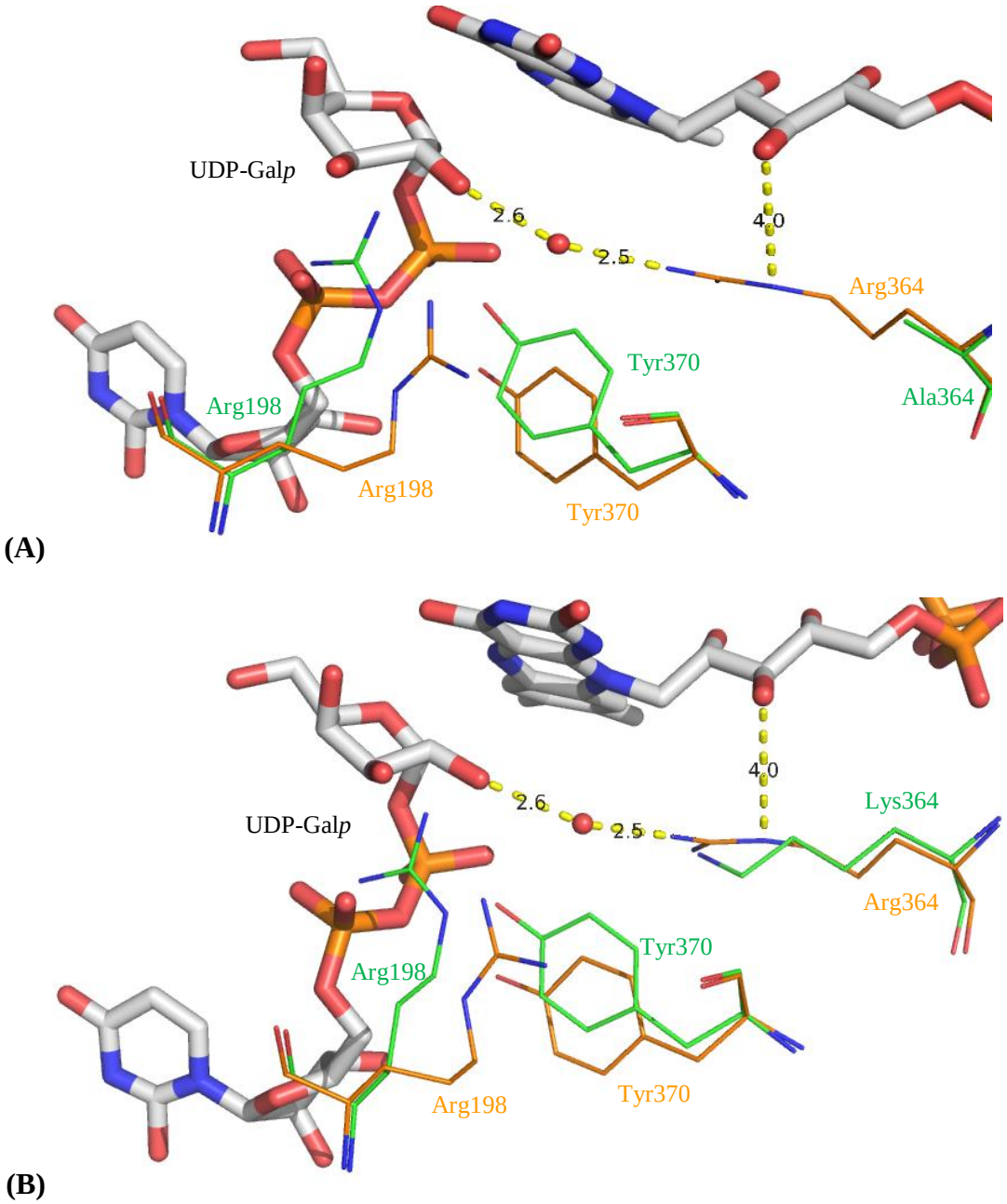


Figure 3-12: Models for R364A and R364K as generated by Rosetta Backrub.
The wild-type residues are in orange, and the mutant residues for (A) R364A and (B) R364K are in green.

3.4.3.3 H88F

His88 makes hydrogen bonds with the isoalloxazine ring of FAD and the C3 hydroxyl of the galactose ring, as shown in Figure 3-13. The model generated for the *Dr*UGM H88F indicates that even though phenylalanine is incapable of making contacts similar to those made by histidine, its aromatic ring is in a relatively similar position to that of the histidine ring (Figure 3-13). This may have some stacking effect on the isoalloxazine ring and galactose sugar, thereby aiding in the positioning of the sugar for nucleophilic attack by FAD.

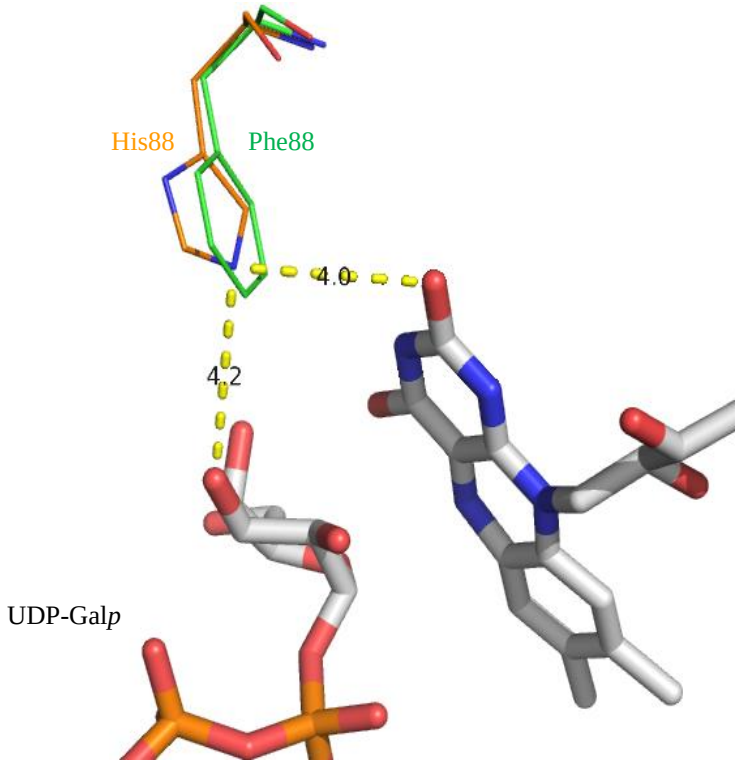


Figure 3-13: Model for H88F as generated by Rosetta Backrub.
The wild-type residue His88 is in orange, and the mutant residue Phe88 is in green

3.4.4 Docking of *Dr*UGM mutants

UDP-Gal*p* was docked into the active site of the wild-type enzyme and the mutant structures, using the ensemble docking feature in GOLD. The process of creating an ensemble in

GOLD is described in Supplementary Section S.2. The advantage of this technique is the ability to produce poses of the docked substrate for direct comparison of the superimposed structures. Each run generates numbers that can be analyzed to yield further information about substrate binding in the various mutants, thereby providing information about the effect of the mutation. The results shown in this section has been tabulated using four different numbers generated from the docking runs. Scores under 'ensemble analysis', explain which protein in the ensemble docks the substrate better; generally the higher the score for the protein, the better the substrate docks in the protein. Each run is asked to generate 100 poses of the substrate for the ensemble. Proteins among the ensemble which provide more favorable interactions for docking of the substrate generate a higher number of docked poses. So, the number of poses generated for each protein in the ensemble is also tabulated for analysis. Average RMSD of the ranked poses gives an idea of how much the atoms of the docked poses have moved with respect to the crystallographic substrate binding pose. Finally, the average distance (Å) between N5 of FAD and C1 anomeric carbon of the Gal*p* (sugar) is also calculated to analyze how the docked poses compare to the distance measured in the crystal structure. This number also gives an idea of how the changes due to mutations affect the productive mode of substrate binding in the active site.

3.4.4.1 Trp184

In general, the docked poses of both mutants (W184A & W184F) are not in the correct conformation necessary for the catalysis of the reaction. Since there is no Trp184 to coordinate the ribose hydroxyls in both the mutants, most docking poses of the mutants are flexible, especially in the uridine-ribose region (Figure 3-14).

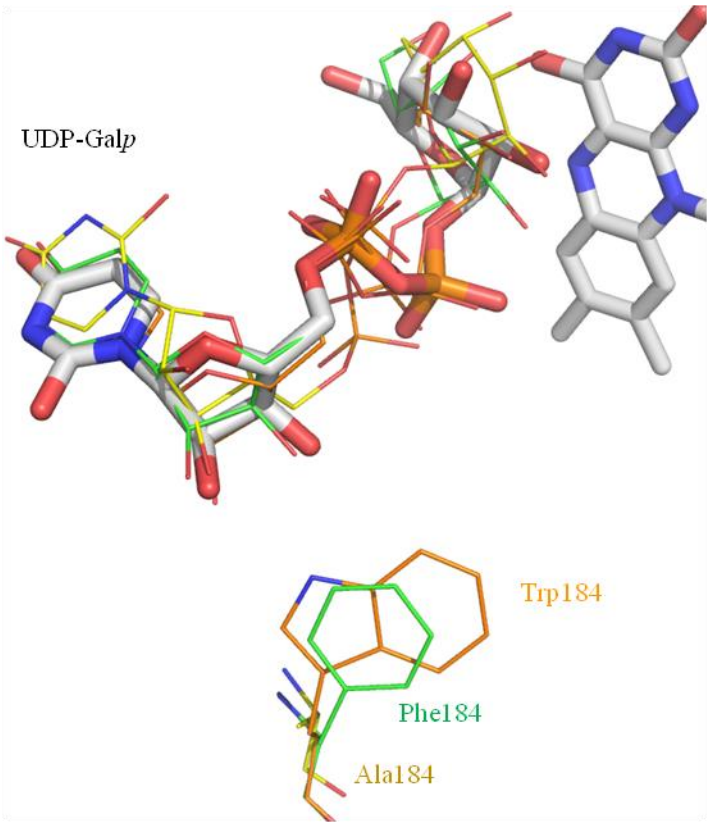


Figure 3-14: The top GOLD docking poses for W184A and W184F.
The top-ranked docking poses from wild-type *Dr*UGM (orange), W184A (yellow) and W184F mutants are shown superimposed on the UDP-Gal*p* (stick; white) from the crystal structure.

Table 3-2 lists the scores from ensemble analysis, percent of poses for each structure in the ensemble and the average RMSD of the docked poses. The wild-type enzyme accounts for 66 % of the docked poses and also has the highest score from this ensemble analysis. W184F has a better score and more % poses than W184A. Although W184F has a lower average RMSD of the docked poses when compared to W184A, the wild-type enzyme still has the lowest value.

3.4.4.2 Arg364

The docked poses obtained for R364 mutants are shown in Figure 3-15. Due to the changes in conformation of residues such as Arg198, Tyr370, and the mutated residues, the Gal*p* portion

of the substrate is observed to have moved further away from its position under the FAD, in the docked poses.

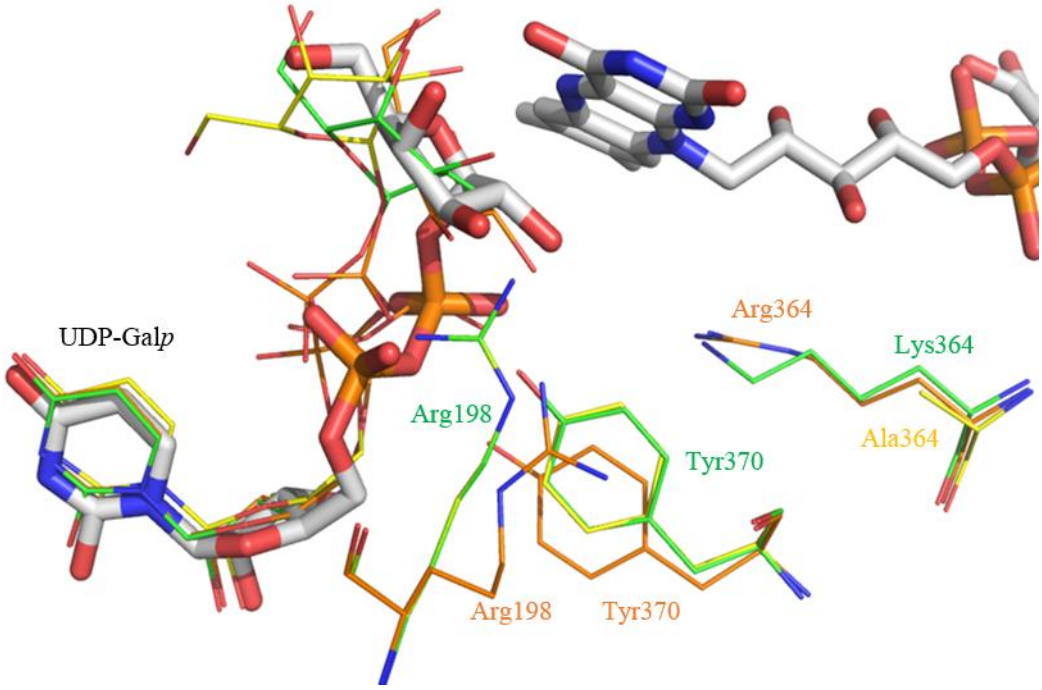


Figure 3-15: The top GOLD docking poses for R364A and R364K.
The top-ranked GOLD docking poses for wild-type *Dr*UGM (orange), R364A (yellow) and R364K (green) mutants shown superimposed on the UDP-Gal*p* (stick: white) from the crystal structure

Table 3-3 indicates that the ensemble analysis scores of the mutants are lower than the wild-type enzyme, and the lysine mutant is better than the alanine mutant. The % of docked poses is much lower, and the average RMSD is higher for the mutants than the wild-type, reflecting the inability of the substrate to bind in a productive conformation. However, an overall comparison of the two mutants shows that R364K scores better than R364A.

Table 3-2: GOLD docking results for Trp184.

Enzyme	Ensemble Analysis score	% Poses	Average RMSD	Average N5 (FAD) C1 distance (Å)
Wild-Type	24.16	66	1.41	3.50
W184A	22.84	14	2.00	4.28
W184F	23.04	20	1.81	4.57

Table 3-3: GOLD docking results for Arg364.

Enzyme	Ensemble analysis score	% Poses	Average RMSD	Average N5 (FAD) C1 distance (Å)
Wild-Type	18.84	50	1.39	3.63
R364A	13.61	16	2.43	4.76
R364K	16.98	34	2.16	4.42

Table 3-4: GOLD docking results for His88.

Enzyme	Ensemble analysis score	% Poses	Average RMSD	Average N5(FAD) C1 distance (Å)
Wild-Type	26.22	48	1.90	4.09
H88F	26.49	52	1.30	3.67

3.4.4.3 His88

Ensemble docking scores of H88F are better than the wild-type enzyme. The docked poses are also similar to the wild-type enzyme, as shown in Figure 3-16. The % docked poses based on the number of solutions in each run, and the average RMSD of docked poses of UDP-Gal*p* is slightly better for the mutant, as shown in Table 3-4. This could be due to the hydrophobic Phe88 being capable of keeping the Gal*p* and the isoalloxazine ring in the correct position without significant changes to the backbone of residues nearby.

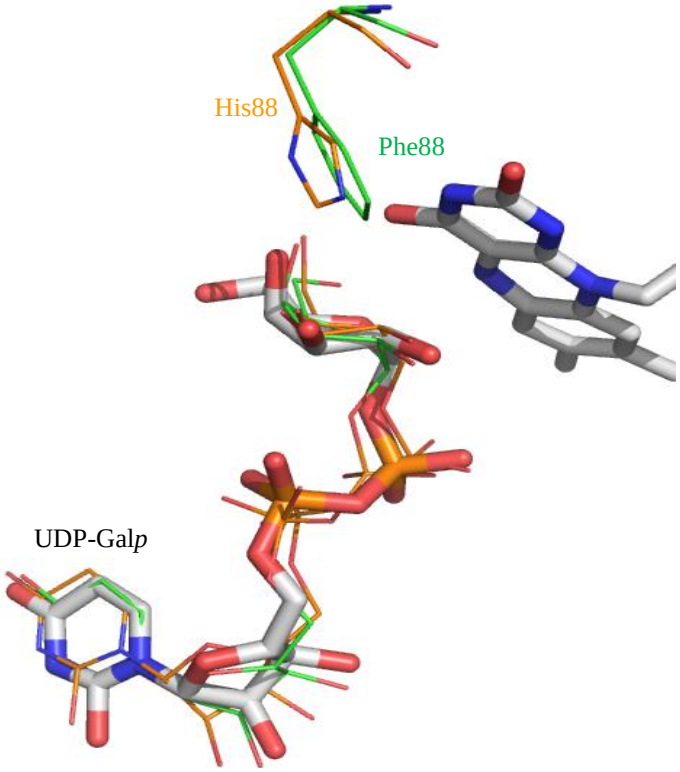


Figure 3-16: The top GOLD docking poses for H88F mutant.
The top-ranked GOLD docking poses for wild-type *Dr*UGM (orange), H88F mutant shown superimposed on the UDP-Gal*p* from the crystal structure

3.4.5 Crystal structure of N372D

The crystal structure of N372D was solved to 2.7 Å by previous members of the Sanders lab.[94] The overall structures of the substrate bound *Dr*UGM and that of the N372D mutant are similar. In this crystal structure, Asp372 occupies a similar position to Asn372 and can make hydrogen-bonding interactions with Arg198. Asp372 is able to make two hydrogen bonds whereas Asn372 is able to make only one with Arg198 (Figure 3-17). However, this mutation has little effect on the binding mode of the substrate, as the conformations of active site residues do not change much when compared to the wild-type enzyme. The FAD_{red} crystal structure of wild-type

*Kp*UGM shows that *Kp*UGM Asp351 is in position to form two bonds to stabilize Arg174, the residue that moves in with the mobile loop.[93]

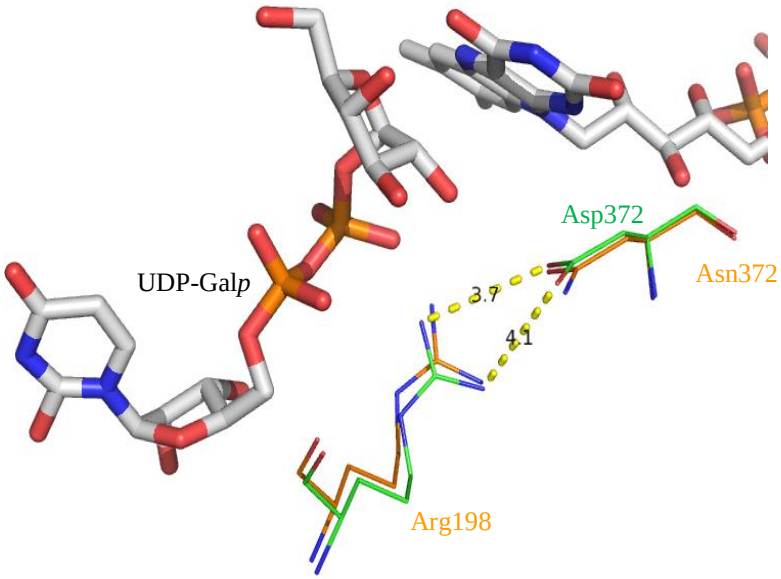


Figure 3-17: Arg198 and Asn372 of wild-type *Dr*UGM and N372D mutant.
The wild-type residues are shown in orange while those of N372D are shown in green.

3.4.6 Discussion

Results from kinetic assays, modeling, and docking experiments were analyzed to understand the role of the chosen *Dr*UGM active site residues in binding UDP-Gal*p* in an active conformation. In the wild-type enzyme, Trp184 interacts with the ribose hydroxyls through hydrogen bonds. Both Trp184 mutants are unable to coordinate the ribose, but the presence of the bulky, aromatic phenylalanine has an influence on substrate binding, as evidenced by the kinetic and docking experiments. The results obtained from docking experiments are comparable to those obtained from kinetic assays. The wild-type enzyme has much higher efficiency than the two mutants. W184A and W184F have displayed a loss of efficiency and a decreased ability to bind

the substrate when evaluated by kinetic assays. This is also reflected in the docking studies as the wild-type enzyme has the highest score from ensemble analysis and a higher number of docked poses with a lower RMSD. Moreover, the higher average RMSD of the poses generated for the mutants suggest that the substrate is mostly not held in the productive conformation within the active site; a factor that could be leading to the loss in kinetic efficiency of the mutants. The aromatic phenylalanine (W184F) may have a positive effect on substrate binding. The W184F mutant has a lower average RMSD for the docked poses and a better kinetic efficiency than the W184A mutant. Therefore, even though the wild-type Trp184 can coordinate the ribose hydroxyls, the aromatic rings and the hydrophobic nature of this residue are responsible for positioning the uridine portion of the substrate via hydrophobic interactions thereby making it an essential residue for ensuring the best substrate binding mode.

Arg364 is the third arginine in the active site, apart from the arginines that stabilize the substrate diphosphate. This residue interacts with the FAD in the absence of the substrate in the active site but moves into the active site to coordinate the incoming substrate. The models predict changes in the active site conformation, as this residue not only influences the positioning of Gal*p* but also influences the position of residues crucial for stabilizing the substrate diphosphate. The possible movement of Tyr370 towards the β-phosphate of the substrate is likely to displace the critical residue Arg198 from its position, thereby causing changes to the substrate binding mode. The docked poses have the diphosphate moiety of the substrate occupying a different position in the active site and Gal*p* is away from the isoalloxazine ring, compared to the wild-type enzyme. The average distance between the N5 (FAD) of C1 anomeric carbon (Gal*p*) is close to 4.5 Å (and higher than wild-type) for the mutants. This trend is also reflected by the kinetic assays performed with the mutants. R364A and R364K were unable to bind enough substrate to achieve the half-

maximal activity of the wild-type and had very low efficiency. This could be due to the inability of both mutants to bind the substrate in a productive conformation within the active site, as suggested by the docking studies. Ensemble analysis scores both mutants lower than the wild-type enzyme, and R364K better than R364A. Therefore, apart from making contacts with FAD, Arg364 helps keep critical residues like Arg198 in the correct position, thereby letting the substrate bind in the most productive mode.

The mutation of His88 with Phe88 was done based on eukaryotic UGM sequences. As evidenced by the kinetic studies, the mutant enzyme is able to bind the substrate in the productive binding mode. The docking experiments are also in favor of this argument. The docked poses of the substrate in the mutant are similar to that of the wild-type enzyme. Moreover, ensemble docking scores H88F marginally better than the wild-type enzyme and the average RMSD of docked poses of UDP-Gal*p* is also better for the mutant. *Dr*UGM H88F thus aids in positioning the sugar under FAD, due to the aromatic group of phenylalanine. It is also noted that Phe88 in the mutant model occupies a similar position with respect to both the FAD and Gal*p* as that of Phe66 in the crystal structure of FAD_{red} eukaryotic *Af*UGM.[35] *Af*UGM Phe66 however, is closer to Gal*p* (~ 3.8 Å) than *Dr*UGM His88 (~ 4.2 Å), as it cannot form any bonds with either FAD or Gal*p*. Thus, the aromatic imidazole ring of His88 also contributes to positioning Gal*p*, serving to orient the sugar for catalysis.

Once the substrate is in the active site, Asn372 hydrogen bonds with Arg198, thereby keeping the mobile loop closed and completely burying the substrate in the active site. The *Dr*UGM N372D mutant is able to form this interaction like the wild-type enzyme, without changes to the orientation of the substrate or other active site residues. However kinetic studies show that the mutation alters the efficiency of the enzyme, making N372D more similar to *Kp*UGM. This

may be a result of the two hydrogen bonds that Asp372 can make with Arg198, which may slow the release of the product formed, ultimately reducing the efficiency of the enzyme.

3.5 Inhibition of *Mt*UGM by MS-208

Since UGM is a potential drug target, as mentioned in the introductory chapter, a constant search for identification of new lead compounds is underway, and a number of studies have been conducted for characterization and analysis of these inhibitors. Inhibitors identified through virtual screening have had more success; although in general have low potency against bacterial UGMs. Among the non-substrate like inhibitors identified from virtual screening, compounds with a 5-hydroxy-pyrozole core were generally more successful in inhibiting UGMs and more importantly were nontoxic to human cells.[54] This section of the chapter will discuss the inhibition experiments performed on *Mt*UGM with a recently identified fungicidal compound, MS-208, having the 5-hydroxy-pyrozole core. Studies described here are focused towards inhibiting prokaryotic *Mt*UGM, since it is a validated drug target.[21] MS-208 demonstrates moderate inhibition against *Mt*UGM (IC_{50} value ~ 64 ± 1 μM). A study conducted by Dr. Pinto's group (Simon Fraser University, British Columbia), hypothesized that MS-208 (Figure 3-18A) may inhibit *Mt*UGM by binding to a novel allosteric site. Based on their molecular dynamics, docking and Saturation transfer difference-nuclear magnetic resonance (STD-NMR) studies, the allosteric site was proposed to be located behind the adenine binding region of FAD [unpublished results].[98] This proposed allosteric site has a small loop with negatively charged residues, called the A-loop and another groove, in between which MS-208 binds, as shown in Figure 3-18B.

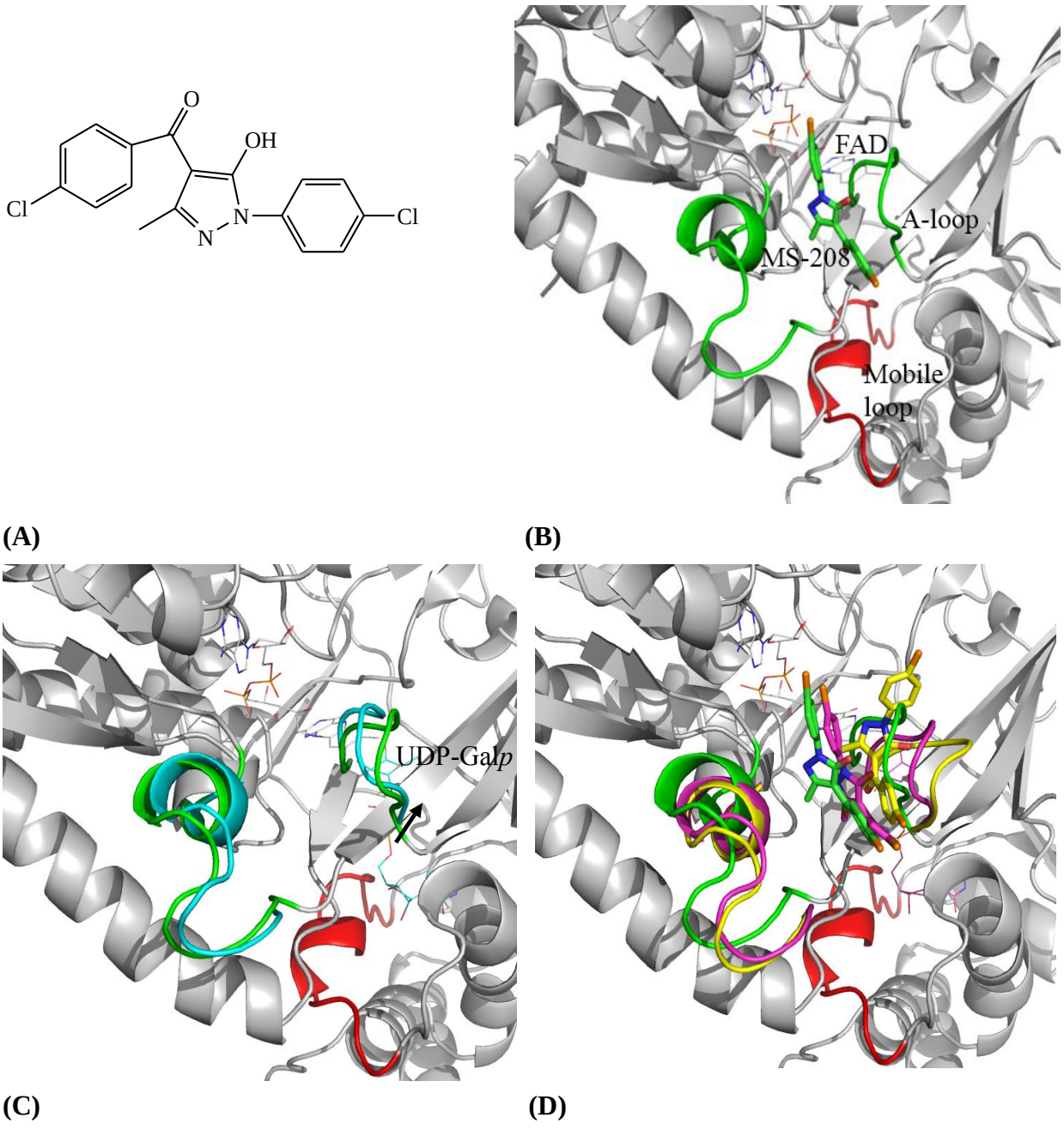


Figure 3-18: MS-208 and its binding to an allosteric site in *Mt*UGM.
(A) The structure of MS-208. (B) MS-208 bound to the allosteric site of *Mt*UGM. The active site mobile loop is open since there is no substrate in the active site. (C) The allosteric site (cyan) when there is substrate in the active site (cyan). (D) Changes in the binding modes of MS-208 (magenta) and A-loop (magenta), when the substrate binds in the active site after MS-208 binds in the allosteric site; the destabilized A-loop of the allosteric site (yellow) when MS-208 (yellow) binds in the allosteric site in the presence of substrate in the active site [unpublished results].[98]

Some of the other significant findings from the study were that MS-208 had a higher affinity to *Mt*UGM than the *Mt*UGM: UDP-Gal*p* complex. The binding of MS-208 weakens the binding of substrate in the active site due to the inability of the active site residues to be in the right conformation. Furthermore, the binding of the substrate also weakens the binding of MS-208 in the allosteric site, because of the destabilization of the A-loop (Figure 3-18D). To determine the mode of inhibition and to generate kinetic numbers, we kinetically evaluated the inhibition of *Mt*UGM by MS-208.

3.5.1 Purification of *Mt*UGM

To overcome the solubility issues, *MtUGM* was cloned into pDESTHisMBP, using Gateway cloning, so that the protein would be expressed with an N-terminal His-MBP-tag.[72] The HisMBP-tagged *Mt*UGM was purified by affinity chromatography and the tag was cleaved using TEV protease. Pure *Mt*UGM was obtained after applying the sample through a Ni-sepharose affinity column.[32] The purity of the protein was assessed using SDS-PAGE (Figure 3-19).

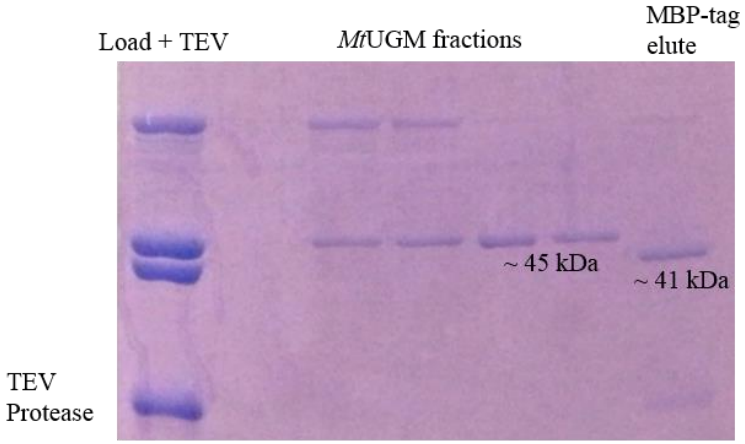


Figure 3-19: SDS-PAGE analysis of purified fractions of MtUGM, after the HISMBP-tag was cleaved using TEV protease.

3.5.2 Kinetic evaluation of *Mt*UGM inhibition by MS-208

The kinetic evaluation of *Mt*UGM inhibition by MS-208 was performed using HPLC - based kinetic assays, as described in chapter 2. The enzyme taken in buffer (both at constant final concentration) was reduced and incubated with MS-208, reacted with the substrate, quenched and analyzed by HPLC to observe product conversion. The final % conversion values were obtained, and rates were determined. Three different inhibitor concentrations (60 µM, 120 µM, and 200 µM) were used for this study.

3.5.2.1 Michaelis-Menten plot

Non-linear regression analysis was performed to determine the kinetic parameters. The K_m and k_{cat} values calculated for *Mt*UGM in the absence of inhibitor were 45.4 ± 3.0 µM and 7.8 ± 0.2 sec^{-1}, respectively. The Michaelis-Menten plot obtained with and without inhibitor is shown in Figure 3-20. At increasing inhibitor concentrations, the K_m of the enzyme increases while the maximum velocity obtained decreases, as shown in Table 3-5.

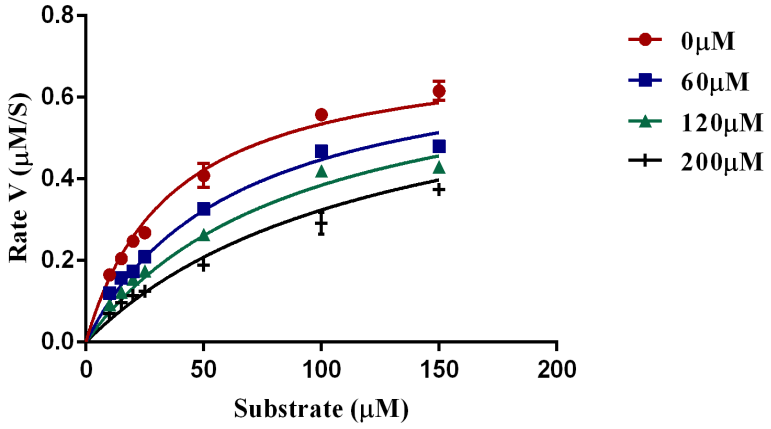


Figure 3-20: Michaelis-Menten plot for inhibition of *Mt*UGM by MS-208.
The plot displays the Rate vs. Substrate (UDP-Gal*f*) concentrations in the absence of MS-208 and the presence of inhibitor at three different concentrations, 60 µM (blue), 120 µM and 200 µM (black)

Table 3-5: Decreasing values of the apparent K_m and maximum velocity (V_{max}) of *Mt*UGM with increasing MS-208 concentrations.

I (µM)	$K_{m,app}$ (µM)	$V_{max,app}$ (µM/S)
0	45.4 ± 3.0	0.78 ± 0.02
60	51.5 ± 4.0	0.67 ± 0.02
120	63.3 ± 6.2	0.63 ± 0.03
200	90.6 ± 12.4	0.57 ± 0.04

3.5.2.2 Lineweaver-Burk plot

Initial efforts to globally fit the data points to an equation for competitive inhibition, using SigmaPlot, failed, as MS-208 did not behave as a competitive inhibitor. Moreover, competitive inhibition was ruled out, since lines corresponding to different inhibitor concentrations did not intersect on the y-axis of the Lineweaver-Burk plot, as shown in Figure 3-21.

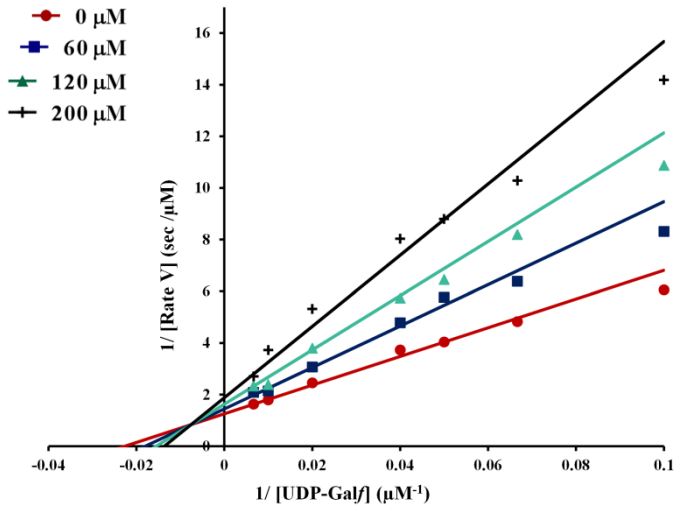


Figure 3-21: Lineweaver-Burk plot for inhibition of *Mt*UGM by MS-208.
This plot shows 1/Rate vs. 1/UDP-Gal*f*, in the absence of MS-208 and the presence of MS-208 at three different concentrations, 60 µM (blue), 120 µM and 200 µM (black).

The Lineweaver-Burk plot shown in Figure 3-21 was obtained by global fitting of the data to the following linear mixed-type inhibitor equation,

$$\frac{1}{V} = \frac{K_m}{V_{max}}\left(1 + \frac{[I]}{K_i}\right)\left(\frac{1}{[S]}\right) + \frac{1}{V_{max}}\left(1 + \frac{[I]}{\alpha K_i}\right) \qquad \textbf{(2)}$$

where, V_{max} is the maximum velocity obtained for each curve at different inhibitor concentration, K_m is the Michaelis-Menten constant, S is substrate concentration, I is the inhibitor concentration. The values generated for both dissociation constants K_i, K_i', and α, from global fitting of the data were 0.13 ± 0.02 mM, 0.4 ± 0.07 mM and 2.99 respectively.

3.5.2.3 Diagnostic Dixon and Cornish-Bowden plots

Dixon and Cornish-Bowden plots were used to confirm that MS-208 is a mixed-type inhibitor of *Mt*UGM. In the Dixon plot, the lines intersected above the x-axis while in the Cornish-Bowden plot the lines intersected below the x-axis. Hence, the two diagnostic plots taken together are indicative of mixed-type inhibition.[99] Furthermore, these plots also allow the determination of K_i and K_i' values. The K_i value as determined from the Dixon plot shown in Figure 3-22A was ~ 135 µM. The K_i' value as calculated from the Cornish-Bowden plot shown in Figure 3-22B was ~ 400 µM. These values are in close agreement with the values calculated from the SigmaPlot global fitting.

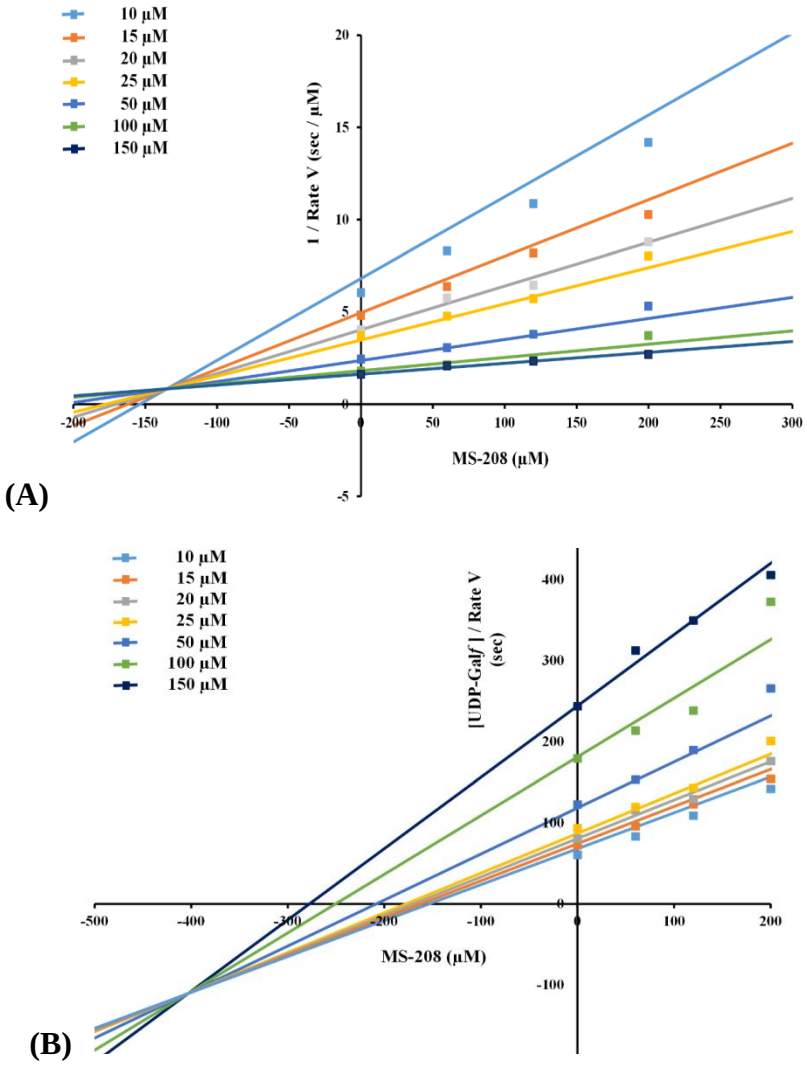


Figure 3-22: Dixon and Cornish-Bowden plots for inhibition of *Mt*UGM by MS-208.
(A) Dixon plot, showing 1/ Rate vs. Inhibitor (MS-208), for various concentrations of UDP-Gal*f*. (B) Cornish-Bowden plot, showing UDP-Gal*f* / Rate vs. Inhibitor (MS-208), for different concentrations of UDP-Gal*f*.

3.5.3 Discussion

The Michaelis-Menten plot displays decreasing V_{max} values with increasing inhibitor concentrations. In the Lineweaver-Burk plot, the lines drawn at various inhibitor concentrations are not intersecting on the y-axis. This suggests that MS-208 is not a competitive inhibitor. The best global fit to the experimental data was obtained with an equation for linear mixed-type

inhibition, using SigmaPlot. Diagnostic Dixon and Cornish-Bowden plots helped further the understanding of the mode of inhibition of *Mt*UGM by MS-208. The lines corresponding to various substrate concentrations intersect at a point above the x-axis (second quadrant) in the Dixon plot and below the x-axis (third quadrant) in the Cornish-Bowden plot, thus confirming that MS-208 is a mixed-type inhibitor.

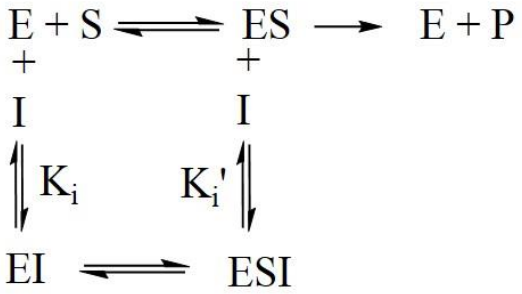


Figure 3-23: General reaction scheme for a mixed-type inhibitor.

Furthermore, these complementary plots also helped determine the K_i and K_i' values. The K_i and K_i' values obtained from these plots were ~ 135 µM and ~ 400 µM respectively. In this case, K_i, the dissociation constant of the EI complex is less than K_i', the dissociation constant of the ESI complex, meaning that MS-208 binds more tightly to *Mt*UGM alone than the *Mt*UGM: UDP-Gal*p* complex (Figure 3-23). These values are in close agreement with those calculated from the global data fit generated using the linear mixed-type inhibitor equation in SigmaPlot. Thus based on the observations from the plots it was concluded that MS-208 works as a mixed inhibitor of *Mt*UGM. Further experiments will be conducted to prove its binding to the suggested allosteric site. Furthermore, improving the inhibition by modifying the structure of MS-208, will also be considered. Some of the suggestions will be discussed in Chapter 6.

Chapter 4: UDP-arabinopyranose mutase

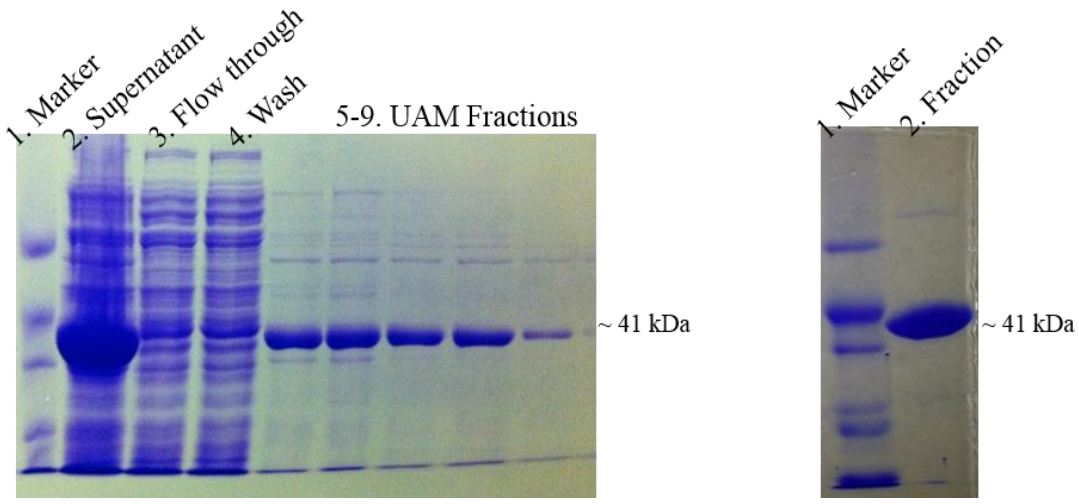
Not much information is available about the 3D structure and catalytic mechanism of plant UAMs. UAM from rice seed extracts, purified by Konishi *et al*. (2007), did not have the characteristic flavin UV-absorption peak at 450 nm. Furthermore, it was observed that *Os*UAMs were active when tested with divalent metal ions; 5 mM manganese chloride ($MnCl_2$) gave the best % activity.[18,100] All three *At*RGPs were also tested with 5mM $MnCl_2$ for mutase activity, using experimental conditions previously developed by Konishi *et al*.[66]

The structural information known about UAM is based on the experimental results and hypothesis generated to describe the possible roles of a few residues. Arg158 was believed to be the site of glycosylation of *Os*UAM1. Konishi *et al*. (2010) explained this based on their observations from experiments conducted using ^{13}C-labelled UDP-glucose to glycosylate this arginine and then analyzing the trypsin digested glycopeptide fragments by Liquid Chromatography-Mass Spectrometry (LC-MS). The alanine mutants of glycosylating arginines from *Os*UAM1 and *Os*UAM3 (R158A and R156A) exhibited poor or no activity. Moreover, other arginines that are in proximity to the site of glycosylation (*Os*UAM1 Arg165 and *Os*UAM3 Arg163) were also proved to be essential for activity. Both, removing only the guanidino nitrogen group of *Os*UAM1 (Arg165 mutating it to a lysine) or removing the side chain completely (Arg165 mutation to alanine) reduced mutase activity by ~ 12 and ~ 17-fold respectively.[101] Similarly, mutating *Os*UAM3 Arg163 to alanine (R163A) had a disastrous effect on activity (~ 50-fold reduction in activity). The sequence of rice *Os*UAMs and *At*RGPs have a highly conserved DXD-motif, characteristic of the glycosyltransferase family. This motif is hypothesized to be involved in binding the metal ion, which is necessary for activity.

In the research undertaken on UAM, as a part of this thesis work, the divalent metal ion dependency of four chosen UAMs *(OsUAM1*, *At*RGP1, *At*RGP2, and *At*RGP3) was investigated, and a comparative analysis was performed. So far, the role of the metal ion in binding the substrate and how the metal ion binds to the active site of the enzyme has not been explored in any of the plant UAMs. Hence, experiments were conducted to pick out the enzyme's metal binding region and determine the residues that may bind the metal ion in the enzyme's active site through SDM. On the basis of these experiments, a possible role for the metal ion in interacting with the substrate (UDP-Ara*p*) is also proposed.

4.1 Purification of *Os*UAM1 and mutants

*Os*UAM1 wild-type and its mutants (H273A, D110A, D111A, and D112A) required for this study were cloned into a pEHISTEV vector and transformed into *E. coli* strain BL-21 gold cells. All proteins over-expressed with the N-terminal histidine tag and were purified using nickel affinity chromatography. Figure 4-1A shows that the fractions (5-9) contain wild-type *Os*UAM1 purified along with other contaminating proteins. The fractions were concentrated and run on a Gel-filtration chromatography, to further purify *Os*UAM1, by removing other contaminating bands. The sample purity was analyzed using SDS-PAGE, and it is shown in Figure 4-1B.

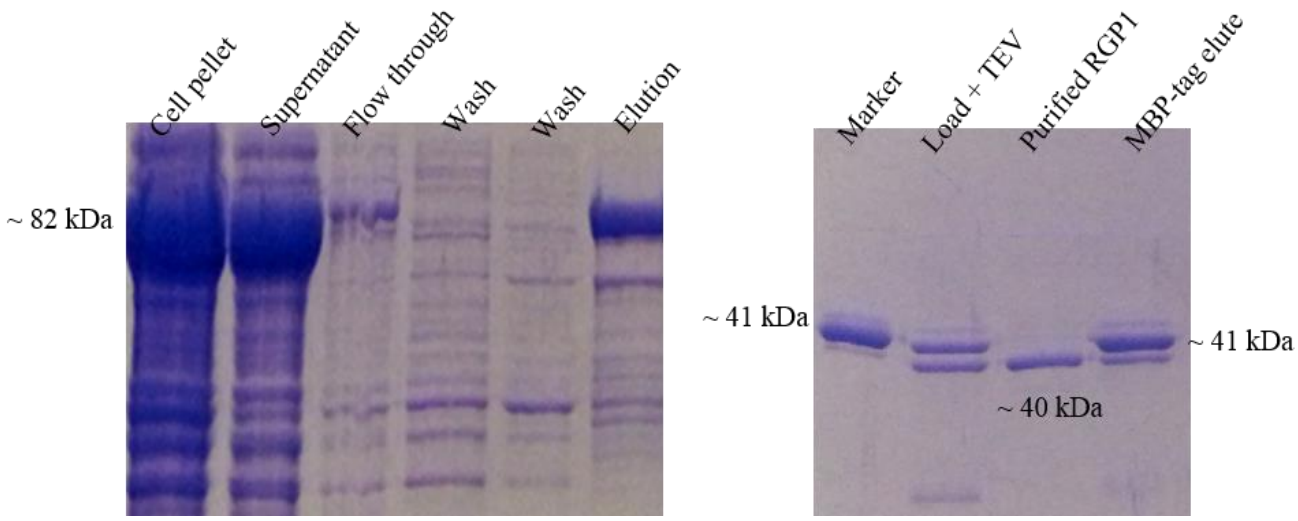


(A) **(B)**

Figure 4-1: SDS-PAGE analysis after *Os*UAM1 purification.
(A) Gel obtained after Nickel affinity chromatography (B) After further purification using Gel-filtration chromatography.

4.2 Purification of AtRGPs

AtRGPs cloned into pEHISTEV vector and transformed into *E.coli* BL-21 gold cells were unable to successfully over-express a desirable amount of soluble protein. Therefore, *AtRGPs* were cloned into pDESTHisMBP, using Gateway cloning technology. The required protein could be expressed with an N-terminal His-MBP-tag, which was cleaved using TEV-protease, due to a recognition site inserted after the His-MBP tag. Pure *At*RGPs were obtained in the flow-through, after applying the TEV-Protease treated sample onto a Ni-sepharose affinity column. The purity of the protein was assessed using SDS-PAGE. The gels obtained with *At*RGP1 are shown as an example in Figure 4-2.



(A) **(B)**

Figure 4-2: SDSPAGE analysis after *At*RGP1 purification.
(A) SDS-gel of *At*RGP1 purification after Nickel affinity chromatography column (B) After cleavage of the His-MBP tag with TEV protease.

4.3 Metal-binding studies on *Os*UAM1 and *At*RGPs

According to a previous study conducted by Konishi *et al.* (2007), rice mutase was activated in the presence of divalent ions (as shown in Figure 4-3). They observed enzyme activity when no divalent metal ions were added to their assay. The value obtained was approximated as 100% relative activity. Based on this approximation, they discovered that rice mutase activity almost doubled in the presence of 5 mM $MnCl_2$, but the same concentration of other divalent cations such as $MgCl_2$ and $ZnCl_2$ displayed no significant increase in activity. Other divalent metal solutions such as $CaCl_2$, $CuSO_4$, and $CoCl_2$ had an adverse effect on relative activity (decreased % relative activity from that obtained with no metal), as shown in Figure 4-3.

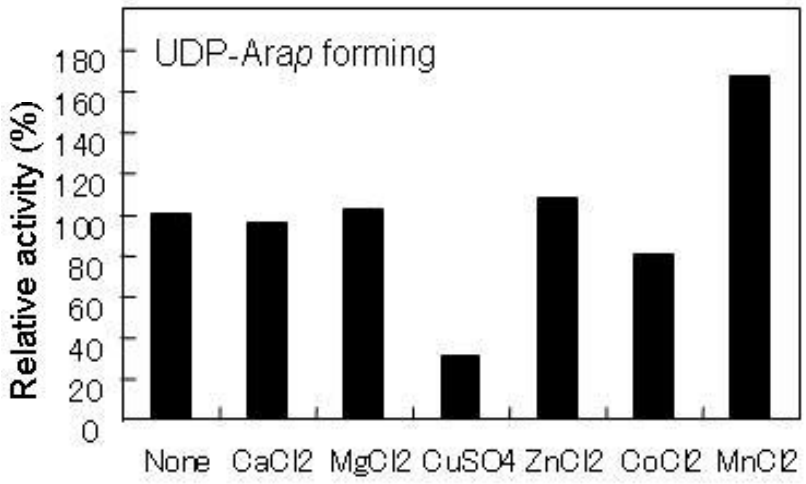


Figure 4-3: Percent relative activity of UAM with divalent metal ions as studied by Konishi *et al.* (2007).*
*Permission to reuse figure attached at the end of this thesis.

4.3.1 Percent activity of *Os*UAM1 with varying concentrations of divalent metal ions

The results obtained here are from reactions performed with constant amounts of substrate and enzyme concentrations, and a manganese concentration ranged from 0 to 5 mM. The results showed that 5 mM Mn^{2+} had an inhibitory effect on enzyme activity and the activity increased with a decrease in Mn^{2+} concentration. The % relative activity obtained at 80 µM concentration of Mn^{2+} was much higher, almost double compared to the activity of the enzyme observed at 5 mM Mn^{2+}. There was an increase in activity observed from 0 µM (11 %) to 80 µM (100 %) Mn^{2+} and 85 - 100 % relative activity was achieved with a Mn^{2+} concentration between 40 -200 µM (Figure 4-4A). Excess Mn^{2+} ion concentration, greater than 320 µM, displayed an inhibitory effect on *Os*UAM1 activity. A second plot obtained with *Os*UAM1 pre-treated with EDTA to remove residual metal, also showed a similar trend (Figure 4-4B). No activity was observed at 0 µM Mn^{2+}, which demonstrates that the enzyme is inactive in the absence of any divalent metal ion. Therefore, Mn^{2+} is essential for *Os*UAM1 activity and is not an activator of the enzyme, as previously stated by Konishi *et al.* (2007).

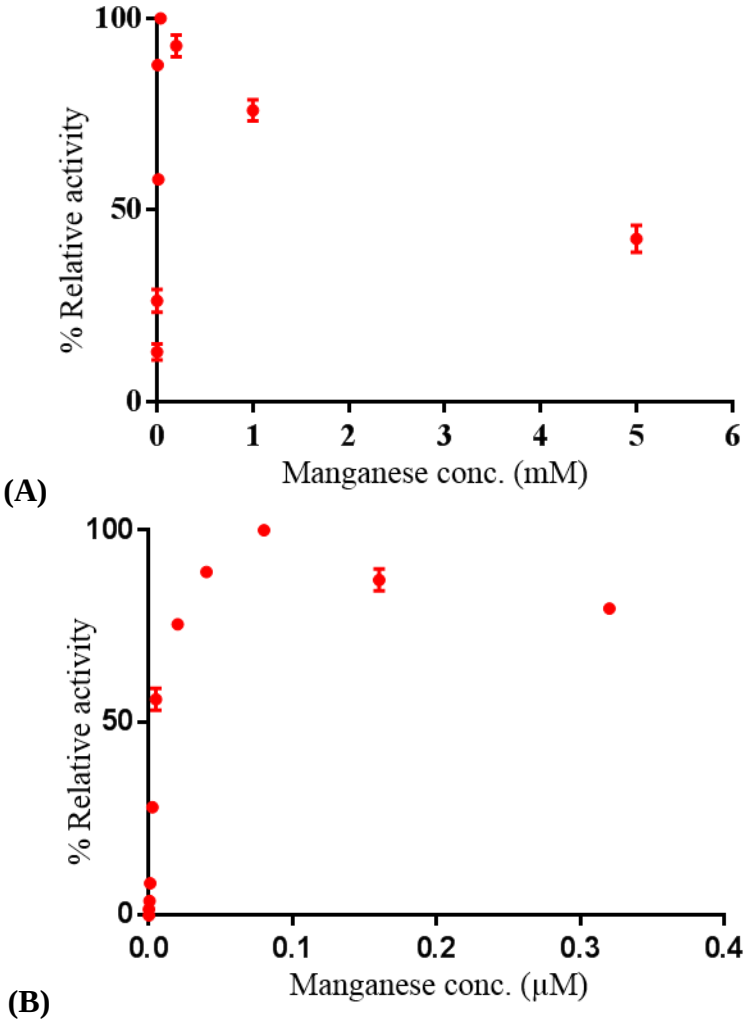


Figure 4-4: Plots showing the change in % Relative activity with Manganese concentration of *Os*UAM1.
(A) Percent relative activity vs. Manganese concentration up to 5 mM of *Os*UAM1 (B) Percent relative activity vs. Manganese concentration up to 320 µM of *Os*UAM1 samples pre-treated with EDTA.

4.3.2 Percent activity of *Os*UAM1 with other divalent metals

Other divalent metals such as Mg^{2+}, Zn^{2+}, Co^{2+}, Ca^{2+}, and Cu^{2+} were used to study the % activity of *Os*UAM1. This was done to understand the effect of these metal ions on mutase activity and to make comparisons with the data obtained by Konishi *et al.* (Figure 4-5).

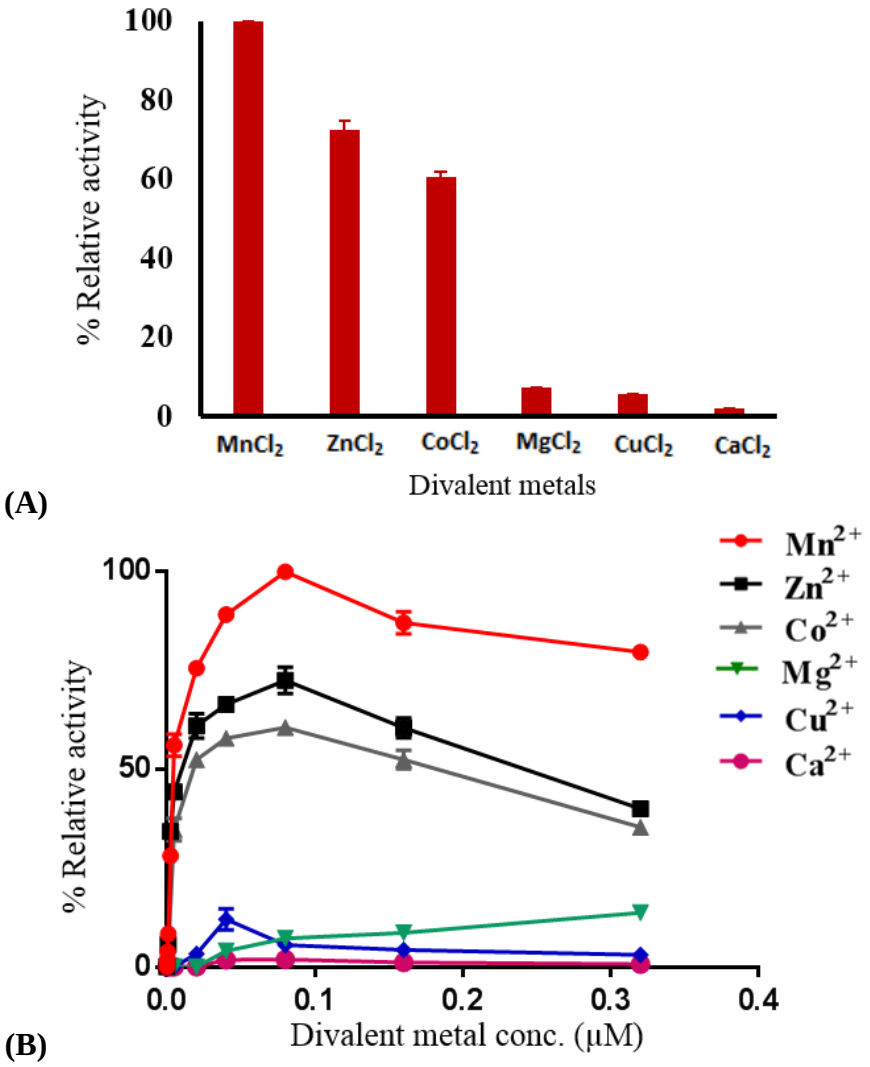


Figure 4-5: Percent activity of *Os*UAM1, in the presence of other divalent metal ions.
(A) The percent relative activity of *Os*UAM1, in the presence of other divalent metal ions at 80 µM concentration in the reaction mixture (B) Percent relative activity of *Os*UAM1with divalent metal ions in the range 0 - 320 µM.

It was observed that Zn^{2+} and Co^{2+} ions were able to show significantly lower % relative activity (close to 70 % for Zn^{2+} and 60% for Co^{2+}) compared to the % activity obtained with Mn^{2+} at 80 µM concentration (Figure 4-5A). However at 5 mM, they had displayed a very low or no effect on activity. Moreover, other divalent metals such as Mg^{2+}, Ca^{2+} and Cu^{2+} displayed low activity, less than 10% increase in activity at 80 µM.

Another plot was constructed by varying the divalent metal concentrations between 0 µM and 320 µM, to see if a trend similar to that of Mn^{2+} was displayed, using *Os*UAM1 pre-treated with EDTA. It was observed that Zn^{2+} and Co^{2+} followed a similar trend with % relative activity increasing to a maximum at 80 µM metal concentration and decreasing as metal ion concentration was increased to ~ 300 µM or higher (Figure 4-5B). There was no significant rise in activity in the case of other divalent metals like Mg^{2+}, Cu^{2+}, and Ca^{2+}. The % relative activity of *Os*UAM1 remained less than 10% with all the three divalent metal ions at various concentrations, compared to the values obtained with Zn^{2+} and Co^{2+}.

4.3.3 Metal binding studies on Reversibly Glycosylated Protein 1

A similar trend of lower % activity with higher concentrations of Mn^{2+} was also noted with *At*RGP1. However, in this case, maximum % activity was obtained at 40 µM concentration of Mn^{2+} (Figure 4-6A & B). Other trends were more similar to what was already observed with *Os*UAM1. At 40 µM, Zn^{2+} and Co^{2+} produced slightly greater than 60% activity (Figure 4-6B). The other divalent ions showed lower activity at this concentration and negligible % activity at concentrations between 0 and 320 µM. As already described with Mn^{2+}, *At*RGP1 displayed a decrease in activity at concentrations higher than 40 µM with Zn^{2+} and Co^{2+} (Figure 4-6C).

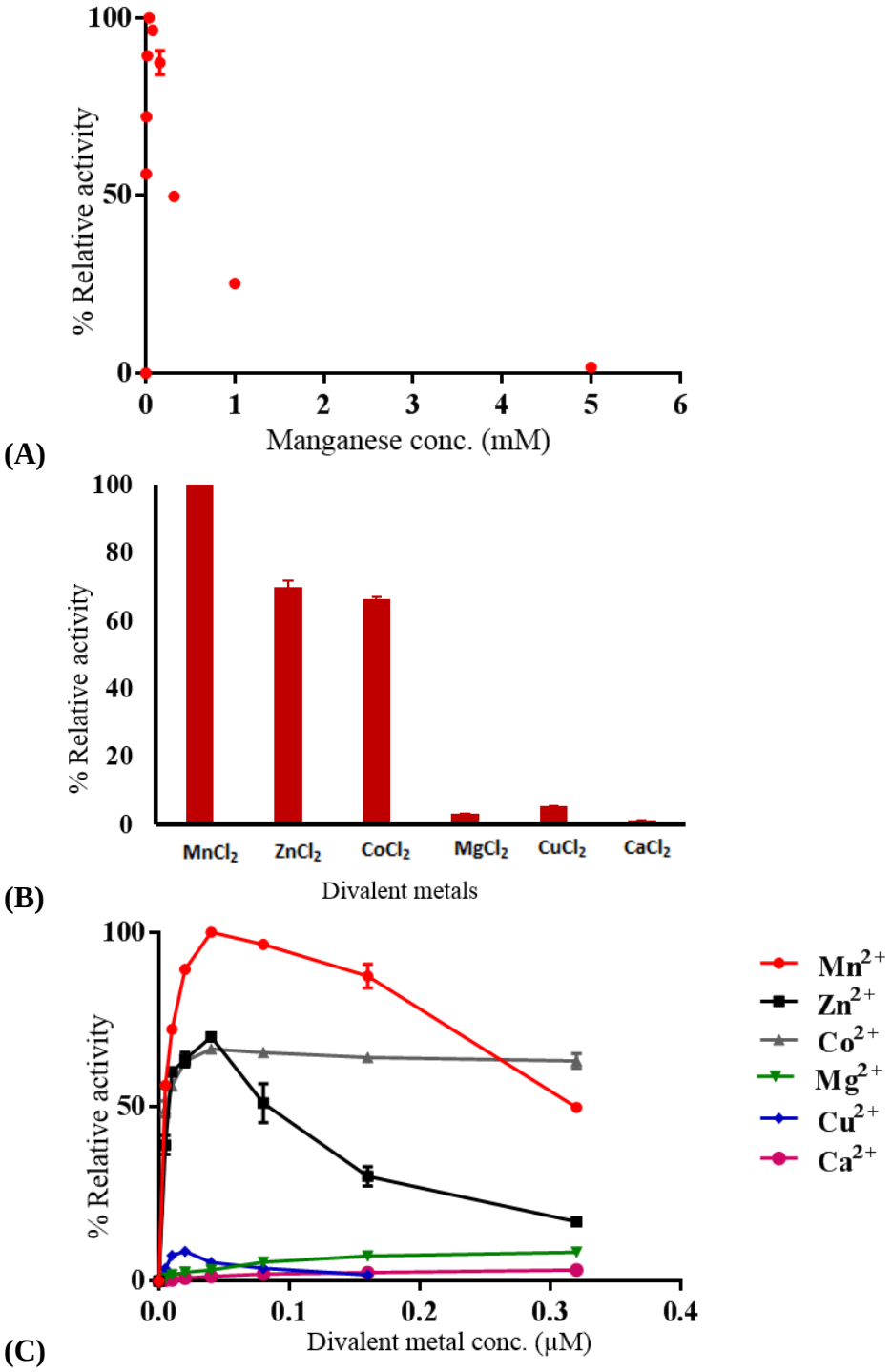


Figure 4-6: Metal binding studies on *At*RGP1.
(A) Percent relative activity vs. Manganese concentration up to 5 mM (B) Percent relative activity of *At*RGP1, with other divalent metal ions at 40 µM in the reaction mixture (C) Percent relative activity of *At*RGP1with divalent metal ions with concentrations ranging from 0 - 320 µM.

4.3.4 Reversibly Glycosylated Proteins 2 and 3 (*At*RGP2 and *At*RGP3)

*At*RGP2 and *At*RGP3 both showed similar trends with Mn^{2+} and other divalent metal ions. In the case of *At*RGP2, the maximum activity was observed at 40 µM (Figure 4-7A & B); Zn^{2+} and Co^{2+} were the other two divalent metal ions that displayed 60 % activity at 40 µM (Figure 4-7B). Though unlike the trend seen thus far, at this concentration, Co^{2+} had a slightly higher activity (3-5 %) than Zn^{2+}. Trends with other divalent ions were similar to *At*RGP1 and *Os*UAM1 from concentrations between 0 - 320 µM (Mg^{2+}, Cu^{2+}, and Ca^{2+} had negligible % activity in this concentration range). (Figure 4-7C).

The graphs for *At*RGP3 are shown in Figure 4-8. Here again, Mn^{2+} displayed the best % activity at 40 µM (Figure 4-8A and B), while divalent ions such as Zn^{2+} and Co^{2+} displayed the next best activity, as seen with *At*RGP1 (Figure 4-8B). In the case of *At*RGP3, there were no aberrations in the trend already established by the other mutases in this study.

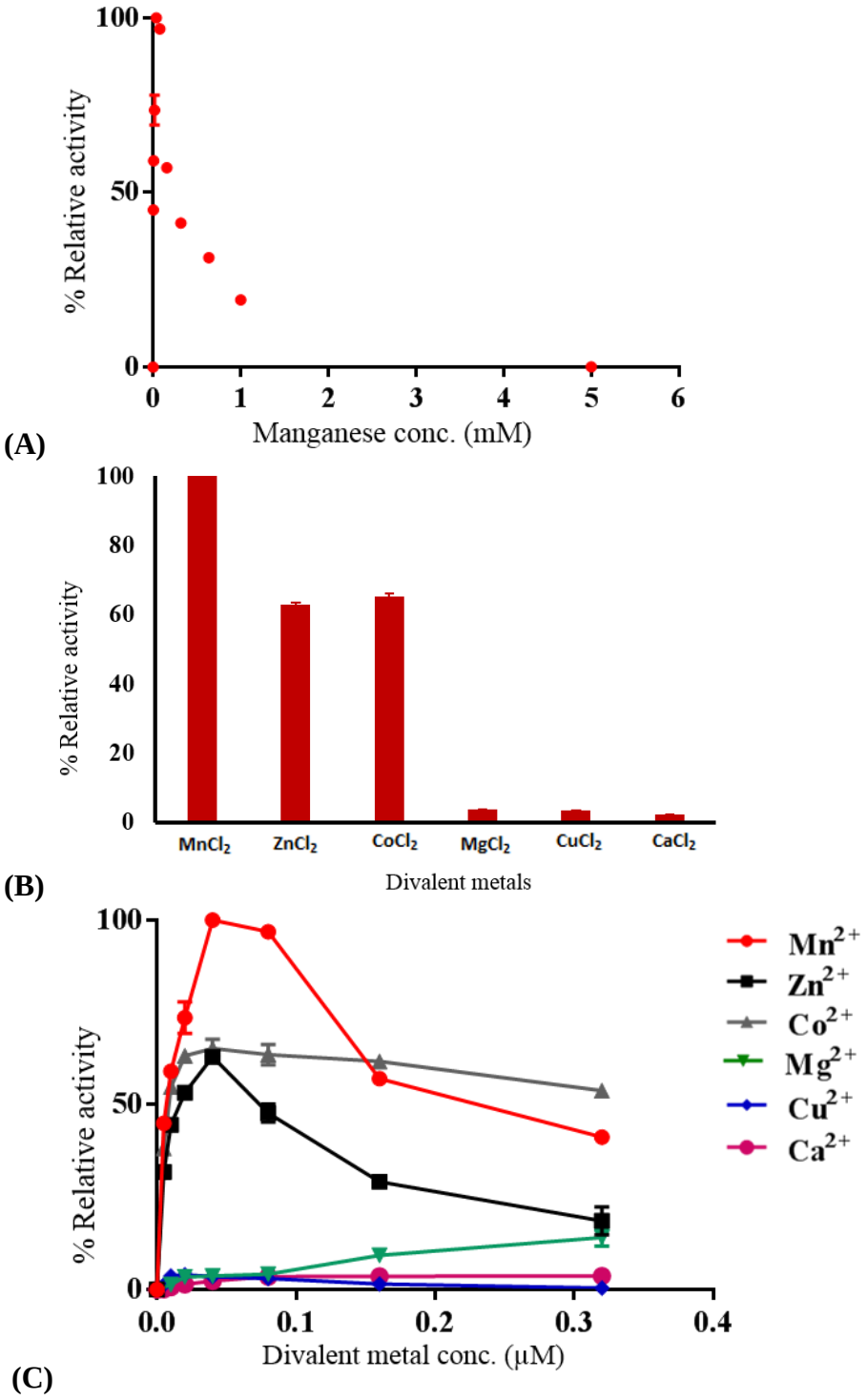


Figure 4-7: Metal binding studies on *At*RGP2.
(A) Percent relative activity of *At*RGP2 vs. Manganese concentration up to 5 mM (B) Percent relative activity of *At*RGP2, with other divalent metal ions at 40 µM in the reaction mixture (C) Percent relative activity of *At*RGP2 with divalent metal ions in the range, 0 - 320 µM.

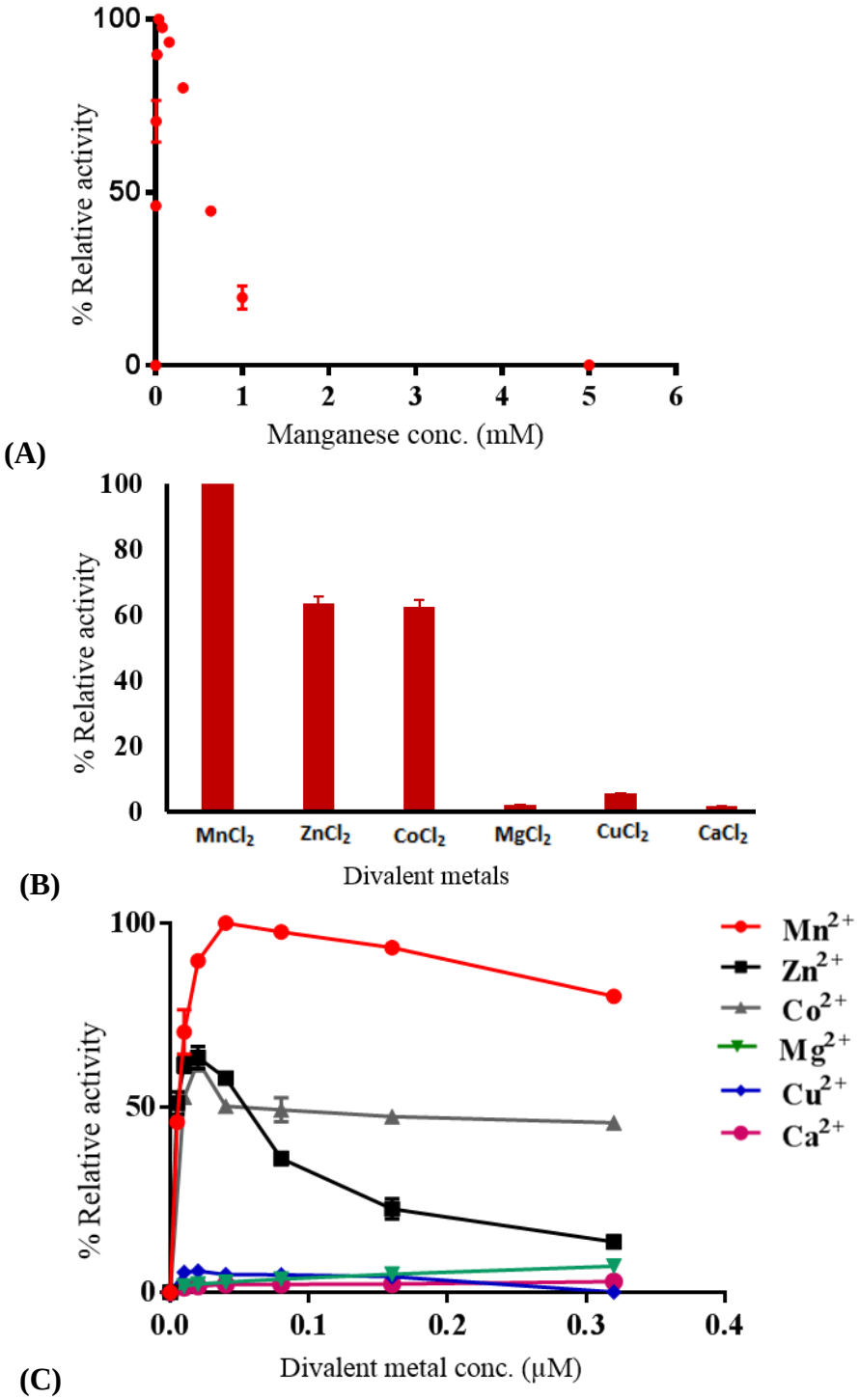


Figure 4-8: Metal binding studies on *At*RGP3.
(A) Percent relative activity of *At*RGP3 vs. Manganese concentration up to 5 mM (B) Percent relative activity of *At*RGP3, with other divalent metal ions at 40 µM in the reaction mixture (C) Percent relative activity of *At*RGP3 with divalent metal ions in 0 - 320 µM range.

4.4 UAM crystal trials

Crystallization experiments were conducted with *Os*UAM1 and *At*RGP1 to obtain structural information. The results from crystal screening of both these enzymes will be discussed here. Purified *Os*UAM1, concentrated to ~ 6.5 mg/ml in 25 mM Bicine pH 8.5 and 1 mM Manganese chloride, was used to set up crystal trials with UDP-Gal*p* (due to commercial unavailability of the actual substrate UDP-Ara*p*). For the initial trials, commercially available crystal screens were used to set up microbatch plates at room temperature and 4 °C. Most of the hits obtained were either spherulites or microcrystalline material; none of the conditions produced single crystals. The spherulites were obtained overnight in multiple conditions with the 2-methyl 2-pentanediol (MPD) screen having 30 - 40% (v/v) MPD concentration in two different buffers; HEPES buffer pH 7.0 and Tris buffer pH 8.0 (Figure 4-9A and B). However, the best hits were observed within three to four days in the PACT screen conditions #2-6, which had thin clusters of needles growing from the initially formed spherulites (Figure 4-9C). These conditions had 0.1 M SPG buffer (having succinic acid: sodium dihydrogen phosphate: glycine in a 2:7:7 molar ratios), pH 5.0, 6.0, 7.0, 8.0 and 9.0, respectively, and 25 (v/v) polyethylene glycol (PEG) 1500.

The conditions were further optimized with the use of grid screens while varying the pH between 4.0 and 9.0 and the PEG 1500 concentration between 5 and 40 % (v/v). The optimization grid screens were performed with both microbatch and hanging drop vapour diffusion techniques. Also, additive screen, having 96 different ingredients (containing different anions, cations, and detergents) was used. For performing the additive screen, the 96 different ingredients were mixed with the PACT screen conditions #2-6, in a 9:1 ratio, (to make 96 different crystallization conditions), before setting up the microbatch plate. The experiments were also repeated at various temperatures, to see if the changes in conditions would increase the likelihood of obtaining

crystals. Most of the results observed did not give single crystals nor did they improve the quality of crystalline material significantly. A grid screen and additive optimization were also performed on the spherulites obtained in the MPD screen. Yet again, only spherulites could be reproduced with 15-25% (v/v) MPD, with 50 mM potassium iodide and HEPES buffer pH 7.5, using both microbatch and vapor diffusion techniques.

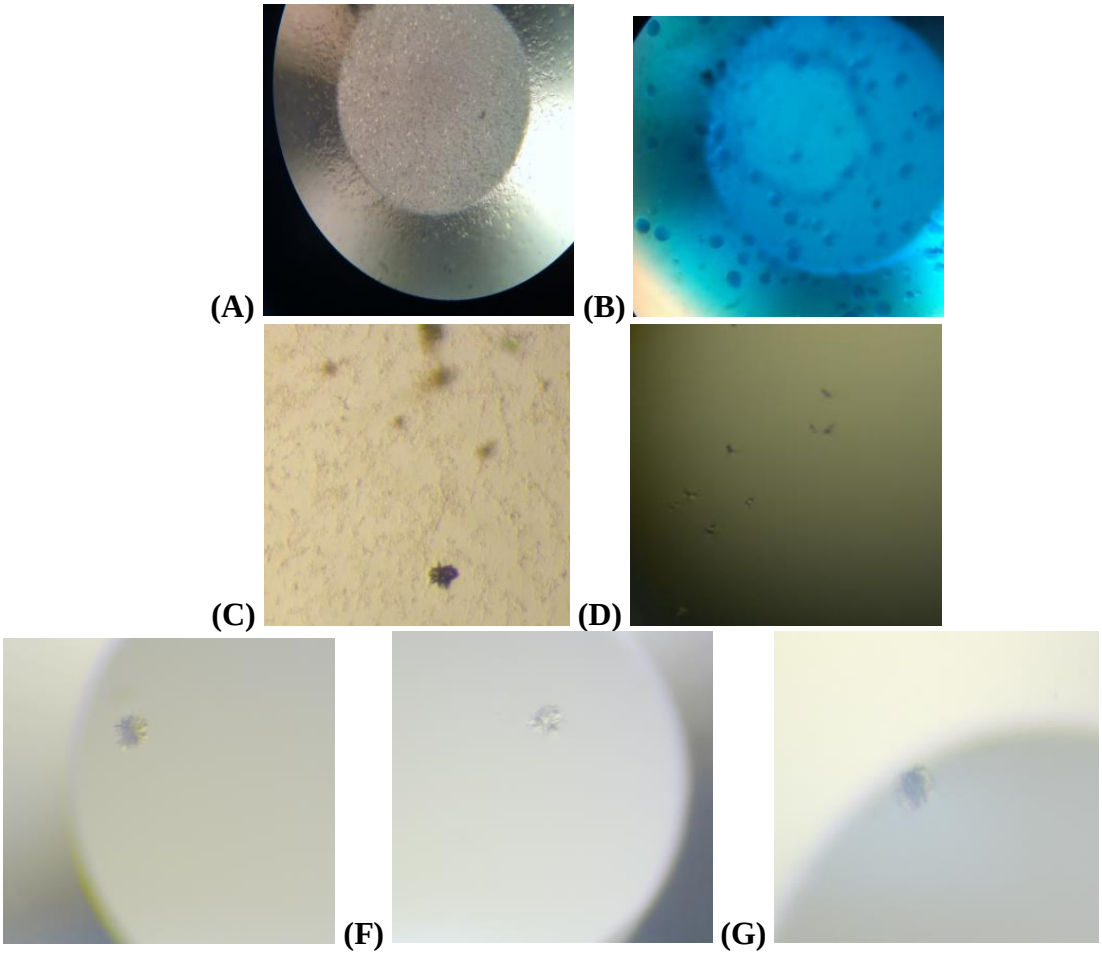


Figure 4-9: Microcrystalline hits observed for *Os*UAM1 crystal screens.
(A) Microcrystalline material observed in the MPD screen for *Os*UAM1 (B) Spherulites of *Os*UAM1 which take up the Izit dye (C) *Os*UAM1 microcrystalline material obtained from hanging drops containing 20% (v/v) MPD, potassium iodide and HEPES buffer pH 7.5 (D) Small crystals obtained from seeding experiments (E) *At*RGP1 needle cluster from PACT screen # 6 (F) *At*RGP1 crystalline material obtained from microseeding in the same condition (G) *At*RGP1, plate-like clusters from microseeding.

Seeding techniques were also used to improve crystal quality. The needle clusters, obtained from the PACT screen were used in performing both microseeding (including streaking) and macroseeding. The crystallization drops were set up using the hanging drop technique and the drops were streaked or macroseeded after sufficient equilibration time (3-4 hours). Small tiny crystals could be obtained within a day, but they did not grow any further to form larger single crystals (Figure 4-9D).

*At*RGP1 was purified and concentrated to ~ 7.0 mg/ml in 25 mM Bicine pH 8.5 and 1 mM manganese chloride. Co-crystallization trials of *At*RGP1 conducted with UDP-Gal*p* produced initial hits in the same conditions (PACT screen conditions #2 - 6), as that of *Os*UAM1 (Figure 4-10E). The needle clusters obtained in this case were slightly longer than those obtained with *Os*UAM1. Grid screens were set up around the condition for optimization. Additive screens and detergent screens were also performed. Hanging drops were also set to see if the needle clusters were reproducible. Seeding experiments were also carried out with this enzyme, using the microseeding technique described in Chapter 2. Generally, the seeding trials yielded slightly better crystals. In some cases, clustered plate-like crystals were also identified. The best results were thin plates in 0.1 M SPG buffer pH 7.0 and 15% (v/v) PEG 1500 (Figure 4-9F & G). The plates were tested for quality; however diffraction results yielded very low resolution.

4.5 Results from modeling studies

Since diffraction quality single crystals of UAM proved to be a difficult task to accomplish, modeling studies were performed to gain some insights into the active site residues involved in substrate and metal binding. There were no known 3D structures for any of the plant UAM or any other proteins with a modest sequence identity (30 - 40 % sequence identity) to UAM, reported in

the literature. Hence, models had to be generated based on the protein sequence of the enzyme. A sequence alignment performed on all four sequences (*Os*UAM1, *At*RGP1, *At*RGP2 & *At*RGP3) showed that all the sequences had greater than 85% identity (Figure 4-10B). Therefore, *Os*UAM1 was chosen for modeling studies.

Two different structure prediction (modeling) programs I-TASSER, and GalaxyWEB were used to predict a model for *Os*UAM1.[87,88,102,103] A description of the strategy utilized by these modeling softwares to predict the model has been included in the Supplementary section S.4. Both these programs produced five structural models for the *Os*UAM1 sequence. Models obtained from these online modeling programs, were manually compared against each other (uploading PDB coordinates generated in PyMOL) to determine similarities in regions predicted to form secondary structures (Figure 4-10A), and also to make an overall comparison of the 3D structure. The comparative analysis of each of the five models, predicted by the programs, showed that regions of the *Os*UAM1 sequence picked to form α-helices and ß-strands were consistent overall. The regions picked to form loops were often slightly different among the models, but generally consistent. The predicted models were also compared to analyze differences in the overall structural architecture. The most consistent models generated by both these programs with minimal overall differences in secondary structural and 3D architecture were selected and are shown in Figure 4-11A & B.

(A)

```
MAGTVTVPSASVPSTPLLKDELDIVIPTIRNLDFLEMWRPFFQPYHLIIVQDGDPTKTIR     60
VPEGFDYELYNRNDINRILGPKASCISFKDSACRCFGYMVSKKKYVFTIDDDCFVAKDPS    120
GKDINALEQHIKNLLSPSTPFFFNTLYDPYREGADFVRGYPFSLREGAKTAVSHGLWLNI    180
PDYDAPTQMVKPRERNSRYVDAVMTVPKGTLFPMCGMNLAFDRDLIGPAMYFGLMGDGQP    240
IGRYDDMWAGWCMKVICDHLSLGVKTGLPYIWHSKASNPFVNLKKEYKGIFWQKDIIPFFQN 302
ATIPKECDTVQKCYLSLAEQVREKLGKIDPYFVKLADAMVTWIEAWDELNPSTAAVENGKAK 364
```

(B)

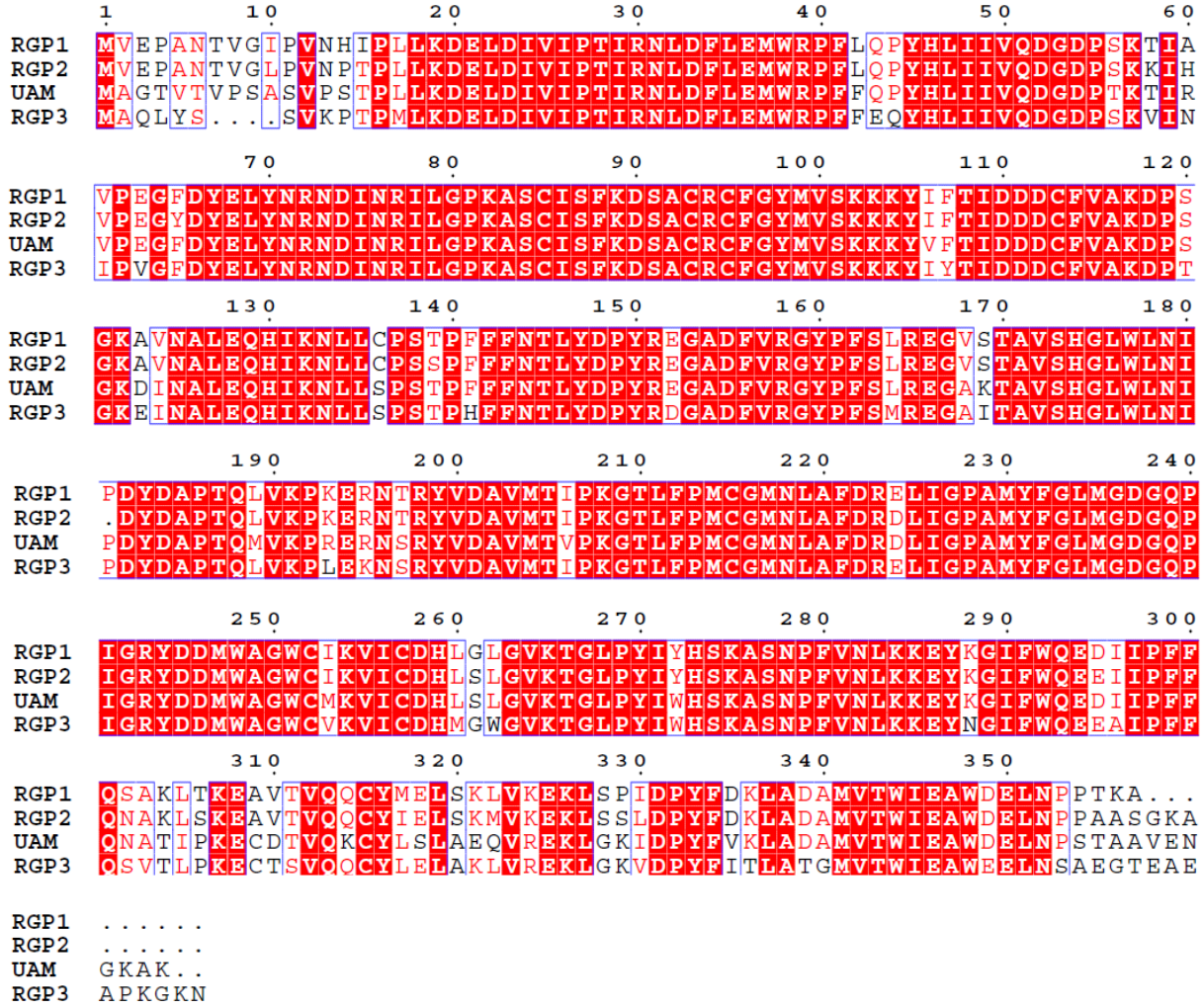


Figure 4-10: High sequence identity between *Os*UAM1 and *At*RGPs.
(A) *Os*UAM1 sequence showing regions predicted to form helices (in red), β-strands (in blue) and coils (in black), as predicted by PredictProtien software. (B) Sequence alignment of OsUAM1 (UAM) & AtRGP sequences (RGP1, RGP2 & RGP3). The sequence alignment was performed using ESPript (Version 3.0)

Further analysis of the *Os*UAM1 model revealed information about the Structural Classification of Proteins (SCOP) that UAM may fall under. UAM was predicted to belong to the class Alpha and beta proteins (a/b). Among the proteins classified under this category, UAM comes under the nucleotide-diphospho sugar transferases fold and superfamily of proteins. The commonly observed prominent features of the secondary structural elements of a protein structure that belong to this fold are, the presence of a β-sheet layer, sandwiched between two layers of α-helices (Figure 4-11). One other notable feature of the sandwiched seven strand β-sheet layer (order of β-strands 3214657), is that strand 6 is antiparallel to the others. UAM falls under the glycosyltransferase family of proteins, which bind a metal co-factor through a DXD-motif and catalyze reactions on nucleotide-diphospho sugars.

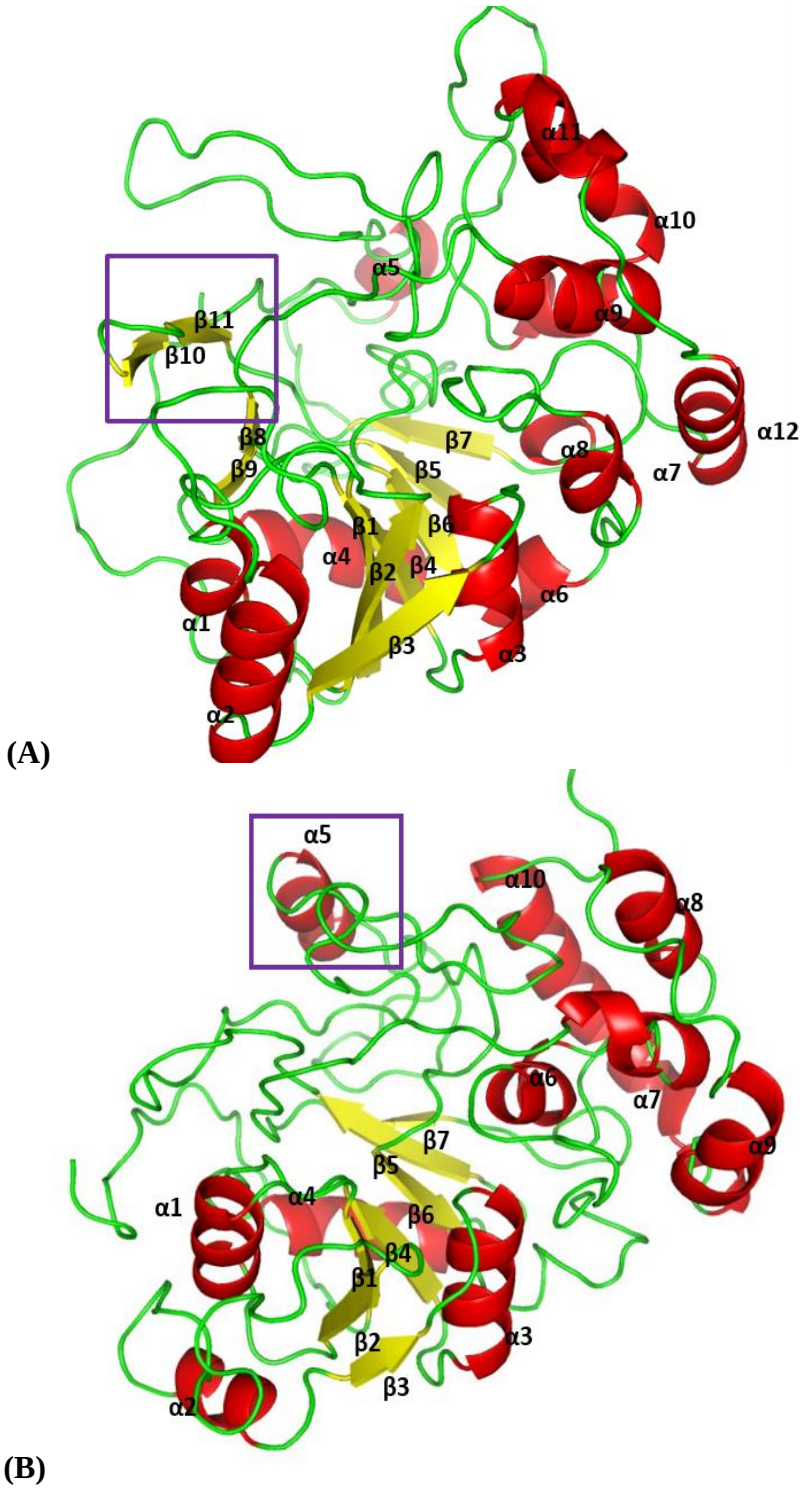


Figure 4-11: Models for *Os*UAM1, predicted by GalaxyWeb and I-TASSER.
(A) *Os*UAM1 model predicted by GalaxyWeb, highlighting the regions different from the I-tasser model in a purple box (B) *Os*UAM1 model as predicted by I-TASSER, highlighting the regions distinct from the GalaxyWeb model in a purple box.

In addition to generating a model, I-TASSER was also able to predict the functional analogs of the model, by using global and local structural similarity search. In the case of the *Os*UAM model, I-TASSER chose Chondroitin polymerase, a bifunctional glycosyltransferase that can bind to UDP-Glucuronic acid (UDP-GlcUA) or UDP-GalNAc and alternatively transfer GlcUA or GalNAc moiety to catalyze the elongation of the chondroitin chain.[104] Like UAM, Chondroitin polymerase also uses Mn^{2+} as a cofactor and UDP-based substrates and has two glycosyltransferase domains. The software also predicts the binding site of UAM by docking UDP-GalNAc (one of the substrates in the crystal structure of Chondroitin polymerase) in the active site of the predicted *Os*UAM model. This process aids in picking out a few active site residues that may bind and interact with the substrate of UAM (UDP-Ara*p*), as shown in Figure 4-12A & B.

Based on the *Os*UAM1 model and its comparison with Chondroitin polymerase, the following residues predicted by I-TASSER may be able to bind the substrate uridine, sugar, and the metal ion cofactor. Residues such as Thr28, Ile29, Asn31, and Asp52 may interact with the uridine portion of the substrate (Figure 4-13A) while residues such as Gly242, Arg243, Asp245, and Asp246 were predicted to interact with the sugar moiety of the substrate. The most important finding regarding this research is the site proposed to bind the metal ion. Residues Asp110, Asp111, Asp112, and His273 were predicted to be in the position to bind the metal ion (Figure 4-12C). The model is in agreement with the hypothesis that, like other members of the glycosyltransferase family, UAM uses its metal ion cofactor to bind and stabilize substrate. The *Os*UAM1 model suggests that the above-mentioned residues bind the metal, holding it in place for the incoming substrate. As part of this study, SDM was performed on these residues to further our understanding of the metal binding site.

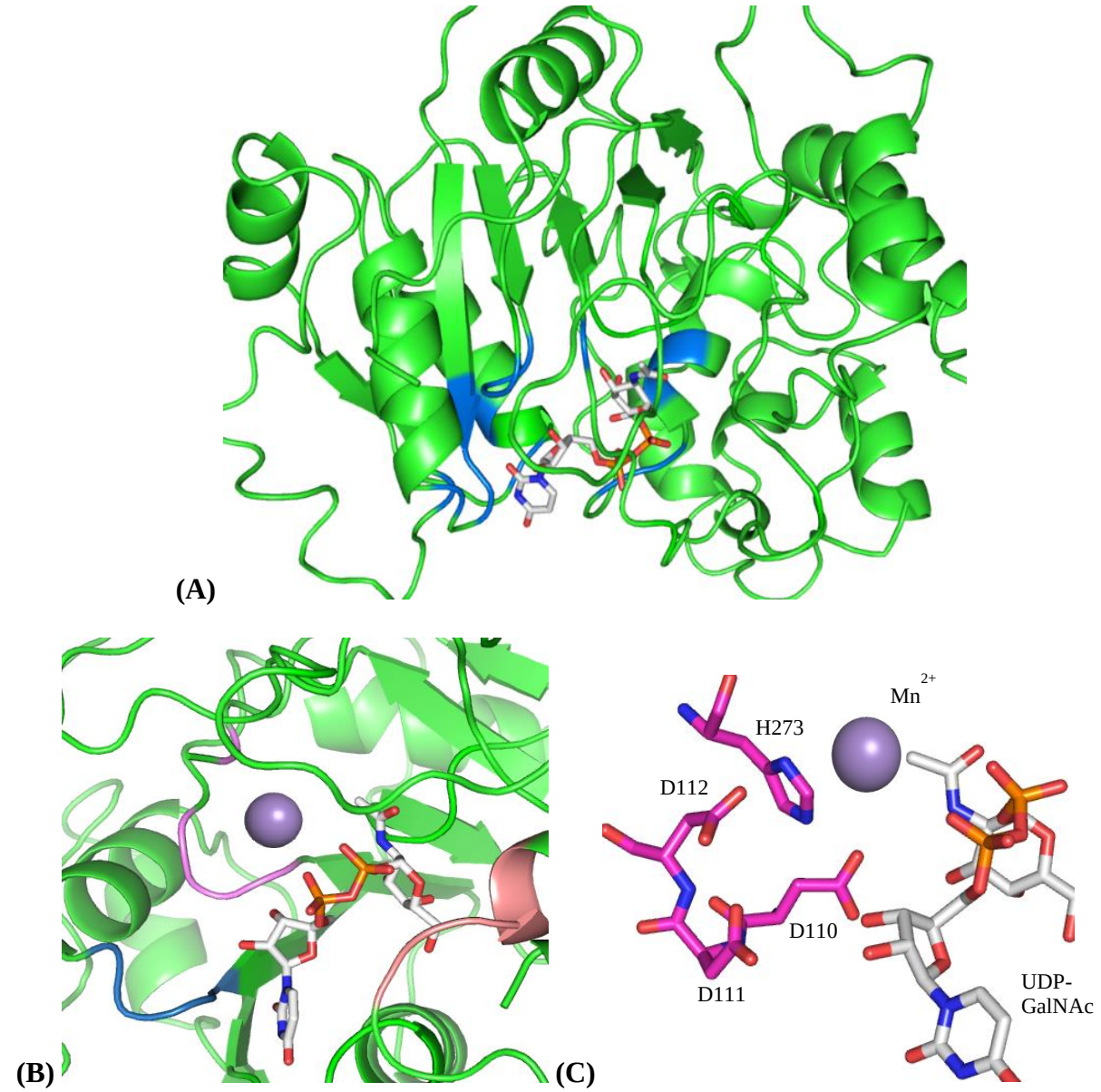


Figure 4-12: I-TASSER model of *Os*UAM1 with UDP-GalNAc in its active site and predicted uridine, metal and sugar regions of the substrate.
(A) I-TASSER model of *Os*UAM1 with UDP-GalNAc in its active site (B) Predicted regions of I-TASSER model that can bind uridine (blue), metal (magenta) and sugar (salmon) regions of the substrate. (C) Close-up view of predicted metal-binding site of *Os*UAM1

4.6 Mutation studies on *Os*UAM1

*Os*UAM1 and *Os*UAM3 contain a DDD-motif in their sequence (Asp110, Asp111, and Asp112 in *Os*UAM1; Asp108, Asp109, and Asp110 in *Os*UAM3); in the case of *Os*UAM2, which

does not perform the mutase function, the third aspartic acid in the motif is replaced by an asparagine (DDN). Mutating Asp112 in *Os*UAM1 to asparagine (D112N) deleted the mutase activity in the enzyme. However, mutating Asn99 in *Os*UAM2 to an aspartic acid (N99D) alone did not succeed in converting *Os*UAM2 into a mutase.[101] As part of this research, four residues identified by the *Os*UAM1 model as having a role in binding the metal cofactor, were mutated to alanines so as to analyze their effect on the enzyme activity. From the model, since His273 was in close vicinity to the DDD-motif, the possible role of this residue in metal binding was also further explored, along with the residues that make up the DDD-motif.

4.6.1 Kinetic study on *Os*UAM1 wild-type

The reaction catalyzed by UAM is reversible, with the equilibrium favoring UDP-Ara*p* formation. The optimal pH for the formation of UDP-Ara*p* was between 5.5 and 6.0 and that for UDP-Ara*f* formation was between pH 7.0 to 7.5.[18] The data obtained from our HPLC-based kinetic assays were fit to a non-linear regression fit and the kinetic parameters were determined. *Os*UAM1 had a K_m of 12 ± 1 μM for UDP-Ara*f*, and a k_{cat} value of 3.2 ± 0.1 sec^{-1}. The calculated specificity constant (k_{cat}/K_m) was 0.26 ± 0.1 $s^{-1}\mu M^{-1}$ (Figure 4-13A).

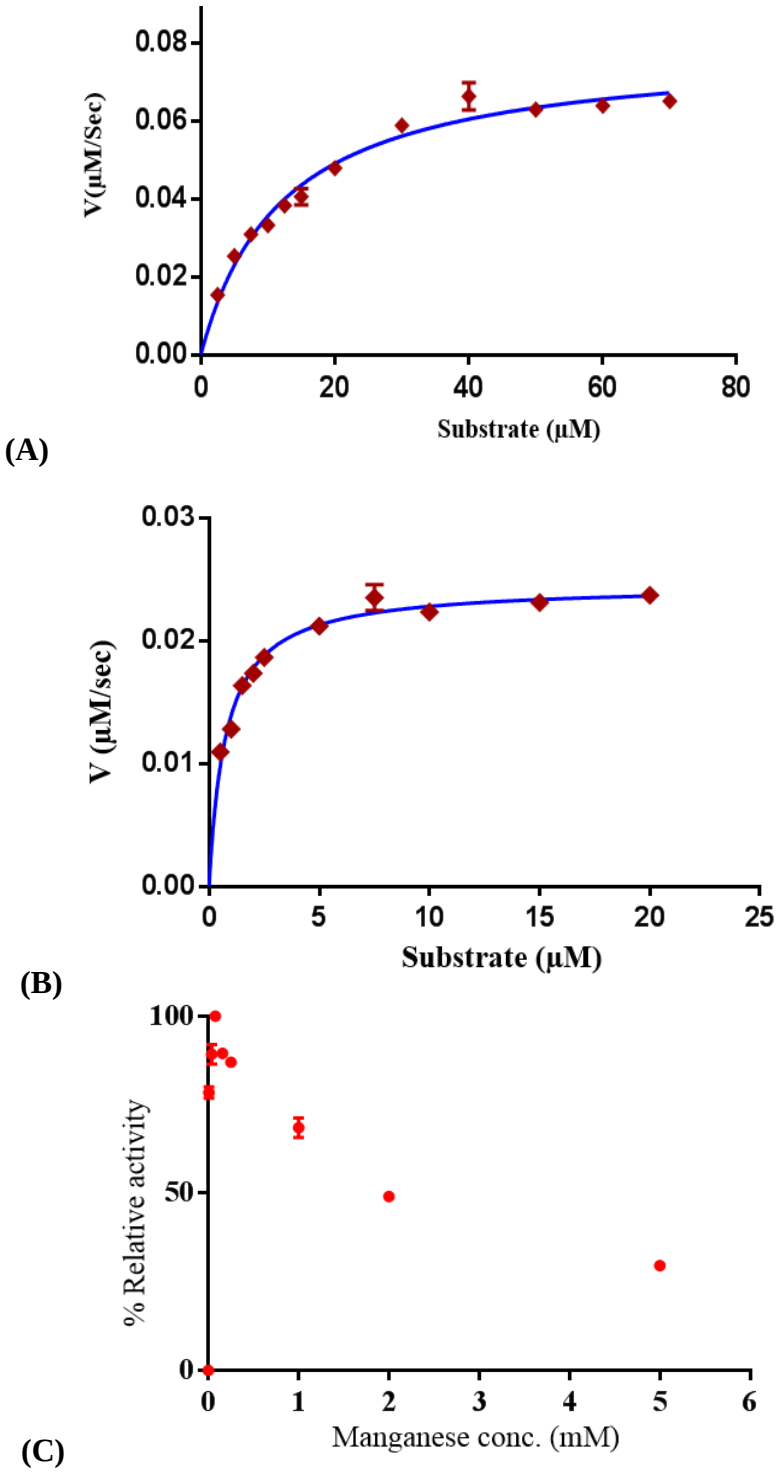


Figure 4-13: **Kinetic curves for *Os*UAM1 wild-type and H273A mutant.**
Michaelis-Menten Kinetic curve for (A) Wild-type *Os*UAM1 (B) H273A mutant and (C) Plot of % Relative activity vs. Manganese concentration for the H273A mutant.

4.6.2 Analysis of *Os*UAM1 mutants - Role of His273

*Os*UAM1 H273A mutant was created by SDM; over-expressed and purified using the same procedure described for the wild-type enzyme. Kinetic analysis was performed to see if this mutation would have any effect on enzyme activity. In general, the mutant displayed lower activity than the wild-type *Os*UAM1. A Michaelis-Menten curve was obtained by plotting rate against substrate (UDP-Ara*f*) concentration. The K_m and k_{cat} values, calculated from the Michaelis-Menten plot (Figure 4-13B) were 0.76 ± 0.05 µM and 0.05 ± 0.001 sec^{-1}, respectively. The specificity constant for this mutant was $0.6 \pm 2.0 \times 10^{-2}$ units. A decrease in both K_m and k_{cat} was noted when compared to the wild-type enzyme (Table 4-1), which meant that the mutant required less substrate to achieve half-maximum activity, but its efficiency was also lower. The metal binding assays conducted showed that this mutant followed a similar trend as that of the wild-type enzyme when treated with various concentrations of Mn^{2+}. Maximum activity was observed at ~ 80 µM Mn^{2+} and about 80 % or higher activity in the 40 - 200 µM range and a decrease in activity at concentrations greater than 320 µM (Figure 4-13C). Since this mutation did not inactivate the enzyme, it can be argued that His273 may not be directly involved in coordinating the metal ion.

4.6.3 Analysis of *Os*UAM1 mutants - Role of DDD-motif

Like the H273A mutant, three alanine mutants of the *Os*UAM1 DDD-motif were created by SDM and purified adopting the same procedure followed for the wild-type enzyme. The purified mutants were concentrated to ~ 1 mg/ml and treated with EDTA to remove any bound metal. Kinetic assays were performed with these mutants, using the same procedure followed for the wild-type enzyme. All three mutants (*Os*UAM1 D110A, *Os*UAM1 D111A, and *Os*UAM1 D112A) were inactive when tested without the addition of Mn^{2+}. Moreover, increasing the metal

ion concentration to 80 µM (concentration of Mn^{2+}, which gave maximum activity in wild-type and H273A mutant) in the reaction mixture, did not improve activity. Furthermore, no conversion of the substrate to product observed, when the reaction time was increased (up to 10-fold). Also, increasing the mutant enzyme concentration in the reaction by 10-fold did not restore activity in these mutants. The complete loss of activity suggests the inability of the DDD-motif alanine mutants to bind the Mn^{2+} cofactor as well as the wild-type enzyme unless the mutations led to detrimental changes to the secondary structure of *Os*UAM1.

Table 4-1: Kinetic parameters of wild-type *Os*UAM1 and its mutants.

Enzyme	K_m (µM)	k_{cat} (s^{-1})	k_{cat} / K_m ($s^{-1}µM^{-1}$)
*Os*UAM1(wild-type)	12 ± 1	3.2 ± 0.1	0.26 ± 0.1
H273A	0.76 ± 0.05	0.05 ± 0.001	$0.6 \pm 2.0 \times 10^{-2}$
D110A	No activity detected		
D111A	No activity detected		
D112A	No activity detected		

To ensure that these mutations did not alter the secondary structure of the enzyme, CD experiments were performed on the wild-type and all of the OsUAM1 mutants created. The results tabulated in Table 4-2, show that all the mutants have comparable % secondary structural elements with the wild-type enzyme. These results suggest that mutating the DDD-motif does not drastically change the secondary structural elements of OsUAM1, but the activity is quenched since the enzyme is probably unable to bind the Mn^{2+} cofactor.

Table 4-2: Secondary structural elements of wild-type *Os*UAM1 and its mutants, as determined using CD.

Enzyme	α- helix (%)	Parallel β- strands (%)	Antiparallel β-strands (%)	β-turns (%)	Random coils (%)
*Os*UAM1(wild-type)	17	13	16	20	48
H273A	19	14	14	20	45
D110A	18	15	14	21	46
D111A	19	14	13	20	45
D112A	16	16	15	21	48

4.7 Discussion

The metal binding studies performed by Konishi *et al.* (2007) were limited, as their studies gave an incomplete picture of the metal ion dependency of *Os*UAM. They reported using 5 mM final concentration of the metal to perform the reaction. Initial experiments demonstrated that most of the enzyme was precipitated (formation of white precipitate in the protein containing tubes) with 5 mM of Mn^{2+}. Further, the group also reported that the enzyme was active without the metal, as they discovered activity when no metal was added to the reaction mixture.[18] Hence, Mn^{2+} was described as an activator of the enzyme. This study claims that Mn^{2+} is necessary for the activity of the enzyme and does not behave as an activator.

In this thesis, metal binding studies were performed on *Os*UAM1 and all three *At*RGPs. All four enzymes, showed maximum activity at lower concentrations, at either 40 µM or 80 µM of Mn^{2+}, but their activity decreased when the metal ion concentration was increased higher than 320 µM. This could be due to the excess metal binding at different sites of the enzyme, interfering with bonds between residues and bringing about detrimental changes to the overall structure of the

enzyme, thus making it less active. Further, the excess metal could also bind to the charged substrate phosphates thereby making them inaccessible to the enzyme.

The EDTA-treated enzymes were inactive when assayed without metal; however, activity was recorded even with metal ion concentration as minimal as 2 µM. This meant that the metal ion was necessary for activity rather than it being an activator of the enzyme. Konishi *et al*. (2007) observed only a fractional increase in activity (under 5%) with Zn^{2+} and Mg^{2+}, compared to the activity they obtained with no metal while the others divalent ions did not show appreciable activity. Contrary to these observations, our experiments revealed that other divalent metal, such as Zn^{2+} and Co^{2+}, retained at least 60 % of the activity of Mn^{2+} at the concentration that gave maximum activity in all four enzymes. It is also pertinent to note that there was no contaminant manganese ions in any of the other divalent metal ion solutions used for the experiment, based on the manufacturer's claim of purity. The high activity of UAM with Mn^{2+} could be accounted to the size of Mn^{2+}, which may be just right to keep UAM's active site in the right configuration to help it coordinate the substrate.

Despite the importance of the metal cofactor, the information known about the role of the metal and its contributions towards the function of the enzyme is minimal, since there is no structural information available for this enzyme. Attempts were made to crystallize *Os*UAM1 and *At*RGP1, but they were unsuccessful in producing crystals of good quality. Purified *Os*UAM produced a single band at ~ 41 kDa on an SDS-gel, but the molecular weight estimated by size-exclusion chromatography was ~ 460 kDa.[18] *Os*UAM was likely to exist as a complex made up of numerous monomeric units. The oligomeric state of the enzyme could be one of the reasons why obtaining diffraction quality single crystals has proven difficult thus far; as larger macromolecules generally tend to have more flexible fragments.

Nonetheless, an initial sequence-based model was proposed for *Os*UAM1, which helped determine the regions that could form secondary structures and also predicted the possible active site residues. The structure models generated from the two different programs, I-TASSER and GalaxyWEB, had similar overall architecture and minimal dissimilar regions. I-TASSER predicted Chondroitin polymerase as a functional analog of the model of *Os*UAM1; this enzyme uses Mn^{2+} metal ion cofactor and UDP-based substrates (UDP-GlcUA and UDP-GalNAc). By docking UDP-GalNAc in the *Os*UAM1 model, I-TASSER proposed three different loops from the starting model that may bind the uridine, diphosphate and sugar regions of the substrate. The model puts the DDD motif (the region previously hypothesized to bind the metal cofactor) and His273, in a position to bind the metal, which in turn is able to stabilize the substrate.

To test the role of these residues, alanine mutants of all four residues were prepared and tested for activity. All three alanine mutants of the DDD-motif inactivated the enzyme. Histidine was not part of the DDD-motif; however, a number of manganese-dependent enzymes have been shown to involve their active site histidines for metal coordination. The H273A mutant reduced enzyme activity but failed to inactivate the enzyme completely. This could mean that His273 is perhaps not directly involved in metal coordination in *Os*UAM1. Our CD experiments also showed that none of the mutants changed the secondary structure of the enzyme, and therefore the loss of activity is due to the inability of OsUAM1 to bind the metal cofactor.

Based on the *Os*UAM1 model predicted, one of the features of the enzymes that belong to the glycosyltransferase family, is the stabilization of the substrate diphosphate by the metal and the binding of the metal to the DXD-motif. Like the other proteins in this protein family, in UAM, the metal's role could be limited to the phosphate stabilization and may not be involved directly in catalysis. In addition, to the work discussed in this chapter, extended X-ray absorption fine

structure (EXAFS) was also performed on *Os*UAM1. The data collection and processing of EXAFS data was performed by our collaborators, Dr. Julien Cotelesage, Research Associate with Dr. Graham George, Dept. of Geology, University of Saskatchewan. Although the data obtained is inconclusive, initial results agree with the above-mentioned role of metal in this enzyme i.e. stabilization of substrate diphosphate. Some of the EXAFS results have been included in the Supplementary Section S.3.

Chapter 5: GDP-6d-*altro*-Heptopyrnaose Mutase

GaHM is the first example of a pyranose-furanose mutase enzyme that interconverts a heptose sugar. The enzyme has low sequence identity (< 21%) to both prokaryotic and eukaryotic UGMs. The sequence alignment shown in Figure 5-1, suggests that GaHM could be a flavoenzyme, due to the presence of conserved residues that are known to covalently bind FAD. In fact, the purified recombinant *Cj*GaHM enzyme was yellow in colour; the protein fractions had the characteristic UV-absorbance at 450 nm indicating the presence of FAD.[58]

GaHM catalyzes the interconversion between GDP-*altro*-Hep*f* and GDP-6d-D-altro-heptopyranose (GDP-*altro*-hep*p)*. According to the studies performed by Dr. Todd Lowary's group at University of Alberta, when GDP-6d-D-*altro*-Hep*f* was incubated with reduced GaHM, the product formed, GDP-6d-D-*altro*-Hep*p* was in a 1:1 ratio and the equilibrium did not favor the formation of the pyranose form as seen with other previously studied mutases.[58]

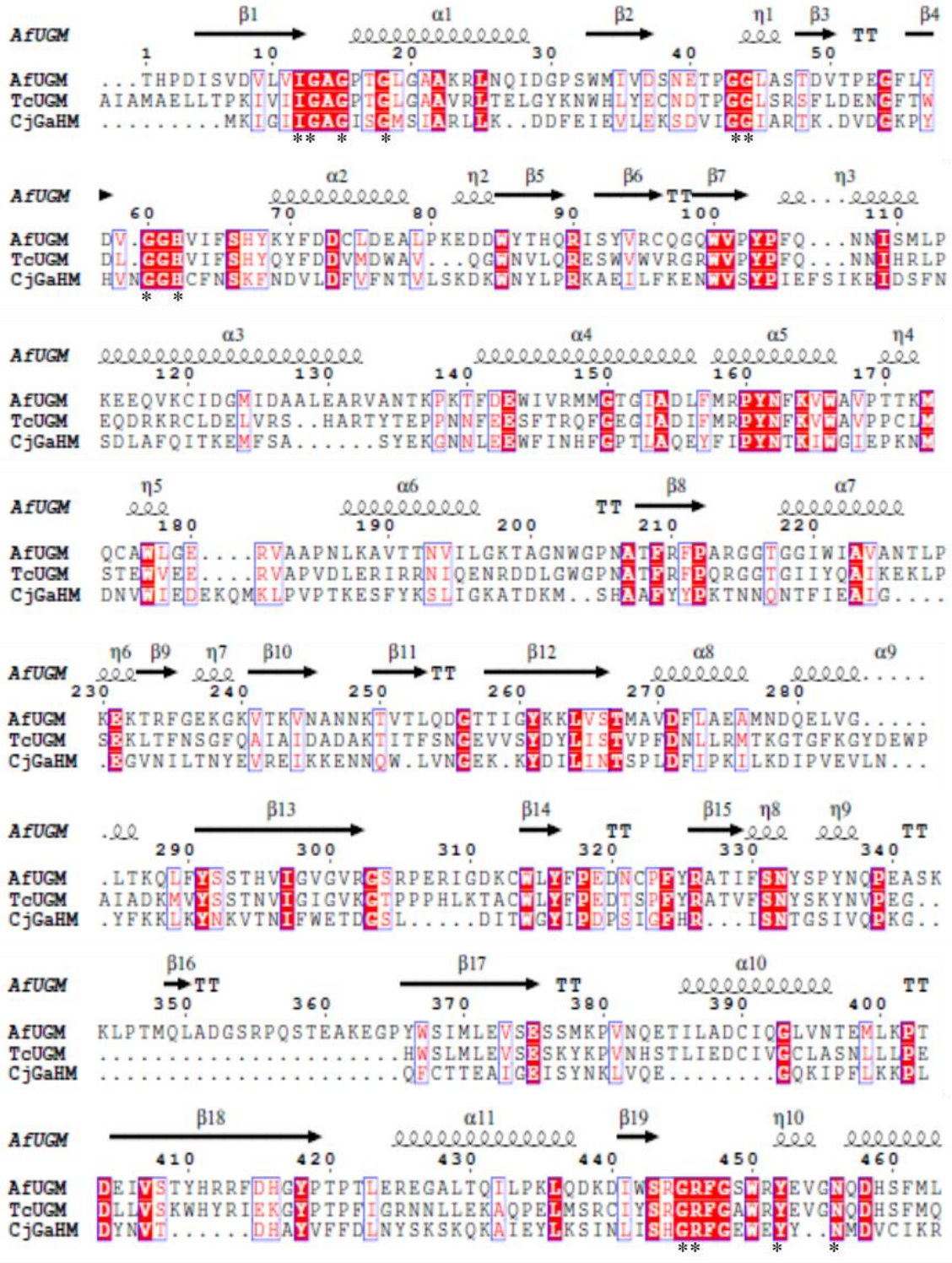


Figure 5-1: Sequence alignment of *Cj*GaHM with *Af*UGM and *Tc*UGM.
Conserved residues are shown in red boxes. The residues that interact with FAD are highlighted with an *. The sequence alignment was performed using ESPript (Version 3.0)

5.1 Structural studies on *Cj*GaHM

A low sequence identity meant that homology models of *Cj*GaHM could not be created, and hence, X-ray crystallography was used to examine the structure. Recombinant *Cj*GaHM enzyme required for crystallization experiments were provided by Dr. Todd Lowary's group. The crystal structure of GaHM with GDP was solved by Dr. Sean Darlymple, a former post-doctoral fellow in the Sanders lab. The crystal structure of *Cj*GaHM, with FAD, both in oxidized and reduced states, was solved initially with GDP in the active site. Since no previous models were available for this enzyme and due to its low sequence identity with other UGMs, molecular replacement could not be used to solve the structure. Crystals were obtained with Se-Met GaHM protein, Figure 5-2 and the structure was solved with single anomalous dispersion (SAD) technique using Se-Met for phasing.

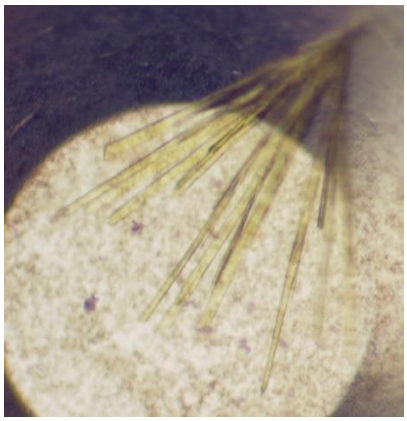


Figure 5-2: Crystals of *Cj*GaHM with GDP.
These co-crystals were obtained by Dr. Sean Darlymple.

5.1.1 Structural features of *Cj*GaHM

The structures of *Cj*GaHM: GDP with FAD_{ox} and FAD_{red} were solved to 2.7 Å and 2.4 Å, respectively. GaHM monomer consists of three domains. Domain 1 is the αßα Rossmann fold which binds the FAD cofactor, Domain 2 is the α-helical domain containing five α-helices, and

Domain 3 is the ß-strand domain, consisting of antiparallel ß-strands, which plays a role in interconnecting the two domains, as shown in Figure 5-3. As seen with the structure of UGMs, the active site of *Cj*GaHM is located in the cleft below FAD. The overall structural features of GaHM compares well with the prokaryotic and eukaryotic UGMs.

A structural comparison of GaHM with *Dr*UGM (prokaryotic) and *Af*UGM (eukaryotic) UGM crystal structures reveals more details about the enzyme. The structure of *Cj*GaHM is comparable to *Dr*UGM (surface area ~ 45600 A°^ 2) but more compact than that of *Af*UGM (surface area ~ 59000 A°^ 2). This is understandable given the length of the sequences of these enzymes; the GaHM sequence is longer than that of *Dr*UGM but not as long as that of *Af*UGM. The GaHM crystal structure however is more similar to eukaryotic UGMs. *Dr*UGM, mentioned in detail in Chapter 3, has only a single flexible loop that moves into the active site when the substrate is present. Eukaryotic *Af*UGM has two mobile loops that move into the active site if the substrate is present in the active site, as shown in Figure 5-3. The structure of *Cj*GaHM also has two mobile loops; however, they could not be modeled into the structure due to little or no electron density.

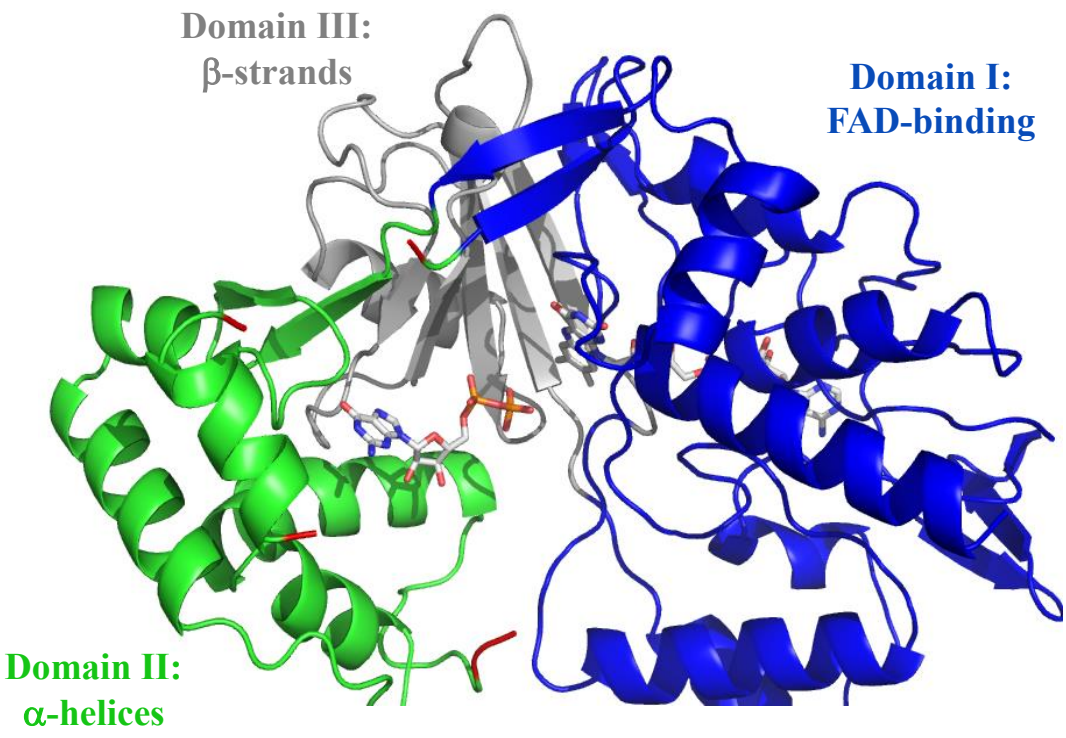
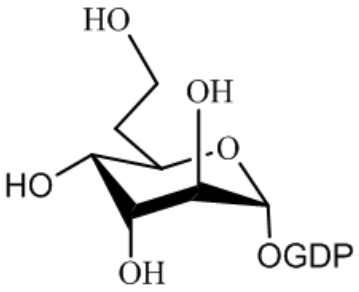
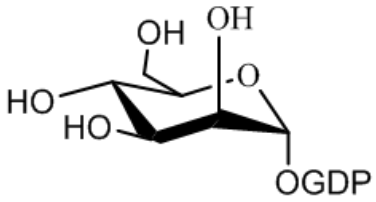


Figure 5-3: The crystal structure of *Cj*GaHM: GDP.
The three domains, Domain 1 in blue, Domain 2 in green and Domain 3 in grey are shown. The start and end of mobile loops 1 and 2 are shown in red.

The objectives of this project were, to understand the structural features of *Cj*GaHM that were responsible for substrate recognition and specificity to a GDP-based substrate, instead of the usually preferred UDP-based substrate, observed with all UGMs studied thus far. In the *Cj*GaHM structure with GDP, there is no electron density available to model the loops. This may be due to the absence of the sugar moiety, which helps in additional stabilization of the flexible loop. Crystallizing *Cj*GaHM with a GDP-sugar will yield information about the binding of the sugar in the active site and the residues that contribute to its stabilization. The substrate GDP-6d-D-*altro*-Hep*p* could not be obtained either commercially or by synthesis. Hence, the crystallization of *Cj*GaHM was attempted with GDP-mannose, chosen based on its similarities with GDP-6d-D-*altro*-Hep*p*, as shown in Figure 5-4.



GDP-6d-D-*altro*-Hep*p*



GDP-mannose

Figure 5-4: Difference between GDP-mannose and GDP-6d-D-altro-Hep*p* structures.
GDP-mannose differs from GDP-6d-D-*altro*-Hep*p* at C3 and also GDP-6d-D-altro-Hep*p* has seven carbons.

5.2 Crystallization of GaHM with GDP-mannose

GaHM was co-crystallized with GDP-mannose in the condition containing, 0.1 M ammonium sulphate, 0.1 M Bis-tris pH 6.5 and 25% PEG 3350 (Figure 5-5). The diffraction data obtained for GaHM: GDP-mannose complex crystals in both FAD_{ox} and FAD_{red} forms were solved to 2.1 Å and 2.3 Å resolution, respectively.

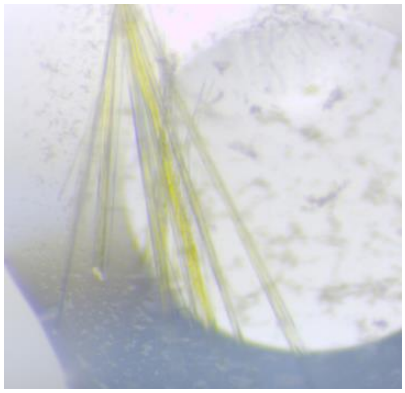


Figure 5-5: Crystals of GaHM: GDP-Mannose.
These crystals were obtained in 0.1M ammonium sulphate, 0.1M Bis-tris pH 6.5 and 25% PEG 3350.

5.3 GDP-6d-*altro*-Heptopyrnaose Mutase structures

5.3.1 $GaHM_{ox}$: GDP-mannose

In the crystal structure, GaHM exists as a monomer. There are two monomer units in the ASU. The active site of GaHM has residues to bind and interact with the guanosine base of GDP. Residues such as Phe147, Pro149, and Tyr146 stack the guanine while Glu98 and Tyr146 make strong hydrogen bonding contacts, as shown in Figure 5-6. Glu98 hydrogen bonds to N1 (2.8 Å) and the amino group at C2 (2.9 Å) of guanosine. Tyr146 hydrogen bonds (2.7 Å) with N7 of guanosine. The guanosine ribose is held in position by a hydrogen bond interaction (2.9 Å) between its C2 hydroxyl group and the amino group of Asn151.

In the sugar binding region, mannose was modeled into the electron density that was observed below the FAD. The mannose C1 anomeric carbon is at a distance of ~ 4 Å from the N5 of FAD. Mannose forms hydrogen bonding interactions with Arg386, Asn394 and His54 residues in the active site. The guanidino nitrogens of Arg386 form hydrogen bonds with the C2 and C3 hydroxyls of mannose at 3.4 Å and 3.3 Å respectively. Hydrogen bonds are also formed between Asn394 and the C4 hydroxyl of the sugar and also between His54 and the C6 hydroxyl of mannose. In the phosphate binding region, Arg307 is in a position to form contacts with the β-phosphate of the substrate.

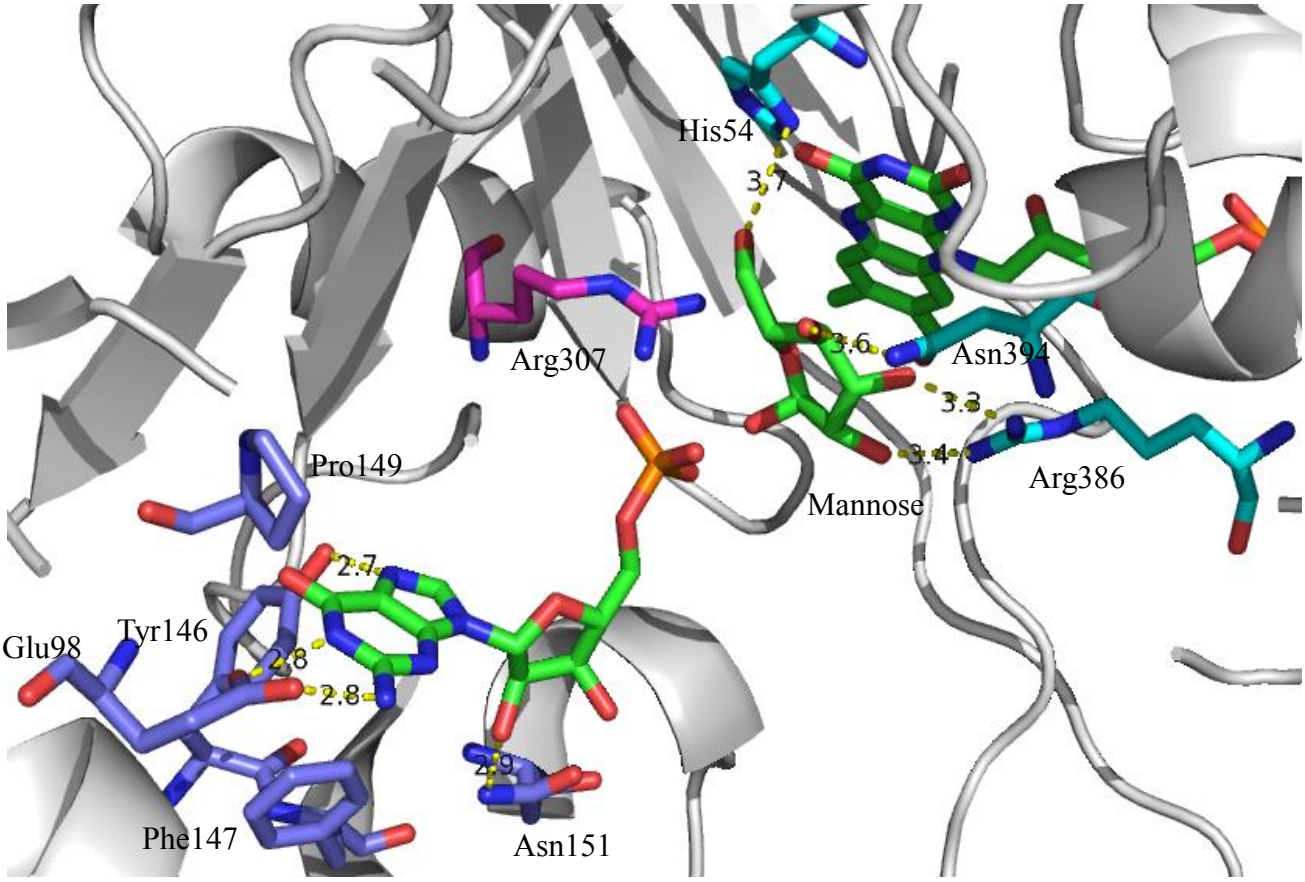


Figure 5-6: A close-up of the active site of GaHM.
The residues that bind guanosine (in blue) and ribose of GMP (in green) are shown. The residues that make contacts with mannose sugar (in green) are shown in cyan. Arg307, which is in position to stabilize the β-phosphate of GDP-mannose is also shown, in magenta.

In the $GaHM_{ox}$: GDP-mannose structure, there is no clear electron density for the two flexible loops (Loop 1 residues 164-176 and Loop 2 residues 191-197). Density is observed for Loop 2 in chain B, but the loop is in an open conformation, stabilized by interactions with the adjacent monomer unit. Moreover, there is no density for the mannose sugar in this chain. In chain A, both loops are not visible; they are not in a closed conformation, and the substrate mimic is exposed to the solvent, as it is not buried completely within the active site. Despite this, density for mannose is observed in this chain.

The problem here is two-fold. Firstly, loop flexibility leads to a loss of information in regards to the residues that may bind and stabilize the phosphate. Arg307 is in a position to form contacts with the β-phosphate of substrate. This is also observed in the GaHM: GDP complex structures solved. However, information regarding the residues which interact with the α-phosphate and locks down the substrate in the active site is still unavailable. Secondly, the flexibility of the loops leads to discontinuous density for the substrate (Figure 5-7A). The guanosine monophosphate (GMP) and mannose portions had to be modeled in separately, and there is little or no density observed for the β-phosphate. The sugar-binding also becomes flexible leading to the observation of two different conformations for the sugar each at on occupancy of 0.5 (Figure 5-7B). Good electron density for one of these loops (loop 190-197) was observed in Chain B of the structure. This was because the loop was stabilized in an open conformation, due to crystal contacts, although this caused the absence of electron density for the mannose sugar in the active site of Chain B.

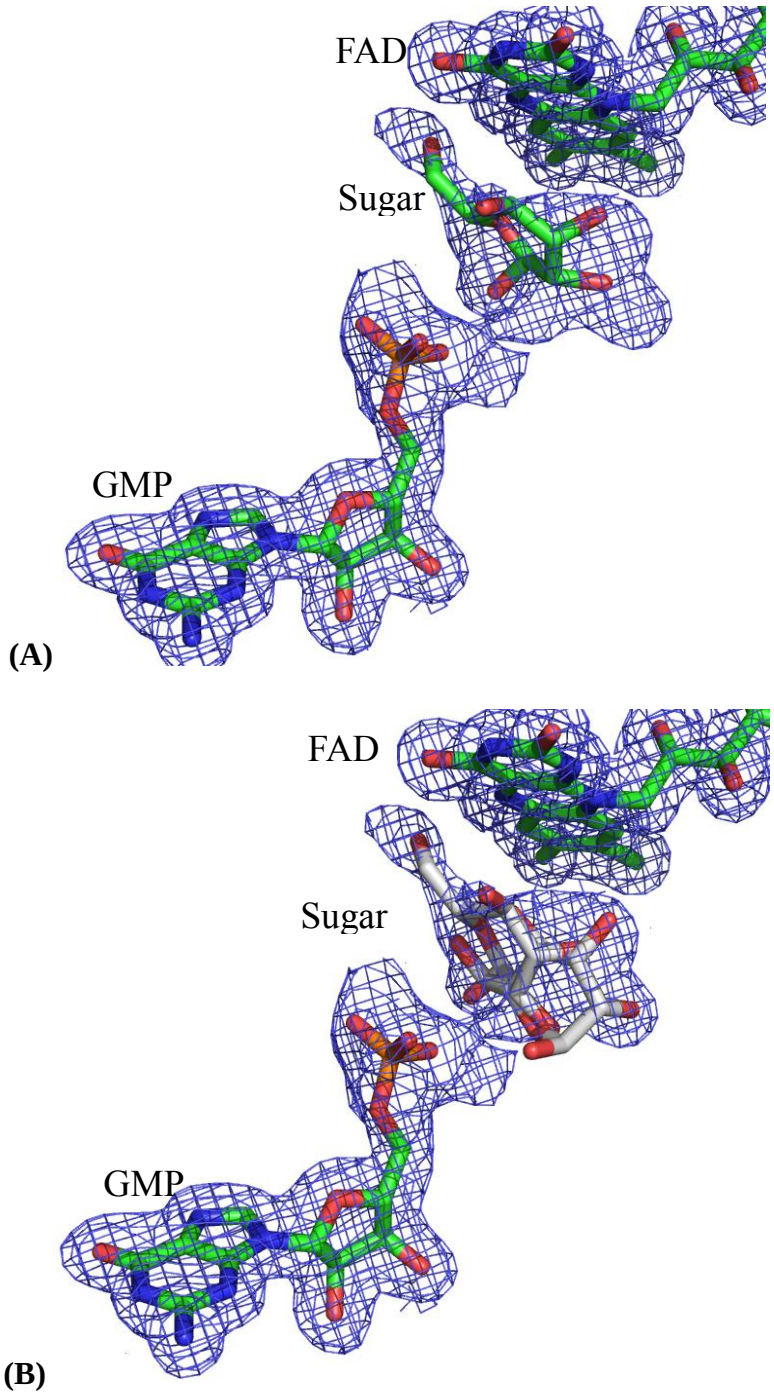
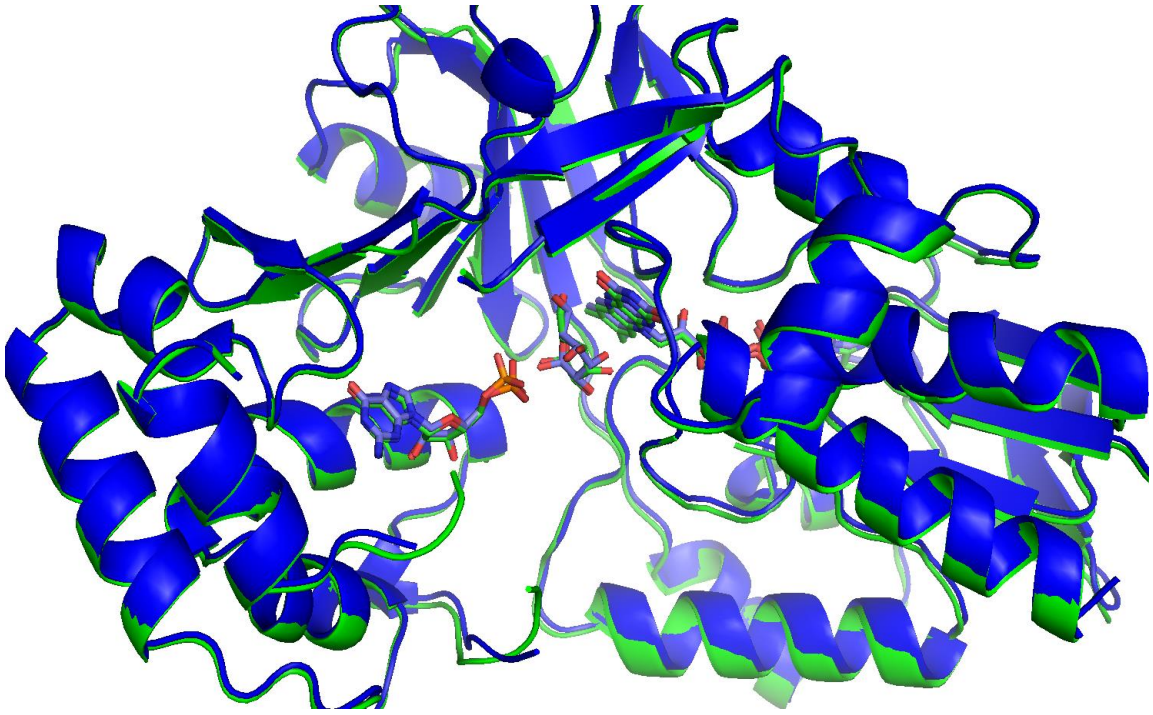


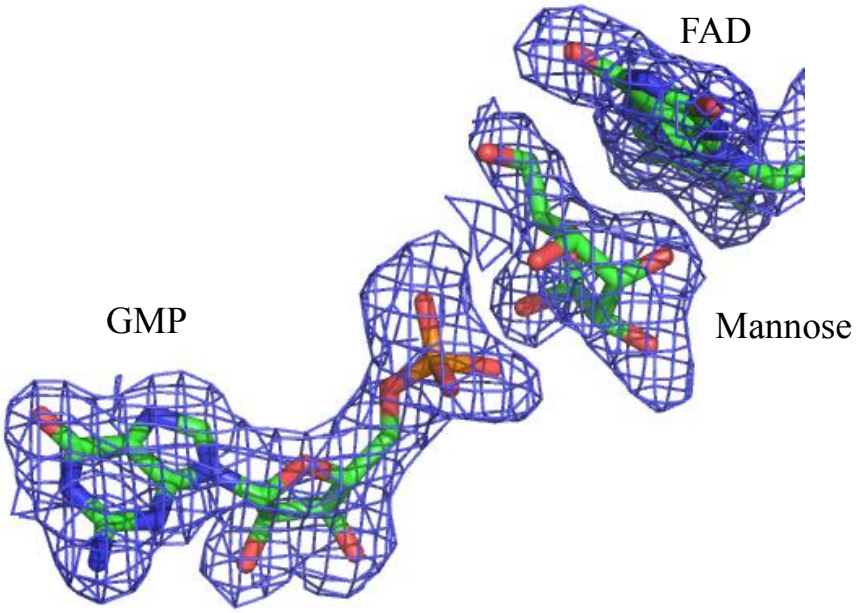
Figure 5-7: Electron density maps for GMP and Mannose in GaHM$_{ox}$: GDP-mannose.
(A) GMP and mannose sugar modeled in the electron density (1.5 σ) separately due to discontinuous density, β-phosphate has no density. (B) The big blob of density surrounding mannose in (A) can accommodate two mannose molecules (shown in white, having two conformations), illustrating the flexible nature of the mannose binding, due to it being a substrate mimic.

5.3.2 $GaHM_{red}$: GDP-mannose

The $GaHM_{red}$: GDP-mannose structure solved to 2.3 Å, shows an overall similar structure to the $GaHM_{ox}$: GDP-mannose complex structure as shown in Figure 5-8A. As with the $GaHM_{ox}$ structure, the electron density for both the loops was missing in monomer A; monomer B had the density for Loop 2, but it was in the open conformation. Electron density for GDP-mannose was discontinuous in the $GaHM_{red}$ structure, and GMP and mannose were modeled in monomer B (Figure 5-8B). Monomer A showed density for GDP but no density under FAD for the sugar.



(A)



(B)

Figure 5-8: Crystal structure of the overall structures of $GaHM_{red}$: GDP-mannose.
(A) Overall structures of $GaHM_{red}$: GDP-mannose (in blue) and $GaHM_{ox}$: GDP-mannose (in green)
(B) The electron density (1.5 σ) and the modeled mannose under the FAD and GMP are shown.

Table 5-1: Data collection and refinement statistics.

	$GaHM_{ox}$: GDP-mannose	**$GaHM_{red}$: GDP-mannose**
Data collection		
Space group	$P2_1\ 2_1\ 2_1$	$P2_1\ 2_1\ 2_1$
Cell dimensions	47.62, 72.50, 272.89	47.48, 71.65, 272.64
a, *b*, *c* (Å)	90, 90, 90	90, 90, 90
Resolution range (Å)	46.9-1.57 (2.10)	46.7-1.87 (2.30)
All reflections	771658	490592
Unique reflections	129116	77769
Redundancy		
$I / \sigma I$	8.88	7.52
Completeness (%)	95.7 (100)	99.3 (100)
No. molecules in ASU	2	2
Data refinement		
Resolution range (Å)	46.9-2.10	47-2.30
R_{work} / R_{free}	0.1564/0.2049	0.1780/0.2287
No. amino acid residues	3×418	3×418
No. of Water (molecules)	606	406
Ligand		
FAD	2×FAD	2×FAD
GMP	2×GMP	2×GMP
Mannose	1× M1P	1× M1P
r.m.s deviations		
Bond lengths (Å)	0.008	0.009
Bond angles (°)	1.096	1.234
Ramachandran favored	95.7 %	95.9%

5.4 Loop modeling

The electron density for Loop 1 (from Asp163 to Pro178), is not clear or is absent in both $GaHM_{ox}$ and $GaHM_{red}$ structures. Based on the observations made from the sequence alignment with eukaryotic UGMs, (*Af*UGM and *Tc*UGM), the Loop 1 of *Cj*GaHM is longer by three residues, as shown in Figure 5-9. The sequence of Loop 1 in *Af*UGM and *Tc*UGM have a methionine (Met173 and Met171, respectively) and tryptophan (Trp177 and Trp175, respectively) residue that is also conserved in GaHM. In all the three sequences, the other end of Loop 1 has proline residues that act as starting points to their respective loops. In both *Af*UGM and *Tc*UGM, a positively charged arginine residue (Arg182 and Arg176, respectively) stabilizes the α-phosphate and also forms a key hydrogen-bond interaction with asparagine (Asn457 and Asn433, respectively) once the substrate is in the active site. In GaHM, the role of arginine is expected to be played by two residues, based on the sequence alignment. Lys174, a positively charged residue is expected to interact and stabilize the negative charge on α-phosphate.

In the case of Loop 2, the electron density is absent from Lys190 to Ala199 in monomer B. Loop 2 of GaHM is shorter by a residue, compared to *Af*UGM and *Tc*UGM sequences. Loop 2 has five conserved residues in both *Af*UGM and *Tc*UGM sequences, only one of them; alanine (Ala207, Ala205, and Ala199 respectively) is conserved in all three sequences.

The crystal structures of both *Af*UGM and *Tc*UGM were used as templates to model the missing GaHM loop regions, due to the similarities in sequences and the crystal structures themselves.[34,36] GaHM residues Met162 through Trp166 were modeled, corresponding to the position of the above-mentioned conserved methionine and tryptophan residues, in Loop 1 of *Af*UGM and *Tc*UGM. The other end of the loop, from Pro178 to Lys174, was also modeled based on the similarities with residues from the eukaryotic UGM sequences, in such a way that Lys174

is in a similar position to that of the Arginines in *Af*UGM and *Tc*UGM.[34,36] The eukaryotic UGMs have a small helix (α-15) located in front of the cleft and isoalloxazine ring of FAD, as shown in Figure 5-9. The helix functions to ensure that the substrate is completely buried in the active site of the eukaryotic UGMs. This small helix is completely absent in GaHM. Therefore the remainder of Loop 1, between residues Ile167 to Met173, was modeled in a position similar to that occupied by this helix.

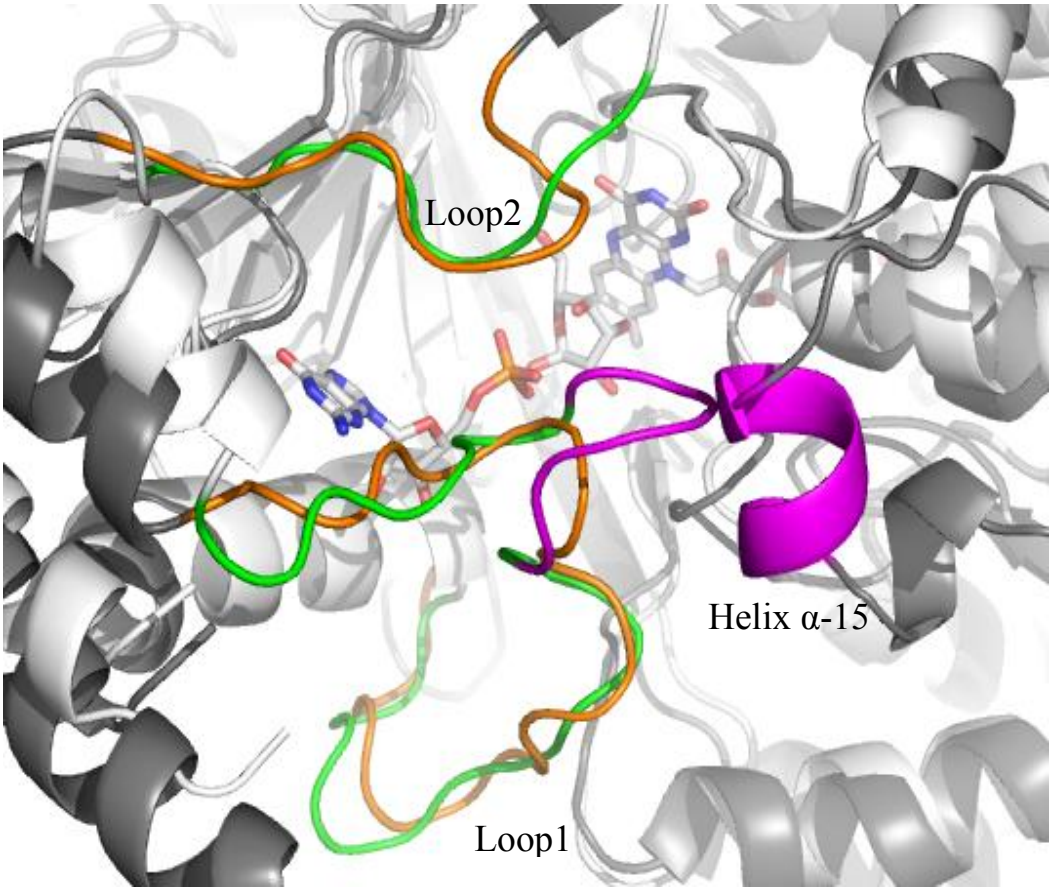


Figure 5-9: Overlay of crystal structures of *Af*UGM and *Cj*GaHM.
Loop1 and Loop2 of *Cj*GaHM (shown in green), modeled on Loop1 and Loop2 of *Af*UGM (shown in orange). The region of *Cj*GaHM Loop1 (residues Ile167 to Met173) modeled in the region occupied by the short helix of *Af*UGM (magenta) is highlighted in magenta.

The role of Loop 2 is to help bury the substrate completely in the active site of the enzyme by moving in towards the active site in the presence of substrate, to contribute residues such as Asn207 in *Af*UGM and Asn200 in *Tc*UGM, which can interact with the hydroxyls of the sugar.[34,36] Despite the dissimilarity of the sequences, Loop 2 of GaHM was also modeled similar to the loops from *Af*UGM and *Tc*UGM.

5.5 Discussion

5.5.1 Specificity for GDP-based substrate

To understand the specificity of *Cj*GaHM to a GDP-based substrate, the residues around the GDP-binding site were compared to those from eukaryotic UGMs, since the overall crystal structures were similar. The sequence alignment showed that the residues corresponding to these regions are similar, with a few exceptions. All three sequences have conserved tryptophan, valine, tyrosine, proline and isoleucine residues. However, it is the substitution of Phe105 in both eukaryotic UGMs with a less bulky non-aromatic Ile97 in GaHM and a further substitution of the basic glutamine residue (Gln106) in eukaryotic UGMs by an acidic glutamate (Glu98) in GaHM, which helps GaHM recognize and bind a GDP-based substrate. Glu106 is able to make two hydrogen bonds with the uracil ring while Gln98 has the ability to make two hydrogen bonds with guanosine, as shown in Figure 5-10. Thus, these subtle changes in the active site aid the binding of a GDP-based substrate by GaHM.

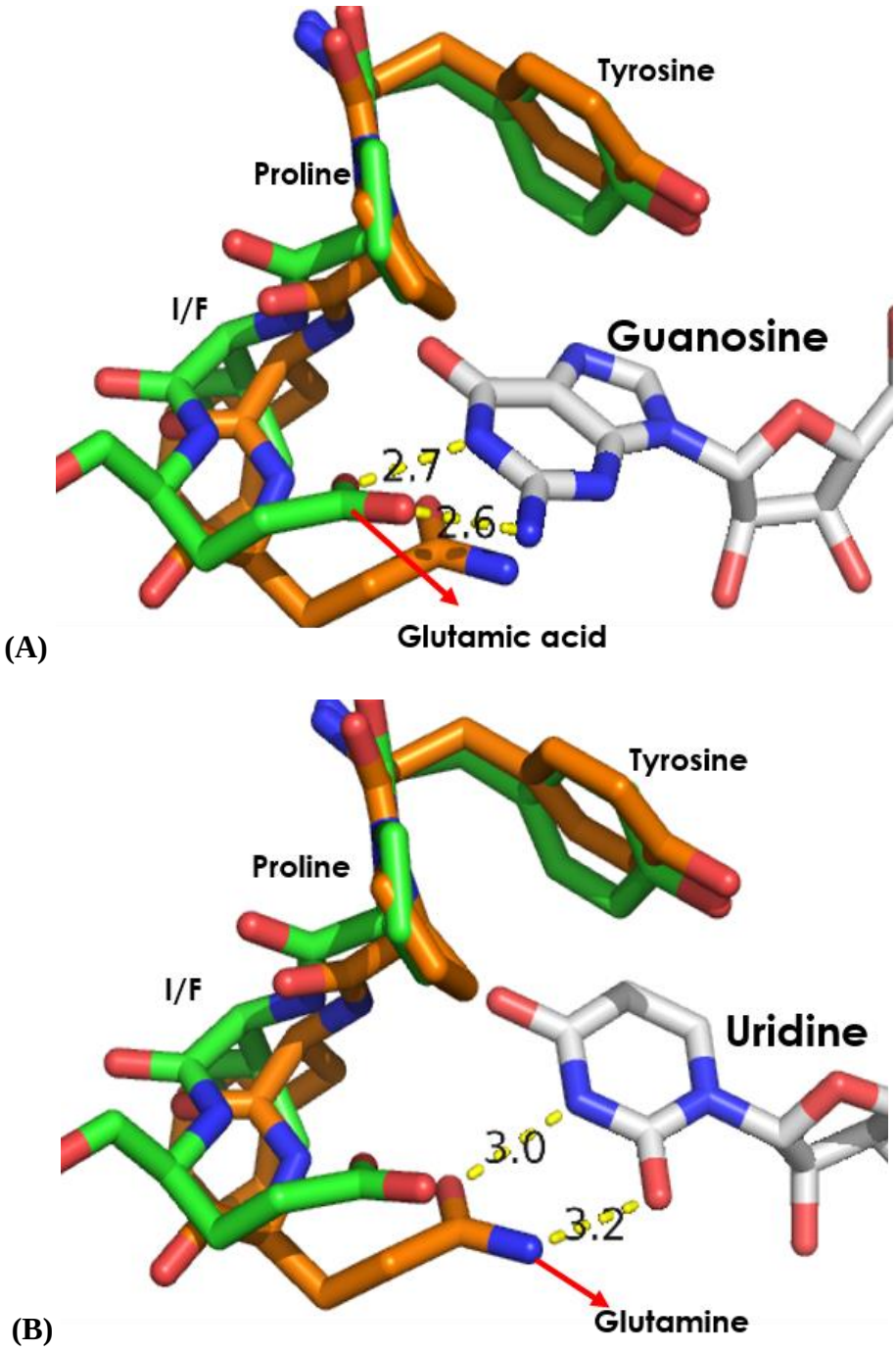


Figure 5-10: Overlay of the nucleotide-binding region of *Af*UGM and *Cj*GaHM.
(A) Glutamic acid is in position to interact with the guanosine in *Cj*GaHM (green) (B) Glutamine interacts with uridine in *Af*UGM (orange).

5.5.2 Stabilization of the α-phosphate

In the crystal structures of *Cj*GaHM, solved with GDP-mannose, the flexible mobile loops that close to bury the substrate in the active site could not be built-in, due to little or no electron density. The overall structural similarity of *Cj*GaHM, with structure of eukaryotic UGMs such as *Af*UGM and *Tc*UGM, enabled the modeling of the flexible loops of *Cj*GaHM,

To identify other residues of the *Cj*GaHM active site that interact with the substrate, the residues contributed by the loop regions of the enzyme were compared to eukaryotic UGMs. The mobile loop brings in an arginine residue (Arg180 in *Af*UGM and Arg176 in *Tc*UGM) which stabilizes the α-phosphate. This residue also interacts with an Asn (Asn457 in *Af*UGM and Asn453 in *Tc*UGM), to keep the enzyme in a closed conformation. In fact, this is also seen in prokaryotic UGMs, as discussed in Chapter 3. Based on the modeled Loop 1, the role of arginine is performed by two residues in *Cj*GaHM. Lys174 moves in with the loop to stabilize the α-phosphate. However since lysine lacks the guanidino group of arginine, it cannot form hydrogen bond interactions with Asn394. The sequence alignment suggests that one of the residues that follows, such as Gln172 should be able to play such a role. The loop may thus be kept in a closed conformation by Gln172, which forms hydrogen bonds with Asn394 of *Cj*GaHM as shown in Figure 5-11.

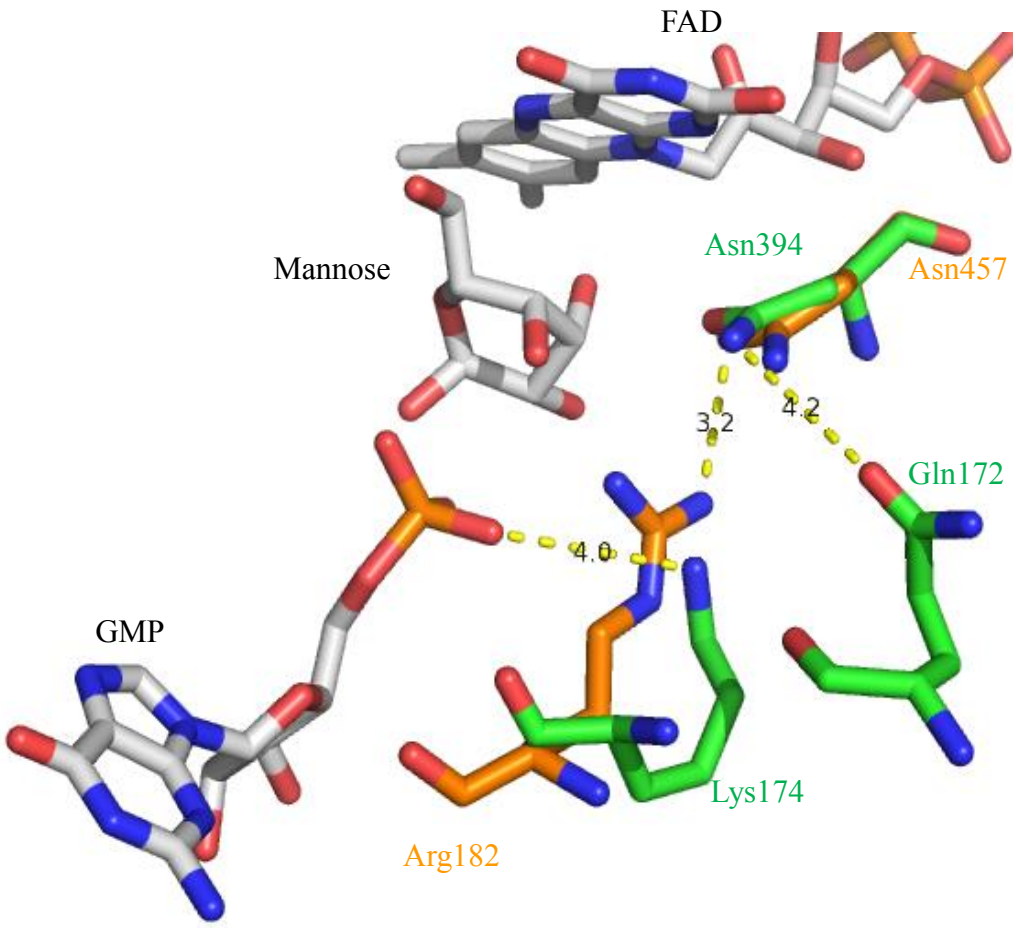


Figure 5-11: Active site residues of *Cj*GaHM predicted to keep Loop 1 in a closed configuration.
The *Cj*GaHM are shown in green. Lys174 stabilizes the α-phosphate while Gln174 forms hydrogen bonds with Asn394 to keep the loop closed and bury the substrate. Arg182 which interacts with Asn457 in *Af*UGM to maintain the loop in a closed configuration is also shown in orange.

Chapter 6: Conclusions and future work

Furanoses are vital components in the cell wall of bacteria, fungi, plants and protozoa. It is important to study the furanose containing glycoconjugates since they are absent in humans and mammals. Their identification in disease-causing pathogens means the enzymes aiding in the production of furanoses are potential drug targets. In this thesis, three different pyranose-furanose mutases (UGM, UAM, and GaHM), employed by their respective pathways leading to the deposition of three different furanoses in the cell wall of various organisms, were studied.

From literature, the understanding of these mutases is to varying degrees. In this thesis, the objectives were outlined so as to focus on the substrate-binding aspects of these mutases. Two of these enzymes, UGM and GaHM, require the same cofactor (FAD) for their respective interconversion reaction, but bind UDP and GDP based substrates, respectively. Also GaHM, unlike any of the mutases described so far, catalyzes the interconversion of a heptose-sugar. Plant-enzyme UAM is a metalloenzyme that does not require FAD for activity but works on a UDP-based substrate. Therefore, the requirement for different substrates and different cofactors meant that the modes of substrate binding were different among the three enzymes.

6.1 UGM

*Dr*UGM was chosen as a model system to study how substrate binding is affected by the active site residues of the enzyme. The importance of only a few residues in binding the substrate in the most productive conformation is known in UGM. Changes were introduced to the active site in the form of point mutations, and experiments were designed to understand the role of these residues. Kinetic assays, modeling, and docking studies were performed on *Dr*UGM mutants. The studies have established the roles of Trp184, Arg364, His88 and Asn372 and residues performing

similar roles in the active sites of other UGMs. Trp184, an important residue required to position the uridine-portion of the substrate, not only makes contacts with the uridine-ribose but also helps stack it in the active site, thereby increasing the efficiency of the enzyme. Arg364 and Asn372 keep critical residues like Arg198 in the correct position. His88 positions the sugar (Gal*p*) under FAD for catalysis. All of these residues contribute to enabling the substrate to bind in the most productive mode in the active site. Minor changes in the active site conformation affected the substrate binding mode of this enzyme and which tends to place the sugar (Gal*p*) further away from the N5 of reduced FAD, thereby causing depletion of activity.

*Mt*UGM is a validated drug target. The inhibitor MS-208 binds to an allosteric site on the enzyme and causes changes to the substrate binding mode, altering the positioning of Arg182 (residue contributed by the mobile loop). MS-208 was tested against *Mt*UGM and the data obtained was globally fit to the equation for mixed-type inhibition. The K_i and K_i' of the inhibitor calculated from the fit were ~135 µM and ~400 µM respectively. Though evidence for the exact conformation of the binding of MS-208 to the allosteric site is yet to be found, the studies have shown the importance of having the active site residues of UGM in the correct conformation so that the substrate binds in the productive mode.

6.2 UAM

Chapter 4 of the thesis discussed the results obtained from the plant metalloenzyme UAM. Konishi *et al.,* (2007), who identified this enzyme in *O. sativa*, described Mn^{2+} as an activator of UAM at 5 mM concentration. Our results from HPLC assays showed that Mn^{2+} is essential for UAM activity, and 5 mM concentration of Mn^{2+} had an inhibitory effect on UAM activity. This was also observed with all three *At*RGPs. All four enzymes showed maximum activity at 40-80

μM Mn^{2+} range. Among the other divalent metal ions tested, Zn^{2+} and Co^{2+} showed 60-70 % activity at this concentration.

More studies were focused around the metal cofactor and its interaction with the substrate. A starting model for UAM was developed based on the *Os*UAM1 sequence, using I-TASSER and GalaxyWEB. On comparison with a functional analog, I-TASSER predicted few possible active site residues that can interact with different regions of the substrate. Asp110, Asp111, Asp112 (DDD-motif) and His273 are in position to bind the metal in *Os*UAM1. The prediction was tested by mutation of the residues mentioned above to alanines. None of the mutations influenced changes in the secondary structure, yet the DDD-motif mutants lost activity, indicating the inability of the enzyme to bind its metal cofactor. SCOP classification of the *Os*UAM1 model also suggests that role of the metal cofactor is to coordinate and stabilize the substrate diphosphate. EXAFS studies were performed on *Os*UAM1 samples and a preliminary model for the metal binding region in *Os*UAM1 was proposed, by our collaborators. The EXAFS data is in agreement with the observations made from the I-TASSER model.

6.3 GaHM

The thesis also discusses results from work on a third pyranose-furanose mutase, *Cj*GaHM. Our crystal structures show that the active site, though similar to UGMs, can adapt to bind and interact with the guanosine of GDP-based substrates. The GaHM domains that bind FAD and substrate are overall similar to eukaryotic UGMs. Crystal structures of GaHM (FAD_{ox} and FAD_{red}) were obtained with the substrate mimic, GDP-mannose. The structures show no electron density for two large flexible loops. This is due to the loops being unable to form enough contacts with GDP-mannose to exist in a closed conformation so as to bury the substrate completely in the active

site. As a result, the sugar and ß-phosphate of GDP-mannose remain flexible in the active site and are unable to form stable interactions with active site residues. This limits the information on active site residues that can stabilize the phosphate and bind the sugar of substrate mimic, GDP-mannose.

The loops were modeled into the GaHM crystal structure based on the eukaryotic UGMs, and the GaHM residues that may interact with the substrate diphosphate and the sugar were identified. The role of the arginine residue that stabilizes the α-phosphate and keeps the loop in a closed conformation, in all the known UGM structures thus far, may be accomplished by two residues Lys174 and Gln172 in *Cj*GaHM. However this is only a predicted model for the closed conformation of the loops and only a crystal structure of the enzyme with its substrate, GDP-6d-*altro*Hep*p*, can provide further evidence.

6.4 Future directions

All three mutases perform similar reactions and yet they are each unique in their own way. Once the substrate is recognized, the active sites of these enzymes prepare to bind the substrate in a productive binding mode for catalysis to occur. This is best exemplified by UGM. Though the sequence identity is very low, these mutases can adapt their active sites to bind their respective substrates and cofactors. Some directions for further studying these enzymes are suggested.

6.4.1 Studies on the allosteric site of *Mt*UGM

The model proposed by the Pinto group (Simon Fraser University, British Columbia) for MS-208 binding to the allosteric site suggests that the inhibitor interacts with hydrophobic residues such as Pro246, Trp260, Ala320 and Pro326 of the allosteric site (Figure 6-1). The backbone carbonyl oxygen of Glu321 forms hydrogen bonds with the hydroxyl group of MS-208. To study

how they affect the binding of MS-208 with the enzyme, alanine mutants of these allosteric site residues can be created and evaluated kinetically with and without MS-208. This study will help gain some insight into the location of the allosteric site, as a complete loss in MS-208 inhibition for an alanine mutant of a residue in the allosteric site may provide an indirect evidence for the location of this site.

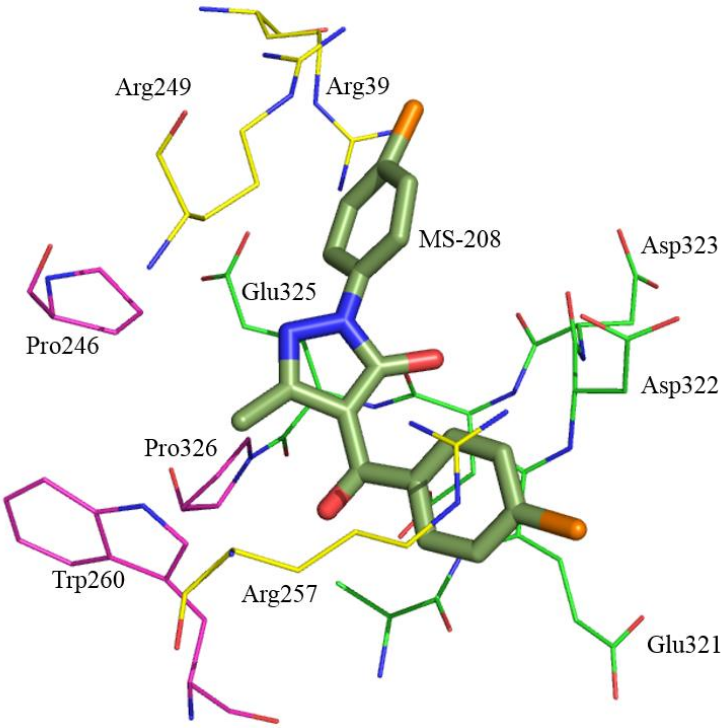


Figure 6-1: The proposed binding site of MS-208 in *Mt*UGM.
MS-208 is depicted as a green stick model. The hydrophobic, positively charged and negatively charged residues around the site are shown in magenta, yellow and green lines respectively.

Since this is the first instance of identification of an allosteric site on UGM, obtaining a crystal structure of *Mt*UGM with the inhibitor MS-208, will not only confirm the presence of the allosteric site but will also indicate how the inhibitor interacts with the enzyme and the changes undergone by the enzyme's active site due to binding of the inhibitor. However, MS-208 is poorly soluble in water, ruling out preparing solutions of high concentrations, a requisite for performing

co-crystallization experiments. More soluble versions of MS-208 can be synthesized and used for crystallization trials, to obtain crystals of *Mt*UGM, complexed with substrate in the active site and the inhibitor at the allosteric site. This will not only confirm the identification of an allosteric site but will also provide indications of whether more than one molecule of MS-208 can bind *Mt*UGM.

Active site residues have been studied as part of this research to understand their role in substrate binding. The active site of UGM is incomplete unless the mobile loop closes around the substrate, as evidenced by the UDP-Gal*p* bound structures of this enzyme. This closure helps position and orients the substrate in a productive conformation within the active site. In the substrate bound crystal structure of *Mt*UGM, Trp166 and Arg261 have been hypothesized as residues which control the closure of the mobile loop. In the open loop conformation, Trp166 and Arg261 form a cation-π interaction, which is broken due to a 180° rotation of Arg261, to let the substrate enter the active site. Hence, to test the hypothesis mutational and kinetic analysis of these critical residues need to be performed.

6.4.2 Crystallization and EXAFS on UAM

Determining the crystal structures of *Os*UAM1 and/or *At*RGPs will in fact be the first reported structures of plant UAM. This will give significant inputs into the mechanism of UAM and how its catalytic mechanism compares to that of UGM. If crystallizing these enzymes still remains tedious, more EXAFS studies can be performed on all four enzymes, to compare and contrast their metal-binding region and ascertain the sugar (Ara*p*) binding residues.

6.4.3 GaHM

Crystallizing *Cj*GaHM with its substrate GDP-6d-*altro*Hep*p* may stabilize its flexible loops, leading to further understanding of the substrate binding mode of this enzyme and how it compares to other known UGMs.

References

1. Peltier, P.; Euzen, R.; Daniellou, R.; Nugier-Chauvin, C.; Ferrieres, V., *Carbohydr. Res.* **2008,** *343,* 1897-1923.
2. Nassau, P. M.; Martin, S. L.; Brown, R. E.; Weston, A.; Monsey, D.; McNeil, M. R.; Duncan, K., *J. Bacteriol.* **1996,** *178,* 1047-1052.
3. Koplin, R.; Brisson, J. R.; Whitfield, C., *J. Biol. Chem.* **1997,** *272,* 4121-4128.
4. Sarvas, M.; Nikaido, H., *J. Bacteriol.* **1971,** *105,* 1063-1072.
5. Brennan, P. J.; Nikaido, H., *Annu. Rev. Biochem.* **1995,** *64,* 29-63.
6. Mergaert, P.; D'Haeze, W.; Fernandez-Lopez, M.; Geelen, D.; Goethals, K.; Prome, J. C.; Van Montagu, M.; Holsters, M., *Mol. Microbiol.* **1996,** *21,* 409-419.
7. Kus, J. V.; Kelly, J.; Tessier, L.; Harvey, H.; Cvitkovitch, D. G.; Burrows, L. L., *J. Bacteriol.* **2008,** *190,* 7464-7478.
8. Ridley, B. L.; O'Neill, M. A.; Mohnen, D., *Phytochemistry* **2001,** *57,* 929-967.
9. Pan, F.; Jackson, M.; Ma, Y.; McNeil, M., *J. Bacteriol.* **2001,** *183,* 3991-3998.
10. Pedersen, L. L.; Turco, S. J., *Cell. Mol. Life Sci.* **2003,** *60,* 259-266.
11. Schmalhorst, P. S.; Krappmann, S.; Vervecken, W.; Rohde, M.; Muller, M.; Braus, G. H.; Contreras, R.; Braun, A.; Bakker, H.; Routier, F. H., *Eukaryotic Cell* **2008,** *7,* 1268-1277.
12. Kleczka, B.; Lamerz, A. C.; van Zandbergen, G.; Wenzel, A.; Gerardy-Schahn, R.; Wiese, M.; Routier, F. H., *J. Biol. Chem.* **2007,** *282,* 10498-10505.
13. Oppenheimer, M.; Valenciano, A. L.; Sobrado, P., *Enzyme Res.* **2011,** *2011,* 415976.
14. Frey, P. A., *FASEB J.* **1996,** *10,* 461-470.
15. Thoden, J. B.; Holden, H. M., *Biochemistry* **1998,** *37,* 11469-11477.

16. Aspinall, G. O.; Cottrell, I. W.; Matheson, N. K., *Can. J. Biochem.* **1972,** *50,* 574-580.

17. Reiter, W. D.; Vanzin, G. F., *Plant Mol. Biol.* **2001,** *47,* 95-113.

18. Konishi, T.; Takeda, T.; Miyazaki, Y.; Ohnishi-Kameyama, M.; Hayashi, T.; O'Neill, M. A.; Ishii, T., *Glycobiology* **2007,** *17,* 345-354.

19. Allos, B. M., *Clin. Infect. Dis.* **2001,** *32,* 1201-1206.

20. Hanniffy, O. M.; Shashkov, A. S.; Moran, A. P.; Senchenkova, S. N.; Savage, A. V., *Carbohydr. Res.* **2001,** *330,* 223-229.

21. Weston, A.; Stern, R. J.; Lee, R. E.; Nassau, P. M.; Monsey, D.; Martin, S. L.; Scherman, M. S.; Besra, G. S.; Duncan, K.; McNeil, M. R., *Tuber. Lung. Dis.* **1997,** *78,* 123-131.

22. Richards, M. R.; Lowary, T. L., *ChemBioChem* **2009,** *10,* 1920-1938.

23. Sanders, D. A.; Staines, A. G.; McMahon, S. A.; McNeil, M. R.; Whitfield, C.; Naismith, J. H., *Nat. Struct. Biol.* **2001,** *8,* 858-863.

24. Partha, S. K.; van Straaten, K. E.; Sanders, D. A., *J. Mol. Biol.* **2009,** *394,* 864-877.

25. Soltero-Higgin, M.; Carlson, E. E.; Gruber, T. D.; Kiessling, L. L., *Nat. Struct. Mol. Biol.* **2004,** *11,* 539-543.

26. Huang, Z.; Zhang, Q.; Liu, H. W., *Bioorg. Chem.* **2003,** *31,* 494-502.

27. Sun, H. G.; Ruszczycky, M. W.; Chang, W. C.; Thibodeaux, C. J.; Liu, H. W., *J. Biol. Chem.* **2012,** *287,* 4602-4608.

28. Gruber, T. D.; Westler, W. M.; Kiessling, L. L.; Forest, K. T., *Biochemistry* **2009,** *48,* 9171-9173.

29. Hanukoglu, I., *Biochem. Mol. Biol. Educ.* **2015,** *43,* 206-209.

30. Rao, S. T.; Rossmann, M. G., *J. Mol. Biol.* **1973,** *76,* 241-&.

31. Beis, K.; Srikannathasan, V.; Liu, H.; Fullerton, S. W.; Bamford, V. A.; Sanders, D. A.; Whitfield, C.; McNeil, M. R.; Naismith, J. H., *J. Mol. Biol.* **2005,** *348,* 971-982.

32. van Straaten, K. E.; Kuttiyatveetil, J. R.; Sevrain, C. M.; Villaume, S. A.; Jimenez-Barbero, J.; Linclau, B.; Vincent, S. P.; Sanders, D. A., *J. Am. Chem. Soc.* **2015,** *137,* 1230-1244.

33. Tanner, J. J.; Boechi, L.; Andrew McCammon, J.; Sobrado, P., *Arch. Biochem. Biophys.* **2014,** *544,* 128-141.

34. van Straaten, K. E.; Routier, F. H.; Sanders, D. A., *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2012,** *68,* 455-459.

35. van Straaten, K. E.; Routier, F. H.; Sanders, D. A., *J. Biol. Chem.* **2012,** *287,* 10780-10790.

36. Dhatwalia, R.; Singh, H.; Oppenheimer, M.; Sobrado, P.; Tanner, J. J., *Biochemistry* **2012,** *51,* 4968-4979.

37. Dhatwalia, R.; Singh, H.; Oppenheimer, M.; Karr, D. B.; Nix, J. C.; Sobrado, P.; Tanner, J. J., *J. Biol. Chem.* **2012,** *287,* 9041-9051.

38. Yuan, Y.; Wen, X.; Sanders, D. A.; Pinto, B. M., *Biochemistry* **2005,** *44,* 14080-14089.

39. Yuan, Y.; Bleile, D. W.; Wen, X.; Sanders, D. A.; Itoh, K.; Liu, H. W.; Pinto, B. M., *J. Am. Chem. Soc.* **2008,** *130,* 3157-3168.

40. Chad, J. M.; Sarathy, K. P.; Gruber, T. D.; Addala, E.; Kiessling, L. L.; Sanders, D. A., *Biochemistry* **2007,** *46,* 6723-6732.

41. Ghavami, A.; Chen, J. J.; Mario Pinto, B., *Carbohydr. Res.* **2004,** *339,* 401-407.

42. Veerapen, N.; Yuan, Y.; Sanders, D. A.; Pinto, B. M., *Carbohydr. Res.* **2004,** *339,* 2205-2217.

43. Desvergnes, S.; Desvergnes, V.; Martin, O. R.; Itoh, K.; Liu, H. W.; Py, S., *Bioorg. Med. Chem.* **2007,** *15,* 6443-6449.

44. Scherman, M. S.; Winans, K. A.; Stern, R. J.; Jones, V.; Bertozzi, C. R.; McNeil, M. R., *Antimicrob. Agents Chemother.* **2003,** *47,* 378-382.

45. Caravano, A.; Sinay, P.; Vincent, S. P., *Bioorg. Med. Chem. Lett.* **2006,** *16,* 1123-1125.

46. Barlow, J. N.; Blanchard, J. S., *Carbohydr. Res.* **2000,** *328,* 473-480.

47. Burton, A.; Wyatt, P.; Boons, G. J., *J. Chem. Soc. Perkin Trans. 1* **1997,** 2375-2382.

48. Caravano, A.; Vincent, S. P., *Eur. J. Org. Chem.* **2009,** 1771-1780.

49. El Bkassiny, S.; N'Go, I.; Sevrain, C. M.; Tikad, A.; Vincent, S. P., *Org. Lett.* **2014,** *16,* 2462-2465.

50. Partha, S. K.; Sadeghi-Khomami, A.; Slowski, K.; Kotake, T.; Thomas, N. R.; Jakeman, D. L.; Sanders, D. A., *J. Mol. Biol.* **2010,** *403,* 578-590.

51. Carlson, E. E.; May, J. F.; Kiessling, L. L., *Chem. Biol.* **2006,** *13,* 825-837.

52. Dykhuizen, E. C.; May, J. F.; Tongpenyai, A.; Kiessling, L. L., *J. Am. Chem. Soc.* **2008,** *130,* 6706-6707.

53. Kincaid, V. A.; London, N.; Wangkanont, K.; Wesener, D. A.; Marcus, S. A.; Heroux, A.; Nedyalkova, L.; Talaat, A. M.; Forest, K. T.; Shoichet, B. K.; Kiessling, L. L., *ACS Chem. Biol.* **2015,** *10,* 2209-2218.

54. Borrelli, S.; Zandberg, W. F.; Mohan, S.; Ko, M.; Martinez-Gutierrez, F.; Partha, S. K.; Sanders, D. A.; Av-Gay, Y.; Pinto, B. M., *Int. J. Antimicrob. Agents* **2010,** *36,* 364-368.

55. Poulin, M. B.; Lowary, T. L., *Methods Enzymol.* **2010,** *478,* 389-411.

56. St Michael, F.; Szymanski, C. M.; Li, J.; Chan, K. H.; Khieu, N. H.; Larocque, S.; Wakarchuk, W. W.; Brisson, J. R.; Monteiro, M. A., *Eur. J. Biochem.* **2002,** *269,* 5119-5136.

57. Poulin, M. B.; Nothaft, H.; Hug, I.; Feldman, M. F.; Szymanski, C. M.; Lowary, T. L., *J. Biol. Chem.* **2010,** *285,* 493-501.

58. Poulin, M. B. Chemical Studies Exploring the Specificity of Bacterial Furanoside Biosynthesis. University of Alberta, 2012.

59. Wang, Q.; Ding, P.; Perepelov, A. V.; Xu, Y.; Wang, Y.; Knirel, Y. A.; Wang, L.; Feng, L., *Mol. Microbiol.* **2008,** *70,* 1358-1367.

60. Carpita, N. C.; Gibeaut, D. M., *Plant J.* **1993,** *3,* 1-30.

61. Ishii, T.; Konishi, T.; Ito, Y.; Ono, H.; Ohnishi-Kameyama, M.; Maeda, I., *Phytochemistry* **2005,** *66,* 2418-2425.

62. Fry, S. C.; Northcote, D. H., *Plant Physiol.* **1983,** *73,* 1055-1061.

63. Konishi, T.; Ono, H.; Ohnishi-Kameyama, M.; Kaneko, S.; Ishii, T., *Plant Physiol.* **2006,** *141,* 1098-1105.

64. Dhugga, K. S.; Ulvskov, P.; Gallagher, S. R.; Ray, P. M., *J. Biol. Chem.* **1991,** *266,* 21977-21984.

65. Langeveld, S. M.; Vennik, M.; Kottenhagen, M.; Van Wijk, R.; Buijk, A.; Kijne, J. W.; de Pater, S., *Plant Physiol.* **2002,** *129,* 278-289.

66. Rautengarten, C.; Ebert, B.; Herter, T.; Petzold, C. J.; Ishii, T.; Mukhopadhyay, A.; Usadel, B.; Scheller, H. V., *Plant Cell* **2011,** *23,* 1373-1390.

67. Goddard, E. A.; Lastovica, A. J.; Argent, A. C., *Arch. Dis. Child.* **1997,** *76,* 526-528.

68. Guerry, P.; Szymanski, C. M., *Trends Microbiol.* **2008,** *16,* 428-435.

69. Guerry, P.; Poly, F.; Riddle, M.; Maue, A. C.; Chen, Y. H.; Monteiro, M. A., *Front. Cell Infect. Microbiol.* **2012,** *2,* 7.

70. Hanniffy, O. M.; Shashkov, A. S.; Moran, A. P.; Prendergast, M. M.; Senchenkova, S. N.; Knirel, Y. A.; Savage, A. V., *Carbohydr. Res.* **1999,** *319,* 124-132.

71. Karlyshev, A. V.; Champion, O. L.; Churcher, C.; Brisson, J. R.; Jarrell, H. C.; Gilbert, M.; Brochu, D.; St Michael, F.; Li, J.; Wakarchuk, W. W.; Goodhead, I.; Sanders, M.; Stevens, K.; White, B.; Parkhill, J.; Wren, B. W.; Szymanski, C. M., *Mol. Microbiol.* **2005,** *55,* 90-103.

72. Nallamsetty, S.; Waugh, D. S., *Nat. Protoc.* **2007,** *2,* 383-391.

73. Chayen, N. E., *Structure* **1997,** *5,* 1269-1274.

74. Karunan Partha, S.; Bonderoff, S. A.; van Straaten, K. E.; Sanders, D. A., *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2009,** *65,* 843-845.

75. Kabsch, W., *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010,** *66,* 125-132.

76. Fodje, M.; Pawel, G.; Janzen, K.; Labiuk, S.; James, G.; Russ, B., *J. Synchrotron Radiat.* **2014,** *21,* 633-637.

77. Vagin, A.; Teplyakov, A., *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010,** *66,* 22-25.

78. Winn, M. D., *J. Synchrotron Radiat.* **2003,** *10,* 23-25.

79. Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S., *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011,** *67,* 235-242.

80. Vagin, A.; Teplyakov, A., *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2000,** *56,* 1622-1624.

81. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.;

Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010,** *66,* 213-221.

82. Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L. W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C., *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2002,** *58,* 1948-1954.

83. Emsley, P.; Cowtan, K., *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004,** *60,* 2126-2132.

84. Moriarty, N. W.; Grosse-Kunstleve, R. W.; Adams, P. D., *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2009,** *65,* 1074-1080.

85. Smith, C. A.; Kortemme, T., *J. Mol. Biol.* **2008,** *380,* 742-756.

86. Lauck, F.; Smith, C. A.; Friedland, G. F.; Humphris, E. L.; Kortemme, T., *Nucleic Acids Res.* **2010,** *38,* W569-575.

87. Ko, J.; Park, H.; Heo, L.; Seok, C., *Nucleic Acids Res.* **2012,** *40,* W294-297.

88. Zhang, Y., *BMC Bioinformatics* **2008,** *9,* 40.

89. Jones, G.; Willett, P.; Glen, R. C., *J. Mol. Biol.* **1995,** *245,* 43-53.

90. Verdonk, M. L.; Cole, J. C.; Hartshorn, M. J.; Murray, C. W.; Taylor, R. D., *Proteins: Struct., Funct., Genet.* **2003,** *52,* 609-623.

91. Nissink, J. W.; Murray, C.; Hartshorn, M.; Verdonk, M. L.; Cole, J. C.; Taylor, R., *Proteins* **2002,** *49,* 457-471.

92. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R., *J. Mol. Biol.* **1997,** *267,* 727-748.

93. Gruber, T. D.; Borrok, M. J.; Westler, W. M.; Forest, K. T.; Kiessling, L. L., *J. Mol. Biol.* **2009,** *391,* 327-340.

94. Karunan Partha, S. Structural and inhibition studies on UDP-galactopyranose mutase. University of Saskatchewan, 2010.

95. Haebel, P. W.; Wichman, S.; Goldstone, D.; Metcalf, P., *J. Struct. Biol.* **2001,** *136,* 162-166.

96. Heras, B.; Martin, J. L., *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2005,** *61,* 1173-1180.

97. Heras, B.; Edeling, M. A.; Byriel, K. A.; Jones, A.; Raina, S.; Martin, J. L., *Structure* **2003,** *11,* 139-145.

98. Yun Shi, Kuttiyatveetil., J. R., Karin E. van Straaten, Cinzia Colombo,David A. R. Sanders,and B.Mario Pinto, **2015**.

99. Cornish-Bowden, A., *Biochem. J.* **1974,** *137,* 143-144.

100. Konishi, T.; Miyazaki, Y.; Yamakawa, S.; Iwai, H.; Satoh, S.; Ishii, T., *Biosci. Biotechnol. Biochem.* **2010,** *74,* 191-194.

101. Konishi, T.; Ohnishi-Kameyama, M.; Funane, K.; Miyazaki, Y.; Konishi, T.; Ishii, T., *Carbohydr. Res.* **2010,** *345,* 787-791.

102. Yang, J.; Zhang, Y., *Nucleic Acids Res.* **2015,** *43,* W174-181.

103. Roy, A.; Kucukural, A.; Zhang, Y., *Nat. Protoc.* **2010,** *5,* 725-738.

104. Osawa, T.; Sugiura, N.; Shimada, H.; Hirooka, R.; Tsuji, A.; Shirakawa, T.; Fukuyama, K.; Kimura, M.; Kimata, K.; Kakuta, Y., *Biochem. Biophys. Res. Commun.* **2009,** *378,* 10-14.

105. Diakun, G. P., *Nature* **1990,** *344,* 83-84.

Supplementary Notes

S.1 Modeling of *Dr*UGM mutants using Rosetta-Backrub

*Dr*UGM mutants were modeled using an online software Rosetta-Backrub.[86,88] This software calculates the best orientation of the mutant residues by taking into account the flexibility of the residues around the point mutation and modeling the backbone chains when the mutant models are generated. The backbone chains are modeled by the "Backrub" step, which picks two amino acids along a chain length of 2-12 residues apart and rotates it along an axis drawn along the Cα of the residues. At the start, the program decides if it can pick only a single residue (residue-only) or select a backbone (backbone move), based on a probability ($P_{rotamer}$) score. The default score for choosing a rotamer-only move is 0.25. If this is chosen, the program just picks up the single residue and applies random rotamers to the residue, based on a backbone-dependant rotamer library. This library is based on crystallographic structures that have already been solved. If the rotamer-only move is not selected, then a segment of the backbone is chosen, and rotation is performed. The backbone rotation is performed using the backrub step and is called as the backbone only move. Next, the algorithm decides if it can stop based on the second adjustable probability, $P_{backbone}$. To make the selection of this move more frequent, the default score for $P_{backbone}$ is set at 0.75. One or a couple of the residues are selected along the chosen backbone segment, and random rotamers are applied. Rosetta full-atom scoring function and Metropolis criterion are then used to evaluate the move. Backrub moves are applied to any heavy atoms within the range of 6 A° around the desired point mutation. A flowchart describing the procedure adopted is shown in Figure S-1.

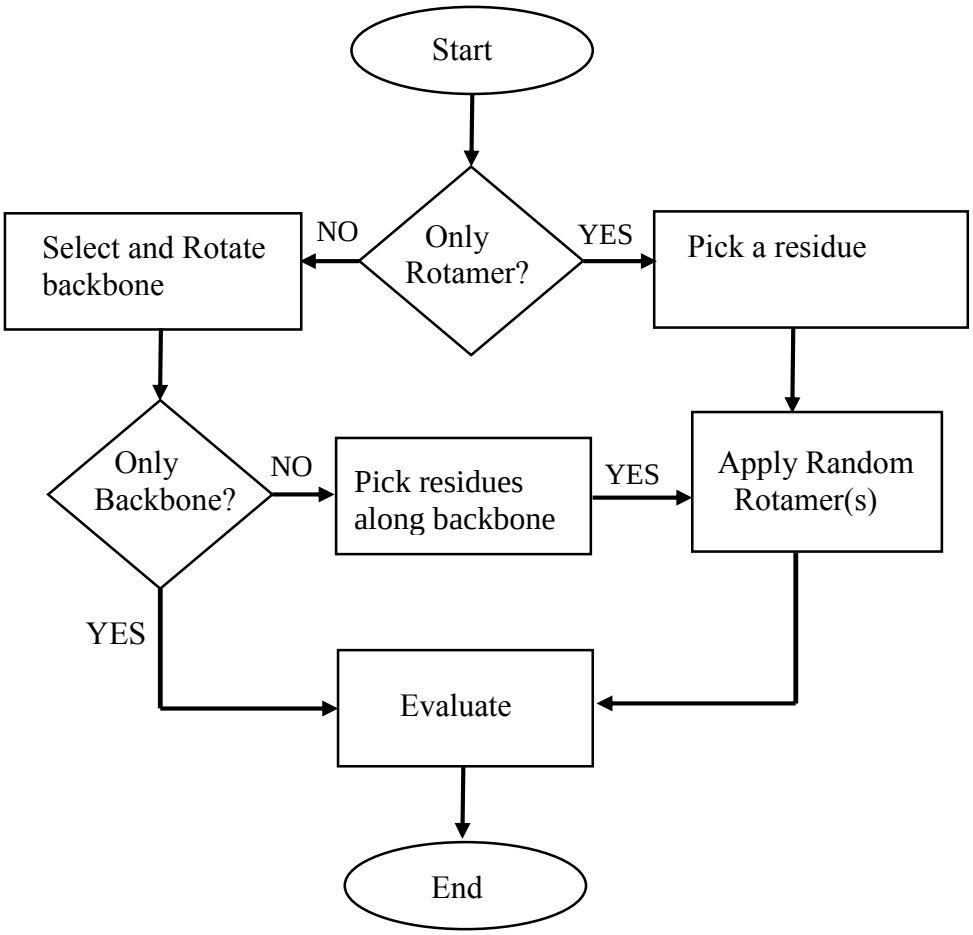


Figure S-1: Flowchart describing the method adopted by Rosetta-Backrub software.

All the requisite information, such as starting angles and backbone parameters, are obtained from the input structure file provided in the Protein Data Bank (PDB) format. Ligands are not included for performing these simulations. Rosetta Backrub can be accessed online at: https://kortemmelab.ucsf.edu/backrub/cgi-bin/rosettaweb.py?query=index.

S.2 GOLD docking for *Dr*UGM mutants

Docking of the substrate (UDP-Gal*p*) into the active site of the *Dr*UGM mutants was performed using the GOLD software (version 5.2.2).[89-92] One of the features of the software,

'Ensemble docking' was used to get a direct comparison and analysis of the results from the wild-type and mutant enzymes. At first, an ensemble of proteins was created by superimposing the structures considered for the docking experiment, and the substrate was then docked into the active site of the ensemble created. Typically, the crystal structure of the wild-type enzyme was superimposed on the modeled structure of mutants considered for the experiment. UDP-Gal*p* was then docked back into the active site of the ensemble, so that it can now simultaneously get docked in the active site of all the protein structures, in a single run. This method is advantageous as it generates numbers that can give a direct comparison of which protein binds the substrate better and generates docking poses for each of the protein in the ensemble.

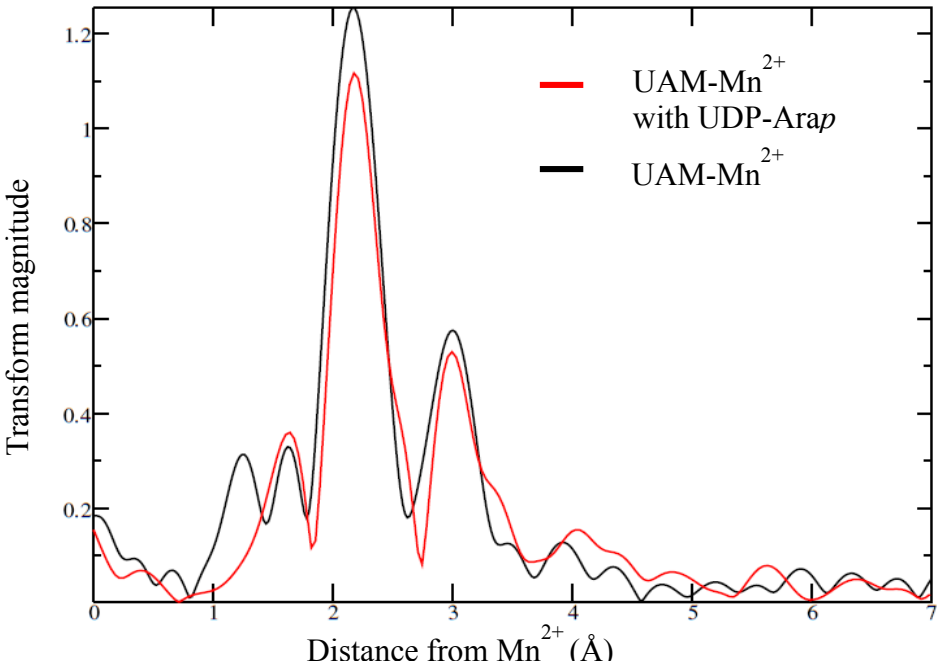
S.3 EXAFS on *Os*UAM1

S.3.1 Sample preparation

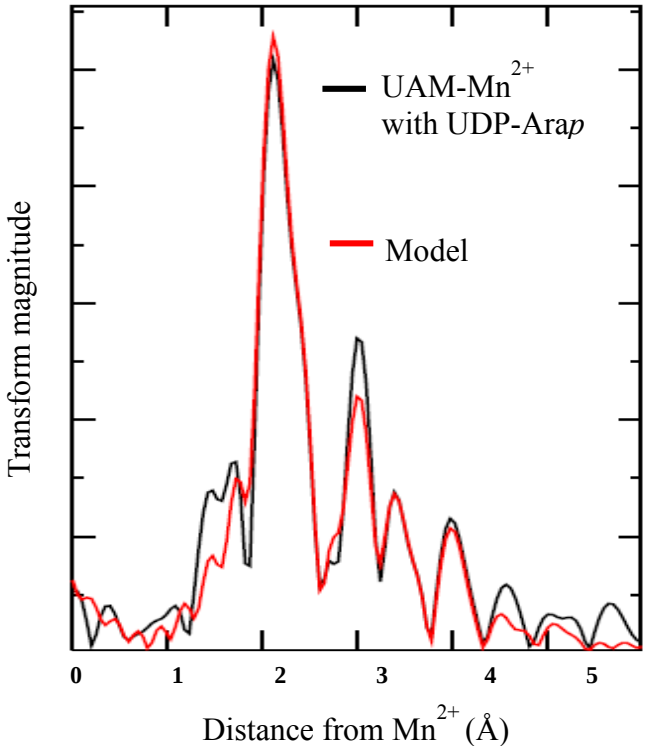
For EXAFS experiments, the histidine tag was removed. TEV protease was added to the purified *Os*UAM1 with the 6xhistidine tag and left to react overnight at 4 °C. The sample was then loaded onto a Protino column previously equilibrated with 25 mM Bicine pH 8.5. *Os*UAM1 without the His-tag was collected in the flow through. The fractions were concentrated to ~ 25 mg/ml in 25 mM Bicine pH 8.5. Two samples were prepared for the EXAFS study. To gain insights about the residues involved in the binding of the metal cofactor, *Os*UAM1 samples were prepared with 0.5 mM $MnCl_2$. To understand the change in metal coordination after addition of substrate, UDP-Ara*p* was added to *Os*UAM1 samples with Mn^{2+}. 30% glycerol was added to all samples before freezing, decreasing the protein concentration to ~ 17 mg/ml.

S.3.2 Results

Since UAM is a metalloenzyme, metal absorption edge EXAFS was used to improve accuracy and supplement the metal binding site prediction made by the I-TASSER model. EXAFS was used to map the active site residues coordinating the metal and predict how the substrate interacts with the metal. A change in spectra in the presence of substrate binding is indicative of changes around the metal when the substrate enters the active site. Additionally, EXAFS was used to determine the distances between the metal and the neighboring atoms.[105] Data obtained for two different samples of *Os*UAM1, UAM-Mn^{2+} and UAM-Mn^{2+}-UDP-Ara*p,* were cut off at a k range of 12. Some important differences in EXAFS spectra obtained with and without substrate were observed at ~2.1 Å, ~2.5 Å and 3.4 - 3.6 Å away from the metal, as shown in Figure S-2A. These differences in the EXFAS spectra are due to the metal making new interactions with the incoming substrate. Theoretical models were designed and fit to UAM-Mn^{2+}-UDP-Ara*p* EXAFS data that was obtained using FEFF, a software used to fit experimental spectra to theoretical calculations. A few theoretical models with different combinations of the DDD motif, metal, diphosphate and water were made, and EXAFS data for all of them were simulated using FEFF and fit to the original EXAFS data obtained. A model that agrees reasonably well with the original EXAFS data is shown in Figure S-2B and C. This model, with residues Asp110, Asp112, the di-phosphate of substrate and two water molecules, accounts for most of the peaks in the experimental data, as shown in Figure S-2B. The metal, Mn^{2+} makes an octahedral coordination. The peak at ~ 2.2 Å is due to six separate contributions; one each from carboxyl oxygens of Asp110 and Asp112 residues, one from each phosphate oxygen of the substrate and two adjacent water molecules. The peak at 3.4 - 3.6 Å can be explained by the distance between phosphorous atoms of the substrate and Mn^{2+}. Further experimental data is required to improve this theoretical model further.


Transform magnitude
1.2
1
0.8
0.6
0.4
0.2
0
1
2
3
4
5
6
7
UAM-Mn2+
with UDP-Arap
UAM-Mn2+
Distance from Mn2+ (Å)

(A)


Transform magnitude
0
1
2
3
4
5
UAM-Mn2+
with UDP-Arap
Model
Distance from Mn2+ (Å)

(B)

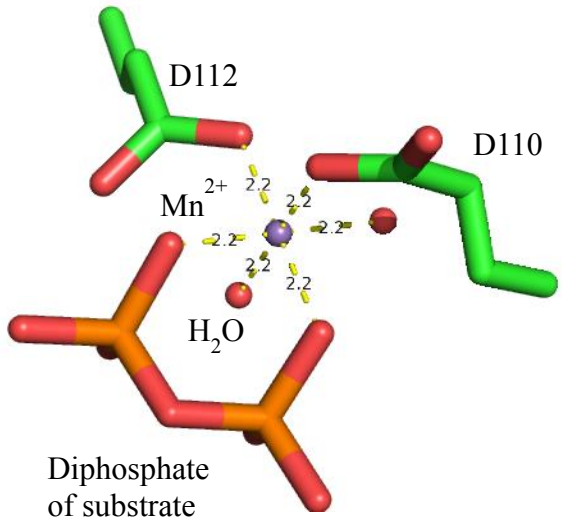


Figure S-2: EXAFS studies on *Os*UAM1.
(A) EXAFS data for UAM with Mn^{2+} (trace in black) and UAM-Mn^{2+}-UDP-Ara*p* (trace in red) (B) EXAFS data for UAM-Mn^{2+}-UDP-Ara*p* (trace in black) overlapped on the FEFF spectra produced from the model (trace in red) (C) Model designed with two aminoacid residues (D110 and D112), both in green; water molecules in red, Mn^{2+} in purple and the substrate diphosphate in orange.

S.4 *Os*UAM1 Modeling

GalaxyWEB and I-TASSER were used to generate 3D-models of *Os*UAM1.[87,88] I-TASSER server predicts the 3D-structure and function of a protein of interest based on multiple threading alignments and structural simulations. The threading procedure is performed by the locally installed meta-threading-server (LOMETS), having individual alignment and scoring programs, which compares regions of the input sequence to structures found in PDB. Regions of the input sequence, for which templates are found, are incorporated into different structural assemblies and the regions of the sequence that do not have a template, mostly loops, are built in by modeling. Restraints from LOMETS, based on the PDB structure-templates chosen to run simulations at different temperatures, are incorporated into the structural assemblies. The various structural assemblies are clustered, and cluster centroids are picked for the next stage, which

focuses on removing steric clashes and refinement of global topology. This is established by incorporating only the restraints from LOMETS to the PDB structures closest to the clusters chosen. Another simulation is performed for hydrogen-bond optimization so that only the lowest energy structures produce the final models. The function of the protein of interest is also determined by matching the final predicted models with structures from PDB having known functions. I-TASSER can be accessed online at: http://zhanglab.ccmb.med.umich.edu/I-TASSER/.

GalaxyWEB predicts the protein structure of a given sequence using multiple template-based modeling (GalaxyTBM) and refines unreliable loop regions by optimization-based refinement (GalaxyREFINE). The software has four different stages such as template selection, sequence alignment, model building and refinement. The multiple template selection and scoring are performed by HHsearch. The matching regions are chosen as templates and the regions of the input sequence that do not have templates are built in. Initial model structures are built using restraints obtained based on templates and multiple sequence alignments. The unreliable loop regions, defined as URLs are then built using loop-closure algorithms. GalaxyREFINE is used for further refinement of the predicted models. GalaxyWEB can be accessed online at: http://galaxy.seoklab.org/.

PERMISSION TO REUSE: FIGURE 4-3

OXFORD UNIVERSITY PRESS LICENSE
TERMS AND CONDITIONS
May 09, 2016

This is a License Agreement between Jijin Raj Kuttiyatveetil ("You") and Oxford University Press ("Oxford University Press") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Oxford University Press, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3864920485040
License date	May 09, 2016
Licensed content publisher	Oxford University Press
Licensed content publication	Glycobiology
Licensed content title	A plant mutase that interconverts UDP-arabinofuranose and UDP-arabinopyranose:
Licensed content author	Teruko Konishi, Takumi Takeda, Yasumasa Miyazaki, Mayumi Ohnishi-Kameyama, Takahisa Hayashi, Malcolm A. O'Neill, Tadashi Ishii
Licensed content date	03/01/2007
Type of Use	Thesis/Dissertation
Institution name	University of Saskatchewan
Title of your work	STRUCTURAL AND FUNCTIONAL ASPECTS OF PYRANOSE-FURANOSE MUTASES
Publisher of your work	n/a
Expected publication date	May 2016
Permissions cost	0.00 USD
Value added tax	0.00 USD
Total	0.00 USD
Total	0.00 USD

Terms and Conditions

STANDARD TERMS AND CONDITIONS FOR REPRODUCTION OF MATERIAL FROM AN OXFORD UNIVERSITY PRESS JOURNAL

1. Use of the material is restricted to the type of use specified in your order details. **2**. This permission covers the use of the material in the English language in the following territory: world. If you have requested additional permission to translate this material, the terms and conditions of this reuse will be set out in clause 12. **3**. This permission is limited to the particular use authorized in (1) above and does not allow you to sanction its use elsewhere in any other format other than specified above, nor does it apply to quotations, images, artistic works etc that have been reproduced from other sources which may be part of the material to be used. **4**. No alteration, omission or addition is made to the material without our written consent. Permission must be re-cleared with Oxford University Press if/when you decide to reprint. **5**. The following credit line appears wherever the material is used: author, title, journal, year, volume, issue number, pagination, by permission of Oxford University Press or the sponsoring society if the journal is a society journal. Where a journal is being published on behalf of a learned society, the details of that society must be included in the credit line. **6**. For the reproduction of a full article from an Oxford University Press journal for whatever purpose, the corresponding author of the material concerned should be informed of the proposed use. Contact details for the corresponding authors of all Oxford University Press journal contact can be found alongside either the abstract or full text of the article concerned, accessible from www.oxfordjournals.org Should there be a problem clearing these rights, please contact journals.permissions@oup.com **7**. If the credit line or acknowledgement in our publication indicates that any of the figures, images or photos was reproduced, drawn or modified from an earlier source it will be necessary for you to clear this permission with the original publisher as well. If this permission has not been obtained, please note that this material cannot be included in your publication/photocopies. **8**. While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by Oxford University Press or by Copyright Clearance Center (CCC)) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and Oxford University Press reserves the right to take any and all action to protect its copyright in the materials. **9**. This license is personal to you and may not be sublicensed, assigned or transferred by you to any other person without Oxford University Press's written permission. **10**. Oxford University Press reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions. **11**. You hereby indemnify and agree to hold harmless Oxford University Press and CCC, and their respective officers, directors, employs and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license. **12**. Other Terms and Conditions: v1.4