Advancing mechanistic understanding of glycosyltransferases

by

Susannah Melanie Lynn Gagnon
BSc, University of Victoria, 2014

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in the Department of Biochemistry and Microbiology

Advancing mechanistic understanding of glycosyltransferases

by

Susannah Melanie Lynn Gagnon
BSc, University of Victoria, 2014

Supervisory Committee

Dr. Stephen V. Evans, Department of Biochemistry and Microbiology
Supervisor

Dr. Alisdair Boraston, Department of Biochemistry and Microbiology
Departmental Member

Dr. Rodney Herring, Department of Mechanical Engineering
Outside Member

Dr. Monica Palcic, Department of Biochemistry and Microbiology
Departmental Member

Abstract

Supervisory Committee

Dr. Stephen V. Evans, Department of Biochemistry and Microbiology
Supervisor

Dr. Alisdair Boraston, Department of Biochemistry and Microbiology
Departmental Member

Dr. Rodney Herring, Department of Mechanical Engineering
Outside Member

Dr. Monica Palcic, Department of Biochemistry and Microbiology
Departmental Member

Glycosyltransferase enzymes synthesize glycosidic linkages, generating carbohydrates and carbohydrate-linked entities ranging from cellulose, starch, and chitin to glycolipids, glycopeptides, and natural product antibiotics. These syntheses involve stereo- and regio-specific sugar transfer from an activated donor molecule, often a UDP-sugar, to an acceptor molecule. Functionally, glycosyltransferases are classified as either "retaining" or "inverting" enzymes depending on whether the stereochemical linkage of the donor substrate is conserved in the product. While inverting glycosyltransfer is mechanistically straightforward, the retaining mechanism remains poorly understood. For retaining glycosyltransferases, the central question is whether transfer occurs via a front-face "S_Ni-like" mechanism or through a 'double displacement' mechanism that invokes a glycosyl-enzyme covalent intermediate.

GTA and GTB are retaining enzymes that catalyze the final step in human ABO(H) blood group A and B antigen synthesis through UDP-GalNAc or UDP-Gal transfer, respectively, to the H-antigen disaccharide acceptor. Although they have been intensively characterized, the processes of substrate recognition, mobile loop organization, and product release in GTA and GTB has long resisted explanation. Further, the question of the retaining enzyme mechanism persists, though the covalent intermediate of the proposed double displacement mechanism has been detected *via* mass spectrometry experiments with GTA/GTB mutants.

Building on previous investigations, we have aimed to characterize and have uncovered details of mechanism, substrate binding, loop organization, and product release using a combined kinetic and structural approach. These investigations are essential not only for understanding GTA, GTB, and retaining glycosyltransferases as a whole, but also for the rational design of inhibitors. Such inhibitors could selectively target, for example, bacterial glycosyltransferases and thus would represent a new class of antimicrobials.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AA	Amino Acid
ADA	N-(2-acetamido)iminodiacetic acid
A-tri	A-antigen terminal trisaccharide
B-tri	B-antigen terminal trisaccharide
BME	β-Mercaptoethanol
BoGT6a	*Bacteroides ovatus* α-N-acetylgalactosaminyl-transferase
CAZy	Carbohydrate Active enZyme database
CCP4	Collaborative Crystallography Project 4
CGEI	Covalent glycosylenzyme intermediate
COOT	Crystallographic Object-Oriented Toolkit
DI	Deoxy inhibitor: α-L-Fuc-(1→2)-β-D-(3-deoxy)-Gal*p*-O$(CH_2)_7CH_3$
DTT	Dithiothreitol
ESI-MS	Electrospray ionization mass spectrometry
Gal	Galactose
GalNAc	N-acetylgalactosamine
GalNAc-T2	Polypeptide GalNAc-transferase 2
GalT1	β-1,4-galactosyltransferase I
GH	Glycoside hydrolases
Glc	Glucose
GlcNAc	N-acetylglucosamine
GT	Glycosyltransferase
GT-A	Glycosyltranfserase fold A
GT-B	Glycosyltranfserase fold B
GT-C	Glycosyltranfserase fold C
GTA	A-specific α-1,3-N-acetylglucosaminyltransferase
GTB	B specific α-1,3-glucosyltransferase
HA	H antigen determinant: L-Fuc-(1→2)-β-D-Gal*p*-O$(CH_2)_7CH_3$
Inv-GT	Inverting glycosyltransferases
LgtC	α-galactosyltransferase
mAbs	Monoclonal antibodies
MAD	Multiple-wavelength anomalous dispersion
MD	Molecular dynamics
ML	Mother liquor
ML_o	Mother liquor initial
MOLREP	Molecular replacement
MOPS	3-(N-morpholino)propanesulfonic acid
MPD	2-methyl-2,4-pentanediol
NMR	Nuclear magnetic resonance
OA	Orthogonal associative
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PEG	Polyethylene glycol
QM/MM	Static hybrid quantum mechanics and molecular mechanics
Ret-GT	Retaining glycosyltransferases
Rh	Rhesus factor
RMSD	Root-mean-square deviation
STD NMR	Saturation transfer difference nuclear magnetic resonance
Tris	tris(hydroxymethyl)aminomethane
UCSF	University of California, San Francisco
UDP	Uridine diphosphate
α3GalT	Bovine α-1,3-galactosyltransferase

Acknowledgements

I gratefully acknowledge my supervisor, Prof. Stephen V. Evans, for his mentorship and support during my time in his lab. I also wish to thank my committee members, Profs. Monica Palcic, Alisdair Boraston, and Rodney Herring, for their extensive support and input during the past five years. I would like to thank all of the members of the Evans lab, past and present, who have made this research possible, especially Ryan Blackler and Omid Haji-Ghassemi, who showed me the ropes when I started.

Dedication

"There is no subject so old that something new cannot be said about it."

—Fyodor Dostoyevsky

Chapter 1: Introduction

1.1 An historical context for the study of the ABO(H) blood group enzymes

At the core of modern-day blood transfusion medicine is the ABO(H) blood group system. The discovery of ABO(H) antigens and their parent enzymes occurred over the course of several centuries, beginning with William Harvey's 1616 discourse on the circulation of blood. Though ancient civilizations theorized the possibility of blood transfusion, the first documented case was not until Richard Lower's experiment on animals in 1665 (Fastag, E., et al. 2013). The Royal Society, which was founded in 1662, published the procedure, entitled "The Method Observed in Transfusing the Bloud out of One Animal into Another", in the 1665 *Proceedings* for the first year. The protocol described how Richard Lower bled one dog nearly to death and then revived this dog *via* a blood transfusion from another, larger dog. Soon thereafter, transfusions were attempted in humans. Interestingly, the impetus for many of these early transfusions was mental health or temperamental problems (Sturgis, C.C. 1942).

It is unclear to whom the first human transfusion ought to be credited: Richard Lower or Jean Baptiste Denis, a physician to Louis XIV of France who had also conducted initial transfusions using dogs. In 1666, Lower, with the assistance of Edmund King, successfully transfused Arthur Coga, an English clergyman who "was a little cracked in the head" (Sturgis, C.C. 1942). However, months earlier Jean Baptiste Denis likely already had performed a blood transfusion on a patient with manic depression and psychosis. Denis used lamb's blood for the transfusion and unfortunately, the procedure failed. The patient reacted violently, experiencing chills, fever, and black urine, symptoms Denis attributed to "a reaction to incompatible blood".

There was a hiatus in blood transfusion research until the work of James Blundell in the early 19th century. Blundell, a professor of physiology and obstetrics, performed the first blood

transfusion from a human donor to a human recipient (Sturgis, C.C. 1942) in an effort to treat a dying patient suffering from stomach carcinoma. This outcome of the first transfusion attempt is uncertain, the patient having succumbed to his illness the following day. Blundell went on to treat ten other patients using the same procedure. Half of these patients died, and eventually Blundell retreated from research to lead a reclusive life in his London home. His work, and the work of other physicians who attempted transfusions in the nineteenth century, highlighted the problem of donor selection.

1.2 The ABO(H) blood group system

Before the discovery of ABO(H) antigens, physicians were performing transfusions without an understanding of blood type compatibility, and so these procedures were just as likely to kill patients as to save them. To make sense of these outcomes, Austrian physician and scientist Karl Landsteiner conducted experiments using the red cells and serum of his staff at the University of Vienna. He observed that the serum of certain individuals agglutinated the red blood cells of others, while yet different combinations, including the mixing of self-serum and self-red cells, produced no agglutination response. Through these tests, Landsteiner deduced the existence of the A, B, and C (later renamed O) blood types in the year 1900, as well as a fourth, AB, the following year with the help of his students. For this work he was awarded the Nobel Prize in Physiology and Medicine, though this recognition came in 1930, nearly three decades after he published his discovery.

Now, we know that the agglutination response Landsteiner observed occurs when circulating antibodies encounter their cognate antigens. For example, circulating anti-A antibodies in the serum of a type B individual interact with and agglutinate the red blood cells of

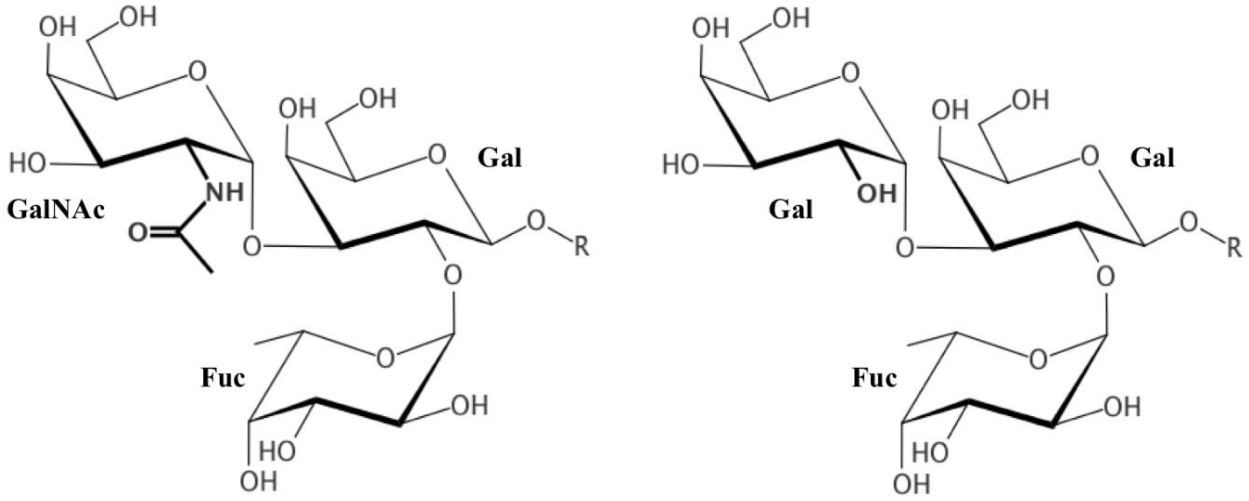
a type A individual. Remarkably, this procedure remains the foundation of modern-day blood typing.

Despite major headway in the field of blood transfusion medicine in the early twentieth century, the chemical structure of the A and B antigens proved elusive until the 1950s and 1960s (Watkins, W.M. 1991a).Several groups, labs mainly directed by Kabat, Morgan, and Watkins, defined the A, B, and O (unmodified H antigen) structures using varied approaches (reviewed in: (Watkins, W.M. 1972, Watkins, W.M. 2001). One approach involved acid and alkaline degradation of A, B, and O(H) oligosaccharides isolated from ovarian cyst fluid (Morgan, W.T.J. and Watkins, W.M. 1959, Kabat, E.A., et al. 1965). Another approach employed "blood-group-splitting enzymes", or exo-glycosidases, to establish the identities of the terminal, non-reducing sugars of the A and B antigens. This was made possible through the finding that fecal extracts (and other bodily fluids) could destroy blood group activity, a process thought to be enzymatic since exposure of the extracts to heat and acid nullified the effect. The researchers behind these experiments speculated that bacterial microorganisms were the source of these enzymes but were unable to obtain enzymatic activity from any cultured isolates. Of course, given current knowledge of bacteria as rich producers of these so-called "splitting" enzymes, or glycoside hydrolases, their insight is rather remarkable, although it remained unsubstantiated for many years.

Together these approaches defined N-acetyl-D-galactosamine, D-galactose, and L-fucose as the terminal non-reducing sugars in A and B blood groups (**Figure 1**) (Watkins, W.M. 1959, Watkins, W.M. 1991b). It was also glycosidase digestion that clarified the H antigen as the precursor to A and B (Watkins, W.M. 1991b). Later, a parallel set of experiments was conducted using red blood cells, which confirmed the experimental results obtained for secreted ABH; the

terminal sugars identified in ovarian cyst fluid were indeed identical to those found within the erythrocyte membrane. The complete chemical structures of ABH, not just their antigenic determinants, were uncovered in the 1970s and 80s (Lowe, J.B. 1993). Despite uniformity in the di- and trisaccharide antigenic determinants, there is some variability in the linkage(s) found within the ABH substructures. There are four main subtypes (1-4) of ABH precursors, and these can exist as free oligosaccharides, covalently linked to proteins through a Ser/Thr/Asn linkage or linked to membrane-associated lipids (Watkins, W.M. 1991a). One of the most fascinating outcomes of these early experiments was the realization that the A and B antigen determinants differ only in the C2 substituent of the terminal sugar (**Figure 1**), a small difference that can have fatal consequences in the event of a mismatched blood transfusion.

Figure 1. *Chemical structure of the A and B blood group determinants*



R is a glycoprotein, glycolipid, or free oligosaccharide. In bold are the substituents that distinguish type A and B. Adapted from (Gagnon et al., 2017) with permission.

Despite extensive investigation into the ABO(H) blood types, their precise biological function is not fully understood. Many studies link human pathogens, including viruses, bacteria,

and fungi, to the human ABO(H) blood group system, where the presence of the A, B, or H antigens on the cellular surface can permit pathogen adhesion and access. This is well-studied in norovirus: one viral sub-group is primarily infectious in type O individuals, since it recognizes the unmodified H-antigen disaccharide. Here, individuals with A or B antigens effectively protect the epitope recognized by the virus through addition of GalNAc or Gal (**Figure 1**). There are similar mechanisms at work in *rotavirus* infection, *Helicobacter pylori*, *Campylobacter jejuni*, *Salmonella enterica*, and *Vibrio cholerae* bacterial infections (Yamamoto, F., et al. 2012, Cooling, L. 2015, Ewald, D.R. and Sumner, S.C. 2016, Heggelund, J.E., et al. 2017). There is also a link between the blood group antigens and certain types of cancer (Khalili, H., et al. 2011, Liumbruno, G.M. and Franchini, M. 2013, Franchini, M., et al. 2016, Huang, J.Y., et al. 2017, Meo, S.A., et al. 2017). Overall, based on the current literature it seems that blood group antigen diversity and polymorphisms could have arisen in response to a multitude of environmental threats in a kind of evolutionary guerrilla tactic that serves to curtail our population-level susceptibility.

1.3 The genetics and enzymology of human α-(1→3)-N-acetylgalactosaminyltransferase GTA and α-(1→3)-galactosyltransferase GTB

In 1910 von Dungern and Hirszfeld established that the A and B blood types are inherited codominantly, and both are dominant over type O, an important milestone in the history of the human ABO(H) blood group system comprehensively reviewed by Watkins *et al.* (2001). Another major development occurred in 1924 with Bernstein's three allele model, which maintains that *A, B,* and *O* are alternative alleles of a gene, such that children inherit one of three alleles from each parent (Lowe, J.B. 1993). This gives six genotypic possibilities, *AA*, *AB*, *AO*, *BB*, *BO*, and *OO*, and four resultant phenotypes, A, B, AB, and O.

Until the 1950s the ABH antigens were thought to be the direct products of the blood group genes. Due to advances in genetics, including the discovery that the translated products of genes were proteins, this notion was discarded – instead, the antigens must represent a secondary gene product. Scientists had identified the terminal blood group structures in part *via* sequential glycoside hydrolase removal of sugar monosaccharides, and so the reverse process might explain their biosynthesis. A new class of enzyme, glycosyltransferases (GTs), was thought to carry out this process. Individual GT enzymes were envisioned to catalyze the addition of specific monosaccharides from a "donor" substrate to an "acceptor" molecule, such that the product of one GT subsequently could act as the substrate for another GT.

Based on this reasoning, the A and B genes were thought to encode an α-N-acetyl-D-galactosaminyltransferase (GTA; EC number 2.4.1.40) and an α-D-galactosyltransferase (GTB; EC number 2.4.1.37), respectively. Experiments in the 1960s confirmed that type A individuals express GTA, type B express GTB, type AB express both, and type O express non-functional enzyme (Morgan, W.T.J. and Watkins, W.M. 1959, Watkins, W.M. 1972).

In 1976 the *ABO* gene was mapped to human chromosome 9 (9q34.2) and found to contain seven exons (Ferguson-Smith, M.A., et al. 1976). Remarkably, there are over 100 allelic *ABO* sequences. The *O* allele, which results in the type O phenotype, characterized by the presence of a terminal α-Fuc and the absence of α-Gal/α-GalNAc, encodes a non-functional GT. The most common *O* allele contains a single nucleotide deletion that shifts the reading frame to introduce a premature stop codon (Yamamoto, F., et al. 1990). The translated product of these genes lacks the GTA/GTB domain required for catalysis. Following the development of recombinant DNA technology, 1990 brought the first recombinant expression of human *GTA* and *GTB*. (Yamamoto, F., et al. 1990). Yamamoto *et al.* (1990) showed that transfecting O cells with

DNA reconstructed with genes encoding GTA and GTB resulted in A and B antigen expression, thus definitively establishing that the *ABO* genes were responsible for AB antigen synthesis (Yamamoto, F., et al. 1990, Yamamoto, F. and Hakomori, S. 1990).

Table 1. Four amino acid differences between GTA and GTB

Residue number	**GTA**	**GTB**
176	Arg	Gly
235	Gly	Ser
266	Leu	Met
268	Gly	Ala

Yamamoto *et al.* also identified the genetic differences between the A and B alleles using cDNA libraries constructed from the RNA of hosts with various blood types (Yamamoto, F., et al. 1990, Yamamoto, F. and Hakomori, S. 1990). They consistently observed four nucleotide substitutions that resulted in four amino acid differences between GTA and GTB (residues 176, 235, 266, and 268; **Table 1**). Following recombinant expression of human GTA and GTB, researchers developed assays to assess the enzymatic activity of GTA, GTB, and mutants thereof, and prioritized characterization of the role of these four 'critical' amino acids, discussed in **Chapter 1.5**.

1.4 Properties and classification of glycosyltransferases

Glycosyltransferases catalyze the formation of glycosidic linkages. The broadness of this definition highlights the diversity and biological importance of GTs. GTs synthesize carbohydrate-containing products, namely oligo- and polysaccharides, glycolipids, glycopeptides, and natural products. Their products, though often overlooked, are involved in many biological processes. Consider the glycosylation of IgG antibodies, whose role in immunity is poorly understood, or the T4 bacteriophage glucosyltransferase, which glucosylates

phage DNA conferring protection against host nucleases. Or consider yet still the work of GTs in the biosynthesis of several antibiotic classes, including the vancomycin and novobiocin classes (Walsh, C.T., et al. 2003). While this thesis examines the human blood group enzymes GTA and GTB, it will aim to draw conclusions or examine relationships within the broader class of GTs.

An informed overview of GTs must include discussion of their biosynthetic counterparts, the glycoside hydrolases. As outlined earlier, GHs served as an indispensable tool permitting the characterization of the blood group antigens. Further, the use of GHs to perform stepwise degradation of sugars made it possible for researchers to envision the reverse pathway. This trend persists in the history of these enzymes, where the findings concerning GHs frequently served as the impetus for hypotheses concerning GTs.

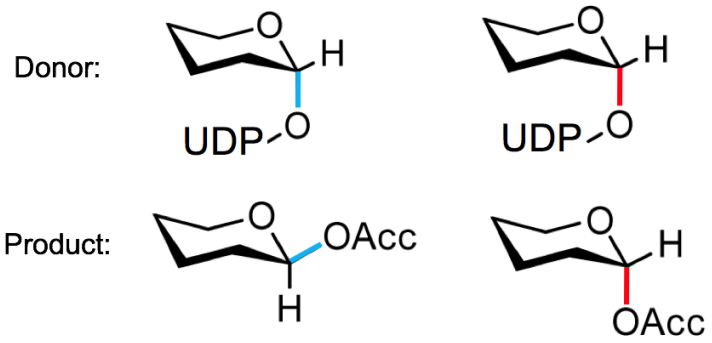
Glycosyltransfer is responsible for the three most abundant homobiopolymers – amylose, cellulose, and chitin (Weadge, J.T., et al. 2007). GTs catalyze what is fundamentally a nucleophilic substitution reaction, where a sugar is liberated from a donor substrate to form a new glyosidic linkage with an acceptor molecule (**Figure 2**). The substitution occurs at the donor sugar anomeric carbon, and it takes place either with retention or inversion of the original donor sugar stereochemical configuration (**Figure 2**). Generally, GTs are highly specific, and the one-enzyme-one-linkage dogma generally holds true (with some rare exceptions).

GTs are classified based on four main characteristics: substrate specificity/product(s) generated, amino acid sequence, protein tertiary structure/fold-type, and glycosyltransfer stereochemistry/mechanism.

For most GTs the donor substrate is a nucleoside diphosphate sugar (*e.g.* UDP-Gal, UDP-GalNAc, GDP-Man) but there are GTs that utilize phosphate- and lipid phosphate-linked sugars. GTs specific for nucleotide-sugar donors are termed “Leloir enzymes” after Luis Leloir, who

discovered the first sugar nucleotide (Leloir, L.F. 1983). GTs utilize a vast range of possible acceptor substrates in comparison to their somewhat limited donor repertoire. Acceptors can belong to any molecular class; GTs specific for nucleic acid (RNA, DNA), peptide, lipid, and small organic molecule acceptors have been identified (Rini, J., et al. 2009).

***Figure 2**. Inverting and retaining glycosyltransfer.*



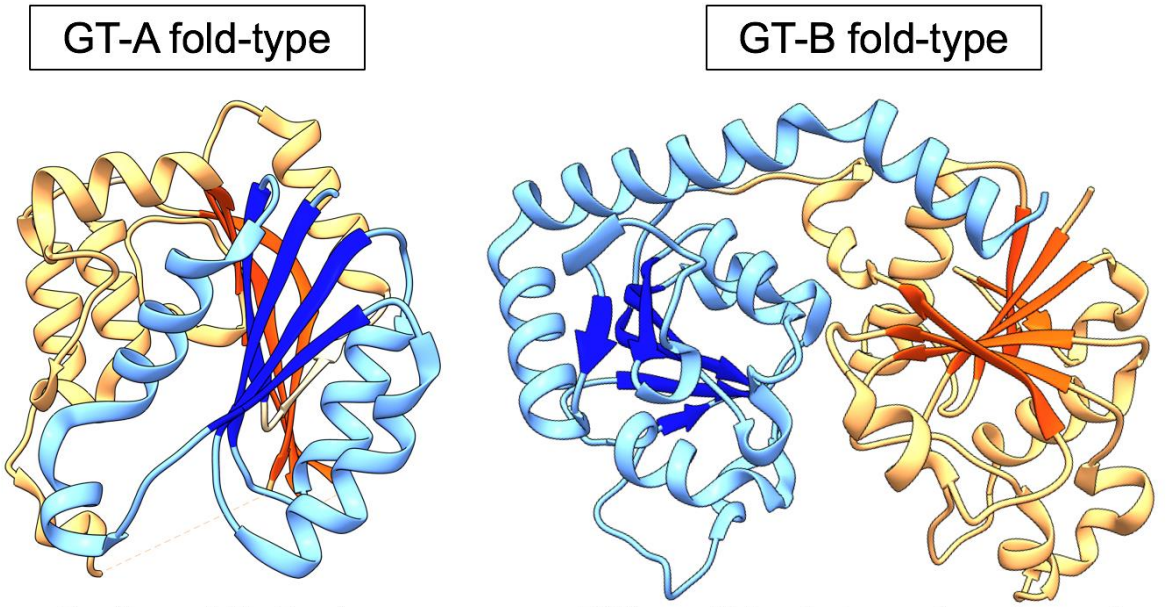
Inverting glycosyltransfer (*left*) alters the stereochemistry about the anomeric carbon of the transferred monosaccharide. This is seen by comparing the linkage of the product relative to the donor substrate (axial to equatorial or α to β linkage and *vice versa*). Retaining glycosyltransfer (*right*) preserves the stereochemistry (axial to axial or α to α, equatorial to equatorial or β to β linkage) of the transferred monosaccharide.

The Carbohydrate-Active enZYmes Database (CAZy; URL: http://www.cazy.org/) (Lombard, V., et al. 2014) originated in 1998 in an effort to compile genomic, structural, and biochemical data for viral, prokaryotic, and eukaryotic protein entities that act on or associate with glycosidic linkages, including GHs, GTs, polysaccharide lyases, carbohydrate esterases, and carbohydrate-binding modules. As of December 2018, CAZy groups GTs into over one hundred families based on amino acid sequence identity (Campbell, J.A., et al. 1997, Coutinho, P.M., et al. 2003). Notably, GT families can exhibit polyspecificity and so genomic or structural information has a limited ability to predict the biochemical activity of a putative transferase. The CAZy families are asymmetrically populated and characterized; while some families, such as

families 100-103, are recent additions and contain few entries, other families, such as family 2, contain a remarkable number of entries (107,871 at the time of writing), though many of these await biochemical verification. Yet still other families, such as family 6, contain comparatively few entries but are incredibly well represented in the Protein Data Bank.

Though GTs vary considerably in sequence, two main fold-types dominate: GT-A and GT-B (not to be confused with the human ABO(H) blood group glycosyltransferases GTA and GTB, which were named first). This is an interesting difference between GTs and GHs, which exhibit a broader variety of fold-types.

Figure 3. *Two main glycosyltransferase fold-types*



The GT-A fold type has two aligned β/α/β Rossmann fold-like domains, while the GT-B fold type has two flexibly-linked β/α/β Rossmann fold-like domains. In each case, one of the two domains is colored orange (light orange for α-helices, dark orange for β-sheets), while the other is colored blue (light blue for α-helices, dark blue for β-sheets).

The X-ray crystal structure of *Bacillus subtilis* SpsA was the first representative of the GT-A fold. The structure reveals two neighboring β/α/β Rossmann fold-like domains (**Figure 3**), which are so close together that they form a central β sheet. The GT-A fold-type also features two distinct nucleotide donor- and acceptor-binding domains in addition to a sequence-predicted Asp-X-Asp (DXD) motif. This latter region participates in metal cofactor or ribose coordination (Lairson, L.L., et al. 2008). The GT-B fold also consists of two β/α/β Rossmann fold-like domains, but in this case there is a flexible linker region between the two, and they are oriented such that they face each other **(Figure 3)**. Bacteriophage T4 β-glucosyltransferase was the first structurally characterized nucleotide-sugar-utilizing GT with the GT-B fold-type. Here, the enzyme active site is situated in the cleft between the Rossmann fold-like domains (Lairson, L.L., et al. 2008). Notably, bacteriophageT4 β-glucosyltransferase and other GT-B fold enzymes are typically metal ion-independent, unlike GTs with a GT-A fold. There is a third, less common fold-type: GT-C. Enzymes with this fold-type are usually hydrophobic integral membrane proteins, such as the *Campylobacter lari* oligosaccharyltransferase (PDB code 3RCE), that utilize lipid-phosphate linked sugar donors (Gloster, T.M. 2014).

Some GTs do not fit into any of these fold-types. For example, researchers have proposed a fourth fold, GT-D, to describe the "domain of unknown function" 1792 (DUF1792), a bacterial glucosyltransferase that catalyzes glucose transfer to Fap1, a *Streptococcus parasanguinis* bacterial adhesin (Zhang, H., et al. 2014). Zhang *et al.* superimposed the structure of DUF1792 with many GTs belonging to GT-A, GT-B, and GT-C fold-types and obtained RMS deviations above 3 Å, which they argue shows poor tertiary structure identity. However, DUF1792 possesses a Rossmann fold-like domain, a UDP/UDP-Glc binding site, and a manganese binding site, similar to other GTs. Unlike most other metal-binding sites in GTs, this site consists of a

DXE rather than a DXD motif, and this appears catalytically important, since mutations to DXD, EXD, and EXE reduce turnover significantly (Zhang, H., et al. 2014).

In addition to these, it is possible, even likely, that other fold-types exist, given that the number of GTs with known structures represents a small fraction of the total number of GTs in all life forms.

Stereochemical classification hinges on whether glycosyltransfer occurs with retention or inversion of the donor sugar anomeric configuration (Sinnott, M.L. 1990). Enzymes that conserve the stereochemistry of the donor sugar glycosidic linkage in the product (α to α, β to β) are retaining (ret-GTs), and enzymes that invert the linkage (α to β, β to α) are inverting (inv-GTs) (**Figure 2**). There is no correlation between fold-type and mechanism, as both ret-GTs and inv-GTs with GT-A and GT-B fold-types have been identified.

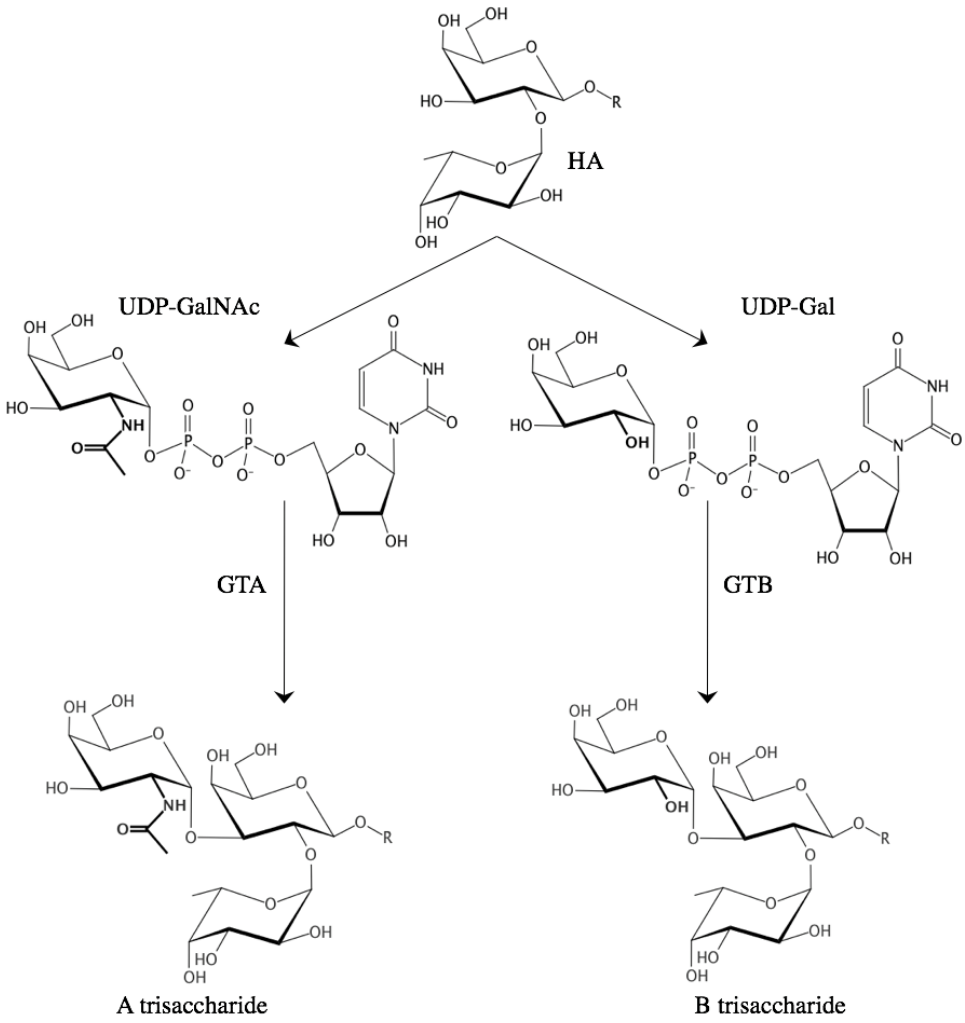
1.5 GTA and GTB

Human ABO(H) blood group enzymes α-(1→3)-N-acetylgalactosaminyltransferase (GTA; EC 2.4.1.40) and α-(1→3)-galactosyltransferase (GTB; 2.4.1.37), CAZy family 6 GTs with a GT-A fold-type, synthesize the A and B antigens (**Figure 4**). These antigens are found on the surface of not only red blood cells, but also on most endothelial and epithelial cells (Szulman, A.E. 1962, Yamamoto, F., et al. 1990, Yamamoto, F. and Hakomori, S. 1990, Hamasaki, N. and Yamamoto, M. 2000, Lombard, V., et al. 2014). These antigens are especially important in transfusion and transplantation medicine, where a blood type mismatch can have fatal consequences.

GTA and GTB are localized to the Golgi complex and are type II transmembrane proteins with a cytosolic N-terminus, a single-pass transmembrane domain, a short stem region, and a

globular, luminal catalytic domain. GTA transfers *N*-acetylgalactosamine (GalNAc) from UDP-GalNAc to the H antigen acceptor terminal disaccharide α-L-Fuc-(1→2)-β-D-Gal-*O*-R (HA) to generate the A antigen, and GTB transfers galactose (Gal) from UDP-Gal to the same acceptor disaccharide to generate the B antigen (**Figure 4**) (Yamamoto, F. and Hakomori, S. 1990). As a result of a frameshift or substitution mutation in the *ABO* gene locus, blood group O individuals express truncated or non-functional forms of these enzymes, in which case biosynthesis terminates at the H antigen (Lee, H.J., et al. 2005). Both GTA and GTB preserve the axial linkage of the donor sugar in the A and B antigen products and so are classified as ret-GTs.

Figure 4. *GTA & GTB glycosyltransfer schematic*



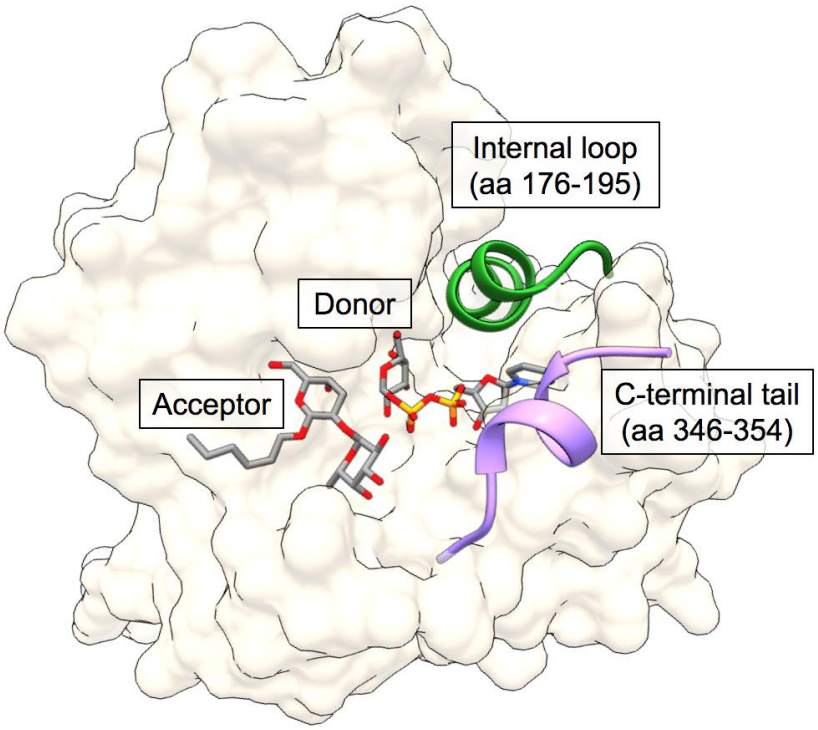
Reproduced from (Gagnon *et al*., 2017) with permission.

There are four acceptor substrate variants, in terms of their peripheral core structures, that give rise to four A and B antigen isotypes. In the 1980s, developments in nuclear magnetic resonance spectroscopy (NMR), mass spectrometry (MS), hybridoma technology, and monoclonal antibodies (mAbs) made identification of these isotypes possible (Clausen, H. and Hakomori, S. 1989). Variants of the H antigen, α-L-Fuc-(1→2)-β-D-Gal-*O*-R differ in the R group as follows: *type I*, α-L-Fuc-(1→2)-β-D-Gal-*O*-(1→3)-β-D-GlcNAc-*O*-R; *type II*, α-L-Fuc-

(1→2)-β-D-Gal-*O*-(1→4)-β-D-GlcNAc-*O*-R; *type III*, α-L-Fuc-(1→2)-β-D-Gal-(1→3)-α-D-GalNAc-*O*-R; *type IV*, α-L-Fuc-(1→2)-β-D-Gal*p*-(1→3)-β-D-GalNAc-*O*-R. In the case of type IV, the minimal H antigen acceptor is linked to a globoside moiety. In all cases, the H antigen terminus, and the moieties corresponding to type I-IV, can be linked to a glycoprotein or glycolipid molecule. The A and B trisaccharide antigens can be present as free oligosaccharides. The presence of these antigenic I-IV subtypes in the human body is not uniform, and their distribution patterns have been heavily reviewed elsewhere (Hakomori, S. 1981, Clausen, H. and Hakomori, S. 1989). Note that there additional A, B, and H antigen polymorphisms. While type I-IV isotypes arise due to core structure differences, other polymorphisms arise from antigen inner core branch processing, glycoconjugate identity (*ie*. glycolipid, N- or O-linked glycoprotein), and biosynthetic interactions with glycosyltransferases, typically belonging to other blood group systems, that act on the same substrate as GTA/GTB (Clausen, H. and Hakomori, S. 1989).

Stephen V. Evans' group published the first crystal structures of GTA (PDB code 1LZ0) and GTB (PDB code 1LZ7) in 2002, nearly the centenary of Karl Landsteiner's breakthrough discovery (Patenaude, S.I., et al. 2002). Selenomethionine-GTB was crystallized following soaks with mercury-containing compound 3-chloro-Hg-2-methoxy-propylurea, and the structure was solved *via* multiple-wavelength anomalous dispersion (MAD) (Patenaude, S.I., et al. 2002). The solution for GTB allowed determination of the GTA structure by molecular replacement. Unfortunately, the use of mercury soaks introduced significant disorder into two regions of polypeptide, preventing complete characterization of the entire enzyme at that time. Subsequent structures were solved using molecular replacement.

***Figure 5**. Mobile polypeptide loops of GTA/GTB*



The internal loop (aa 176-195; green helix) and C-terminal tail (aa 345-354; purple helix) are organized around the active site, where donor and acceptor substrates are bound. Substrates are colored by element with carbon gray, nitrogen blue, oxygen red, and phosphorous orange.

These enzymes have two mobile polypeptide loops, an internal loop (aa 176-195) and a C-terminal tail (aa 346-354), that recognize and sequester substrate during catalysis, a common feature among GTs (**Figure 5**) (Gastinel, L.N., et al. 2001, Persson, K., et al. 2001, Patenaude, S.I., et al. 2002, Qasba, P.K., et al. 2005, Yazer, M.H. and Palcic, M.M. 2005, Letts, J.A., et al. 2007, Alfaro, J.A., et al. 2008, Schuman, B., et al. 2010). Both loops are disordered in the unliganded, "open" state and, when fully liganded, transition to a highly ordered, catalytically competent "closed" state where they serve to occlude water from the active site to prevent non-specific donor hydrolysis. In the 'semi-closed' state donor (or UDP) and Mn^{2+}, but not acceptor, are bound to the enzyme. Here, the internal loop is predominantly ordered, but the C-terminal tail (which has been shown to be involved in acceptor recognition) usually remains disordered

(Patenaude, S.I., et al. 2002, Alfaro, J.A., et al. 2008). The "closed" state, where both the internal loop and C-terminal tail are ordered about the substrates, has been observed when both donor and non-reactive acceptor analogues are present in the active site.

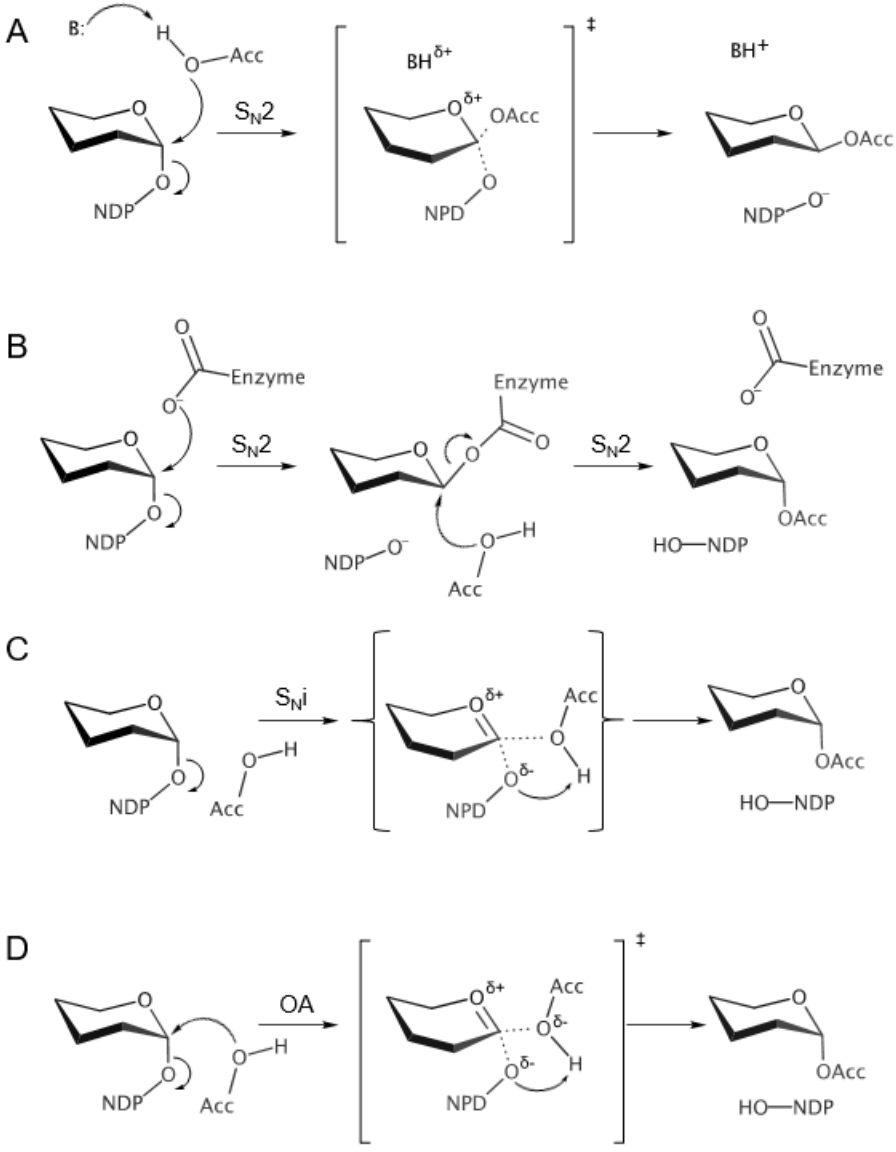
GTA and GTB possess a DXD motif that consists of Asp211, Val212, and Asp213. As with many other divalent metal cation-binding GTs, these aspartate residues help coordinate Mn^{2+}, which in turn stabilizes the negatively-charged pyrophosphate moiety of the UDP sugar donor substrate.

Remarkably, GTA/GTB differ in just four so-called "critical" amino acids out of 354 (Arg/Gly176, Gly/Ser235, Leu/Met266 and Gly/Ala268) and are therefore the most homologous naturally-occurring GTs that use distinct naturally-occurring donor substrates (**Table 1**). Two of these (Leu/Met266 and Gly/Ala268) are involved in donor specificity (Patenaude, S.I., et al. 2002), while Gly/Ser235 influences acceptor binding, and Arg/Gly176 affects internal polypeptide loop ordering (Letts, J.A., et al. 2006). Much of the work exploring the roles of these four critical amino acids exploited GTA/GTB chimeras. Four-letter codes delineate chimeric identities, where each letter in the code corresponds to one of the four critical residues. Using this notation, AAAA represents GTA, BBBB represents GTB, and AABB represents a chimera with the first two critical residues of GTA, Arg176 and Gly235, and the last two critical residues of GTB, Met266 and Ala268.

1.6 The unresolved catalytic mechanism of retaining glycosyltransferases

Glycosyltransferases have been found in all classes of living organisms, where they are responsible for the most abundant post-translational modification in nature, yet the mechanism by which about half of these enzymes function remains in dispute (Hurtado-Guerrero, R. and Davies, G.J. 2012, Schuman, B., et al. 2013, Ardèvol, A. and Rovira, C. 2015).

***Figure 6**. Mechanistic schemes proposed for inverting and retaining glycosyltransferases*



(**A**) S_N2 mechanism of inverting GTs. (**B**) Double displacement mechanism proposed for retaining GTs. (**C**) Front face S_Ni-like mechanism, which begins with leaving group dissociation (Lairson, L.L., et al.

2008). (**D**) Front face orthogonal associative (OA) mechanism, which proceeds in a single step (Schuman, B., et al. 2013).

While inverting glycosyltransfer is understood to proceed *via* acceptor S_N2 attack of the anomeric donor carbon (**Figure 6A**) in a manner analogous to the inverting GH enzymes, there are several candidate mechanisms for retaining glycosyltransfer, and it is unclear which of these is biologically in effect. The proposed mechanisms of retention include double displacement (**Figure 6B**) and variations of a front-face (also called front-side) S_Ni-like mechanism (**Figure 6C&D**), each requiring the presence of a nucleophilic group proximal to the anomeric donor sugar carbon.

Double displacement involves an initial nucleophilic attack of the donor sugar, performed by an enzyme nucleophile, to generate a covalent glycosyl-enzyme intermediate with inverted stereochemistry (**Figure 6B**). In a second displacement reaction, an acceptor nucleophile attacks the intermediate, again inverting stereochemistry, resulting in net retention of anomeric stereochemistry in the saccharide product (Soya, N., et al. 2011). The internal return S_Ni-like mechanism, hereafter described as "S_Ni", has dissociative character and consists of the following events: UDP departure, which yields an oxocarbenium ion-like transition state, and subsequent nucleophilic attack of the sugar by the acceptor nucleophile (**Figure 6C**) (Sinnott, M.L. 1990). There is an offshoot of S_Ni, known as the "orthogonal associative" mechanism, that involves simultaneous UDP loss and nucleophilic attack by the acceptor at a right angle to the C1-leaving group axis (**Figure 6D**) (Schuman, B., et al. 2013). Though its delineation may seem unnecessarily fastidious, in the analysis of Schuman *et al.* (2013) the orthogonal associative variant best explains the observed data for retaining GTs. For example, Schuman *et al.* stipulate that the measured distances between the acceptor nucleophile and the anomeric carbon of the

donor sugar would not permit the dissociation required for S_Ni, at least for LgtC, the *Neisseria meningitides* α-(1→4)-galactosyltransferase, and GTA/GTB (Schuman, B., et al. 2013).

The observation that family 6 GTs GTA/GTB, β-galactoside α-(1→3)-galactosyltransferase (α3GalT; EC 2.4.1.87), and the *Bacteroides ovatus* α-N-acetylgalactosaminyltransferase (BoGT6a; EC 2.4.1.40) each possesses a putative active site nucleophile (Glu303, Glu317, and Glu192 respectively) fostered early speculation of a double displacement mechanism (Patenaude, S.I., et al. 2002). Evidence for a glycosyl-enzyme intermediate has accumulated over the past two decades of study: the detection of a covalent adduct with E303C mutants of GTA/GTB by Soya *et al*. (2011), the chemical rescue of an inactive Glu317Ala mutant of α3GalT (Monegal, A. and Planas, A. 2006), and the crystal structures of a BoGT6a Glu192Gln mutant in which hydrolyzed GalNAc may be linked covalently to Gln192 (Pham, T., et al. 2014), though in this case the resolution (3.42 Å) is too poor to allow for a definitive evaluation. For α3GalT, static hybrid quantum mechanics and molecular mechanics (QM/MM) studies have shown that nucleophile-assisted front side attack and double displacement can occur at equal rates (Gomez, H., et al. 2012). Subsequent hybrid QM/MM and molecular dynamics (MD) simulations provided evidence that retaining glycosyltransfer *via* the putative nucleophile (*e.g.* Glu303 and Glu317 in the case of GTA/GTB and α3GalT, respectively) may only proceed through a double displacement mechanism with a dissociatively-formed covalent intermediate (Rojas-Cervellera, V., et al. 2013).

It is unclear which of the proposed mechanisms, front-face attack (both the S_Ni dissociative and orthogonal associative variants) and double displacement, is most relevant to *in vivo* catalysis. Computational studies have shown that multiple mechanisms are energetically feasible, and the assignment of one all-encompassing mechanism may be unreasonable.

Additional static hybrid QM/MM calculations have indicated that both S_Ni and double displacement mechanisms are feasible for wild-type GTB (Bobovska, A., et al. 2014, Bobovská, A., et al. 2015)

Ardèvol and Rovira (2015) have summarized the available QM/MM studies of GHs and GTs, including the somewhat ambiguous results obtained for the retaining GTs. The energetic similarity of these mechanisms and the evidence accumulated to date raise the possibility that subtle active site perturbations, such as those imposed methodologically, have mechanistic consequences and may even determine which mechanism, out of all energetically accessible options, is in effect (Hurtado-Guerrero, R. and Davies, G.J. 2012, Ardèvol, A. and Rovira, C. 2015).

Chapter 2: Primary aims & rationale

Despite their profound biological importance, many aspects of GTs remain unknown, including questions of reaction mechanism, substrate-binding and conformational changes, and product release. Given that some GT saccharide products are linked to human disease processes (outlined in **Chapter 1**), among them genetic disorders, bacterial infection, and cancer, there is an advantage to improving our fundamental understanding of how GTs operate. Structural and kinetic studies of model GT enzymes can provide insight into these areas and can inform the design of specific inhibitors against bacterial GTs. As well, since GTs are stereo- and regio-selective, knowledge of their precise functions can permit their use in chemoenzymatic synthesis, circumventing the problem of multiple protection and deprotection steps, which are tedious and generate waste.

The human ABO(H) blood group A and B enzymes GTA and GTB are well-characterized and can be readily mutated and produced in high milligram quantities *via* recombinant bacterial expression. These qualities make GTA and GTB excellent model enzymes for studying glycosyltransfer, and thus they are the subject of the work reported in this dissertation.

One area of uncertainty for GTA/GTB and other retaining GTs is the mechanism of glycosyltransfer, as discussed in **Chapter 1**. The core question is whether retaining GTs use a double displacement mechanism, which invokes a covalent acyl-enzyme intermediate, or a variant of front-face attack, which proceeds in a single step and does not generate an enzyme-linked intermediate. For GTA and GTB, Glu303 is the candidate catalytic nucleophile that becomes linked to the donor sugar during double displacement. For GTA/GTB Glu303Cys

mutants there has been MS detection of the double displacement intermediate species (Soya, N., et al. 2011). However, there has been no direct detection of the intermediate species.

Glu303 is not only significant for its role as a putative nucleophile, but also because it makes key stabilizing hydrogen bonds with both donor and acceptor substrates, as indicated in earlier structural studies (Alfaro, J.A., et al. 2008). Thus, Glu303 may be a critical amino acid for both catalysis and substrate binding. **Aim 1** of this dissertation is to advance understanding of the retaining glycosyltransfer mechanism by examining the role of Glu303 using a site-directed mutagenesis strategy coupled with kinetic and X-ray crystallography structural analysis. My approach and progress toward this aim, as well as additional relevant background information, are outlined in **Chapter 4**.

Another area of uncertainty for GTs is substrate binding. Among GTs with published structures, relatively few have been collected in complex with intact substrate. This has been a challenge for GTA and GTB in particular, and there are several factors that may explain these difficulties, including substrate hydrolysis, oxidation, and conformational changes in both the enzyme and substrate, all of which can shatter crystal lattices (Boix, E., et al. 2001, Angulo, J., et al. 2006, Blume, A., et al. 2006, Alfaro, J.A., et al. 2008). Without liganded crystal structures, our understanding of substrate binding and enzyme conformational changes is lacking. We cannot know precisely how the substrate interacts with the enzyme prior to catalysis, nor can we assess whether the transferring monosaccharide of the donor sugar undergoes conformational changes required for turnover.

Aim 2 of this dissertation is to determine how donor substrate binds and examine how it achieves the constrained, high-energy conformation associated with catalysis. The approach and progress toward this aim, as well as specific background information are described in **Chapters**

5 and **6**. **Chapter 5** focuses on a series of X-ray crystal structures of GTA, GTB, and their chimera in complex with three different donors, UDP-Gal, UDP-Glc, and UDP-C-Gal. The third donor is a non-hydrolysable substrate analogue, whose use presumably could circumvent the issue of substrate-hydrolysis, one of the major impediments to crystallization. In **Chapter 6** the focus is on mutants of GTA/GTB Arg188 and Asp302, residues known to interact with the donor sugar based on the data reported in **Chapter 5** and past structural studies (Alfaro, J.A., et al. 2008, Gagnon, S.M. et al. 2015).

Though GTA/GTB were first characterized structurally in 2002 (Patenaude, S.I., et al. 2002), nearly two decades prior to this current work, little is known about the product release mechanism. It is unclear how the A/B trisaccharide product and UDP leaving group departure is coordinated. More broadly, there is scarcity of GT structures collected in complex with their reaction products. Any forthcoming insights from such structures, including those of model enzymes GTA and GTB, could inform the development of product analogues for use as GT inhibitors.

Aim 3, the final aim of this dissertation, is to characterize trisaccharide product formation and release from the active site following glycosyltransfer. The relevant background information, experimental approach, and progress are outlined in **Chapter 7**.

Through a thorough structural and kinetic examination of each of these facets of glycosyltransfer, residue 303 and the catalytic mechanism (**Chapter 4**), substrate binding and conformational shifts (**Chapters 5&6**), and product departure (**Chapter 7**), *via* model enzymes GTA/GTB, we gain a more complete understanding of these biologically crucial class of enzymes. The advances discussed in this dissertation will spur future work on related GTs and will inform their use in practical applications.

Chapter 3: Experimental approach

3.1 Generation of mutants, protein expression, purification

The -10 GTA/GTB chimera genes (amino acids 63-354; AABB, ABBA, ABBB chimera) were made by PCR amplification using the wild-type human ABBB and AABB genes as templates. The forward primer 5′-ATA TGA ATT CAT GGT TTC CCT GCC GCG TAT GGT TTA CCC GCA GCC GAA-3′ (MIN2) introduced an EcoRI site in the 5′ end, and the reverse primer 5′-ATA ATT AAG CTT CTA TCA CGG GTT ACG AAC AGC CTG GTG GTT TTT-3′ (PCR-3B) introduced a HindIII site in the 3′ end. The PCR profile used was 94 °C/3 min (94 °C, 30 s, 55 °C, 30 s, and 72 °C, 1 min) for 30 cycles. After gel purification, the PCR products were digested with EcoRI and HindIII for 2 h at 37 °C and were ligated into pCWΔ*lac*, which had been opened with EcoRI/HindIII. Each ligation was transformed into *Escherichia coli* BL21 chemically ($CaCl_2$) competent cells. The DNA sequences were confirmed on both strands.

Previously described methods (Marcus, S.L., et al. 2003) were used to construct all mutant enzymes by recombinant PCR except the E303Q mutants. These were produced (QuikChange) using GTA/GTB (residues 63-354) plasmid DNA as a template. The first PCR was performed using the outside forward primer MIN2 and the internal reverse primer for the mutant, which contains a single codon substitution. A second PCR was performed PCR3B and the internal forward primer, which contains the codon substitution corresponding to the desired mutation. The two overlapping fragments were isolated, annealed by PCR mediated 3' extension, and amplified using the 5’ and 3’ primers MIN2 and PCR3B. The resulting mutated DNA fragment was digested, and ligated into pCWΔ*lac* as above.

For the E303Q mutants, the forward primer was 5’-GCT GTT TGG CAC GAC **CAG** TCC CAC CTG AAC AAA TAC-3’, and the reverse primer was 5’-GTA TTT GTT CAG GTG

GGA **CTG** GTC GTG CCA AAC AGC-3. For E303A, the forward primer was 5'-CAC GAC **GCT** TCC CAC CTG AAC AAA TAC CTG CTG-3', and the reverse was 5'-CAG GTG GGA **AGC** GTC GTG CCA AAC AGC TTC GAT AC-3'. For E303C, the forward primer was 5'-TGG CAC GAC **TGC** TCC CAC CTG AAC AAA TAC CTG-3', and the reverse was 5'-CAG GTG GGA **GCA** GTC GTG CCA AAC AGC TTC GAT AC-3'. For E303D the forward primer was 5'-TGG CAC GAC **GAC** TCC CAC CTG AAC AAA TAC CTG-3', and the reverse was 5'-CAG GTG GGA **GTC** GTC GTG CCA AAC AGC TTC GAT AC-3.

For the D302A mutants the forward primer was 5'-TGG CAC **GCT** GAA TCC CAC CTG AAC AAA TAC CTG-3', and the reverse was 5'-GTG GGA TTC **AGC** GTG CCA AAC AGC TTC GAT ACC-3'. For the D302C mutants the forward primer was 5'-GTT TGG CAC **TGC** GAA TCC CAC CTG AAC AAA TAC-3', and the reverse was 5'-GTG GGA TTC **GCA** GTG CCA AAC AGC TTC GAT ACC-3'. For the D302L mutant the forward primer was 5'-TGG CAC **CTG** GAA TCC CAC CTG AAC AAA TAC CTG-3', and the reverse primer was 5'-GTG GGA TTC **CAG** GTG CCA AAC AGC TTC GAT ACC-3'. For the D302E mutants the forward primer was 5'-TGG CAC **GAG** GAA TCC CAC CTG AAC AAA TAC CTG-3', and the reverse primer was 5'-GTG GGA TTC **CTC** GTG CCA AAC AGC TTC GAT ACC-3'.

For the R188K mutants the forward primer was 5'-G CAG GAC GTT TCC ATG CGT **AAA** ATG GAA ATG ATC AGC GAC-3' and the reverse primer was 5'-GTC GCT GAT CAT TTC CAT **TTT** ACG CAT GGA AAC GTC CTG CC-3'.

All insert and plasmid purifications were performed with the Qiagen plasmid purification system (Qiagen, Chatsworth, CA). All ligations were made using T4 DNA ligase (Invitrogen) at room temperature for 1 h. All restriction enzymes were purchased from New England Biolabs.

Recombinant wild-type and mutant GTA and GTB were expressed in *E. coli* BL21 cells as previously described (Seto, N.O.L., et al. 1999, Marcus, S.L., et al. 2003). Purification was performed following the published two-step procedure, using an SP-Sepharose cation exchange column followed by a UDP-hexanolamine affinity column as described previously

3.2 Kinetics

Through a long-standing collaboration with Prof. Stephen Evans, Prof. Monica Palcic's research group conducted all kinetic assays using the following parameters.

Kinetic characterization of the mutants was carried out with a radiochemical assay using a hydrophobic acceptor α-D-Fuc-(1→2)-β-D-Gal-*O*-$(CH_2)_7CH_3$ and radiolabelled sugar donors (Palcic, M.M., et al. 1988, Alfaro, J.A., et al. 2008). In this assay radiolabelled reaction products are isolated from unreacted donor by adsorption onto reverse-phase C18 cartridges.

3.2.1 Kinetics of 303 mutants of GTA/GTB (Chapter 4)

All kinetics for these mutants were performed through collaboration with Prof. Monica Palcic's research groups, and the methods are described in Blackler *et al.* (2017).

3.2.2 Kinetics of 302 and 188 mutants of GTA/GTB (Chapter 6)

All kinetics for these mutants were performed through collaboration with Prof. Monica Palcic's research groups, and the methods are described in Gagnon *et al.* (2018).

3.3 Crystallization

All crystallization was performed by Svetlana N. Borisova except where spontaneous crystals formed in stock solutions as indicated in the relevant subsections below. The final pH for all crystallization conditions described below was pH 7-7.5.

3.3.1 Structures of 303 mutants of GTA/GTB (Chapter 4)

All of the mutant enzymes were crystallized in the presence of mercury using conditions identical to the wild-type GTA and GTB enzymes decribed in (Patenaude, S.I., et al. 2002). Only two mutant enzymes could be crystallized in the absence of mercury ions: GTA/E303D and GTB/E303C, and both were crystallized under conditions as previously reported (Alfaro, J.A., et al. 2008).

The atomic coordinates and structure factors (5CMF, 5CMG, 5CMH, 5CMI, 5CQL, 5CQM, 5CQO, 5CQP, 5CQN) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

3.3.2 Structures of GTA, GTB, and chimera in complex with different donors and acceptors (Chapter 5)

UDP-C-Gal, L-Fuc-(1→2)-β-D-Gal*p*-O$(CH_2)_7CH_3$ (HA) and α-L-Fuc-(1→2)-β-D-(3-deoxy)-Gal*p*-O$(CH_2)_7CH_3$ (DI) were synthesized as previously reported (Lowary, T.L. and Hindsgaul, O. 1993, Kamath, V.P., et al. 1999, Partha, S.K., et al. 2010). Donor substrates UDP-Gal and UDP-Glc were purchased from Sigma-Aldrich.

Crystallization—Chimeric AABB, ABBB and ABBA proteins were crystallized in protein stock solution as described by Alfaro *et al.* (Alfaro, J.A., et al. 2008). Native GTA/GTB proteins were

grown at 4°C from a much higher concentration of protein (30-40 $mgml^{-1}$ for GTB and 16-20 $mgml^{-1}$ for GTA) along with 1% polyethylene glycol (PEG) 4000, 4.5–5% 2-methyl-2,4-pentanediol (MPD), 100 mM ammonium sulfate, 70 mM sodium chloride, 50 mM N-[2-acetamido]-2-iminodiacetic acid (ADA) buffer pH 7.5, 30 mM sodium acetate buffer pH 4.6 and 5 mM manganese chloride ($MnCl_2$) for GTB crystallization and 5-8 mM $CoCl_2$ for GTA crystallization. Drops of 10-15 µl were placed against a reservoir containing 3.7% PEG 4000, 7% MPD, 0.3 M ammonium sulfate, 0.25 M sodium chloride, 0.2 M ADA buffer and 0.1 M sodium acetate. Crystals grew for 5-10 days at 4 ° C.

Protein crystals were washed with artificial mother liquor (ML)-2 containing 3.5% PEG 4000, 50 mM ammonium sulfate, 40 mM sodium chloride, 35 mM ADA buffer and 15% MPD prior to substrate addition. Crystals of AAAA, AABB, ABBB, ABBA and BBBB in complex with UDP-Gal, UDP-C-Gal, UDP-Glc, DI and HA were obtained by soaking them in mother liquor ML-2 with 15% MPD, 60-70 mM UDP-Gal, 60-80 mM UDP-Glc, 40-60 mM UDP-C-Gal, 12-20 mM DI, 20 mM HA, and 10 mM $MnCl_2$ for 2-5 days at 4°C. Short soaking times (24-48hrs) and reduced concentrations of UDP-C-Gal and DI were also used for ABBA and ABBB chimeras. Donors, acceptors, and analogues thereof were added incrementally over a period of a few minutes to hours to avoid crystal fracture. The data sets and structures presented were from those crystals which diffracted to highest resolution. Before freezing crystals for data collection, the cryoprotectant concentration was adjusted to 20% MPD.

The atomic coordinates and structure factors (5C36, 5C38, 5C3A, 5C3B, 5C3D, 5C47, 5C48, 5C49, 5BXC, 5C1G, 5C1H, 5C1L, 5C4B, 5C4F, 5C8R, 5C4C, 5C4D, 5C4E have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

3.3.3 Structures of 302 and 188 mutants of GTA/GTB (Chapter 6)

GTA, GTB, their chimera and mutants generally crystallize more readily as mercury derivatives, but this has been shown to disorder the internal polypeptide loop (Letts, J.A., et al. 2007, Alfaro, J.A., et al. 2008, Schuman, B., et al. 2010). To explore the structural effects of Arg188 and Asp302 mutations the enzymes were crystallized in the absence of mercury. Diffraction-quality crystals were generated for GTB/R188K and four of the Asp302 mutant enzymes: GTA/D302C, GTB/D302A, GTB/D302C, GTB/D302L.

Spontaneous crystals were recovered from concentrated stocks of GTB/D302A (68 mg/mL), GTB/D302L (40-50 mg/mL), GTB/D302C (55 mg/mL), and GTA/D302C (95-100 mg/mL) stored in 50 mM MOPS pH 7.00, 0.1M NaCl, 1mM DTT, 5 mM $MnCl_2$ and kept at 4^{o}C. GTB/R188K crystals were obtained by hanging drop vapor diffusion at 4^{o}C, where 3μL of concentrated stock (66-68 mg/mL) was mixed with 1 μL of reservoir solution containing 0.3 M sodium acetate and 0.3 M NaCl.

Prior to freezing, crystals were washed in mother liquor initial (ML_o) containing 6.8% PEG 4000, 40 mM sodium acetate pH 4.6, 30 mM ADA pH 7.5, 20 mM MES pH 6.5, 40 mM ammonium sulfate, 9 mM $MnCl_2$, 30% glycerol.

The atomic coordinates and structure factors have been deposited in the Protein Data Bank (6BJI, 6BJJ, 6BJK, 6BJL, 6BJM), Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

3.3.4 Structures of GTA/GTB in complex with trisaccharide products (Chapter 7)

The first crystals of GTA/GTB were grown as mercury derivatives (Patenaude, S.I., et al. 2002). Crystals were washed with artificial mother liquor ML-1 consisting of 10% PEG 4000, 30 mM ADA buffer pH 7.5, 30 mM sodium acetate buffer pH 4.6, 100 mM ammonium sulfate, 10 mM $MnCl_2$ and 30% glycerol as the cryoprotectant. Crystals of GTA/GTB in complex with the A and B trisaccharides were obtained by soaking them in mother liquor ML-1 with 30% glycerol and 45-50 mM substrates for 2-5 days at 18°C.

Native crystals of GTA/GTB lacking any heavy metals were grown at 4°C from much higher concentrations of protein (30-40 mg/mL for GTB and 16-20 mg/mL for GTA) along with 1% PEG 4000, 4.5-5% MPD, 100 mM ammonium sulfate, 70 mM sodium chloride, 50 mM ADA buffer pH 7.5, 30 mM sodium acetate buffer pH 4.6 and 5 mM $MnCl_2$ for GTB crystallization and 5-8 mM $CoCl_2$ for GTA crystallization. 10-15 µL drops were placed against a reservoir containing 3.7% PEG 4000, 7% MPD, 0.3 M ammonium sulfate, 0.25 M sodium chloride, 0.2 M ADA buffer and 0.1 M sodium acetate. The crystals were usually grown for 5-10 days at 4°C.

Before making complexes, crystals of GTA/GTB were washed with modified mother liquor ML-2 consisting of 3.5% PEG 4000, 50 mM ammonium sulfate, 40 mM sodium chloride, 35 mM ADA buffer and 15% MPD or glycerol. Crystals of native GTA and GTB in complex with the A and B trisaccharide, α-L-Fuc*p*-(1→2)[α-D-GalNAc*p*-(1→3)]-β-D-Gal*p*-O$(CH_2)_7CH_3$ and α-L-Fuc*p*-(1→2)[α-D-Gal*p*-(1→3)]-β-D-Gal*p*-O$(CH_2)_7CH_3$, respectively, were obtained by soaking them in mother liquor ML-2 with 15% glycerol or MPD and 45-50 mM substrates for 2 - 5 days at 4°C. Before freezing the crystals for data collection, the concentration of the cryoprotectant was made 30% glycerol or 20% MPD respectively.

The atomic coordinates and structure factors (codes 3IE6, 3IE8, 3IDK, 3IDL, 3IDR and 3IDT) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

3.4 Data collection & structure determination

For all structures, data were collected on a Rigaku RAXIS IV++ area detector at a distance of 72 mm and exposure times between 4.0 and 5.0 minutes for 0.5° oscillations, and processed with d*trek (Pflugrath, J.W. 1999). X-rays were produced by a MM-002 generator (Rigaku/MSC) coupled to Osmic "Blue" confocal x-ray mirrors with power levels of 30 watts. The crystals were frozen and maintained under cryogenic conditions at a temperature of -160 °C using a CryoStream 700 crystal cooler (Oxford). All structures were solved using molecular replacement techniques (MOLREP) (Vagin, A. and Teplyakov, A. 1997) with wild-type GTA (PDB code 1LZ0) or GTB (PDB code 1LZ7) as a starting model. Data refinement was performed using the REFMAC5 module (Murshudov, G.N., et al. 1997) in the CCP4 program suite (Winn, M.D., et al. 2011). Model and density visualization and editing was done using the program *Coot* (Emsley, P., et al. 2010). Data and refinement statistics for all structures are included in each relevant chapter.

3.5 Saturation transfer difference (STD) nuclear magnetic resonance (NMR) spectroscopy

In collaboration with Profs. Stephen V. Evans and Monica Palcic, Prof. Thomas Peters' group at the University of Lübeck Institut für Chemie und Metabolomics conducted all experiments involving STD-NMR using a protocol described in Gagnon *et al.* (2017).

3.6 Figure generation

Figures 3, 5, 7, 8, 9, 10, 16A-B, 19, and 22G-H were generated using UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311) (Pettersen, E.F., et al. 2004). Figures 12, 14, 15, 16C, 18A-B, 20, 21, 22A-F, and 24 were generated using SetoMac, an unpublished development of SETOR (Evans, S.V. 1993). Figures 1, 2, 4, 6A-D, 11, 13 and, 18C-E were produced with Marvin Sketch, a ChemAxon program (http://www.chemaxon.com).

3.7 Permissions

Much of the information provided in the experimental approach, results, and discussion sections is adapted from authored publications. Where indicated, tables and figures are reproduced or modified versions of those that appear in the relevant publications (Gagnon, S.M., et al. 2015, Blackler, R.J., et al. 2017, Gagnon, S.M.L., et al. 2017, Gagnon, S.M.L., et al. 2018). For this, the necessary permissions have been secured from *Glycobiology* and *The Journal of Biological Chemistry* and are included comprehensively in **Appendices I-IV**.

3.8 Collaborations

Data that arose solely *via* the efforts of collaborators is indicated in the relevant sections above and includes: all kinetic experiments on GTA/GTB and their mutants and chimera, all STD NMR experiments on GTA/GTB with their product trisaccharides.

Chapter 4: Role of Glu303 in substrate binding and catalysis

4.1 Introduction to Chapter 4

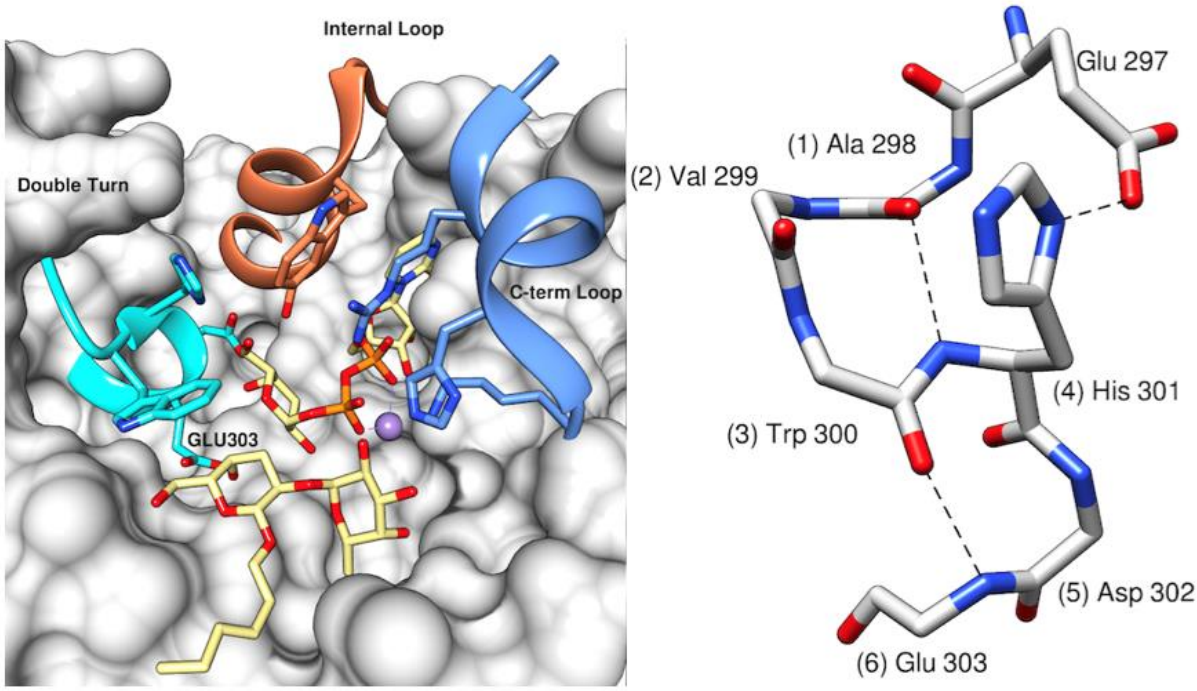
This chapter addresses the first aim of this dissertation: to advance understanding of the retaining glycosyltransfer mechanism. Model retaining GTs, human ABO(H) blood group A and B enzymes GTA and GTB, synthesize the A and B blood group antigens. These enzymes are ideal probes for mechanistic studies, since they have been well characterized and can be readily mutated and produced in high milligram quantities. An intermediate consistent with the double displacement mechanism has been detected in mass spectrometry experiments of GTA and GTB Glu303Cys mutants (Soya, N., et al. 2011). Analysis has shown the enzyme-glucosyl intermediate to be stable on the order of days, which provides an excellent opportunity for crystallization and X-ray diffraction experiments. Further, studies examining the energetics of glycosyltransfer in wild-type and E303C mutants *via* quantum-mechanics and molecular-mechanics methods have provided additional evidence for the stability of the acyl-enzyme intermediate (Bobovská, Tvaroška, and Kóňa 2015).

Despite extensive theoretical and MS support, there has been no direct detection of the intermediate species in the wild-type or mutant GTA/GTB. A crystal structure would provide more convincing mechanistic proof and would allow determination of the conformation of the covalently linked sugar. Further, we would be able to view the architecture of the enzyme itself in the intermediate stage of glycosyltransfer. There is a strong interest in the rational design of specific GT inhibitors, a feat that has proved challenging, since theoretically these could target essential bacterial GTs while leaving mammalian GTs unimpeded. Despite ongoing efforts, few effective GT inhibitors have emerged, and there is a growing demand for such compounds given the problem of antibiotic resistance (Compain, P. and Martin, O.R. 2001, Compain, P. and

Martin, O.R. 2003, Bella, M., et al. 2015). Knowledge of the intermediate structures of GTA and GTB bound to their cognate sugars could inform the design of such molecules.

Glu303 is not only significant for its role as a putative nucleophile, but it is also important in substrate binding, as indicated in earlier structural studies (Alfaro, J.A., et al. 2008). In structures determined in complex with natural substrate, Glu303 is observed to form hydrogen bond interactions with both donor and acceptor hydroxyl groups. In the latter case, Glu303 interacts with the Gal-4-OH, which has been identified as a "key polar group" in that it is critical for catalysis in GTA/GTB, an interesting binding feature that these enzymes share with CAZy family 6 bovine α-1,3-galactosyltransferase (α3GalT) (Lowary, T.L. and Hindsgaul, O. 1993, Sujino, K., et al. 1997, Mukherjee, A., et al. 2000).

Figure 7*. GTA/GTB double turn motif*



Left: GTA/GTB active site with the internal loop, C-terminal tail and the double-turn motif displayed as ribbons and the remainder of the enzyme surface with UDP-Gal donor and DI acceptor bound (PDB code 2RJ7). *Right*: Close up of the residues comprising the double-turn motif in GTA/GTB (PDB code 2RJ7). Dashed lines represent hydrogen bond interactions. Adapted from (Blackler *et al*., 2017) with permission.

Glu303 also resides within a six amino acid residue "double-turn" motif, which consists of two interpenetrating hairpin turns (**Figure 7, right**): the N-terminal Type I hairpin turn (Ala298, Val299, Trp300, and His301) and the C-terminal Type II′ hairpin turn (Trp300, His301, Asp302, and Glu303) (Creighton, T.E. 1993). The double-turn is structurally strained – in the Type I turn, His301 adopts an uncommon, traditionally disallowed P_{II}' conformation (Hollingsworth, S.A. and Karplus, P.A. 2010). This is reminiscent of the nucleophile elbow motif of α/β hydrolases (Ollis, D.L., et al. 1992), though in this second case the catalytic nucleophile itself is strained, unlike GTA/GTB, which prevents direct comparison of these double-turns.

To identify the function of Glu303 in catalysis, substrate binding, and double-turn stability, we produced E303C, E303D, E303Q and E303A mutants of GTA and GTB for kinetic and structural analysis.

Toward the published data reported in **Chapter 4**, my contributions include refining X-ray diffraction data, depositing the structures, analysis and interpretation structural and kinetic data, and writing and editing a significant portion of the paper.

4.2 Results

Kinetics and structural data for 303 mutants of GTA and GTB

Data collection and refinement statistics are shown in **Table 2** for GTA mutants and **Table 3** for GTB mutants, while **Table 4** summarizes the kinetic results obtained by Professor Palcic's group for all mutant enzymes. As mercury derivatives, GTA, GTB, and their chimeras and mutants generally crystallize more readily (Patenaude et al. 2002; Alfaro et al. 2008), and so four mutants of GTA and four of GTB were produced and crystallized in the presence of mercury (denoted with asterisks): GTA/E303C*, GTA/E303D*, GTA/E303Q*, GTA/E303A*,

GTB/E303C*, GTB/E303D*, GTB/E303Q*, and GTB/E303A*. Only GTA/E303D and GTB/E303C also crystallized in their native form.

Other than the differences described below, the mutant enzyme structures do not deviate from those of wild-type GTA/GTB (*e.g.* PDB codes 1LZ0/1LZ7). Compared to the wild-type enzymes, GTA and GTB mutants had elevated temperature factors and greater or complete disorder in residues 297-302 (**Table 2**). For all mutants, the C-terminus (residues 346-354) was disordered, consistent with previous native and Hg-derivatized structures of unliganded wild-type GTA/GTB (Patenaude et al. 2002; Letts et al. 2007; Alfaro et al. 2008). The C-terminal tail is labile and in structures of GTA/GTB can remain disordered, whether the enzyme is unliganded or with donor substrate and manganese bound in the active site. Generally, the C-terminus only becomes fully ordered upon addition of the acceptor substrate with which it makes several key hydrogen bond interactions as observed by Alfaro *et al.* (2008).

In all mercury derivatized structures, the internal loop (residues 176-195) is always observed to be disordered, consistent with earlier studies of these enzymes (Patenaude et al. 2002; Letts et al. 2007; Alfaro et al. 2008). In native GTA/E303D the internal loop displayed weak electron density and was in an open conformation distinct from that observed in earlier work (Alfaro et al. 2008). For native GTB/E303C, there is electron density for residues 188-195, while residues 176-187 are disordered – there is slightly greater disorder than in the structure of native wild-type GTB, where amino acid residues 177-183 were disordered (PDB code 2RIT) (Patenaude et al. 2002; Alfaro et al. 2008).

Table 2. Data collection and refinement results for GTA/Glu303 mutants

	GTA/E303A*	**GTA/E303C***	**GTA/E303D***	**GTA/E303Q***	**GTA/E303D**
Resolution (Å)	20.00 – 1.95	20.00 – 1.83	20.00 – 1.61	20.00 – 1.73	20.00 – 1.85
Data collection wavelength (Å)	1.5418	1.5418	1.5418	1.5418	1.5418
Space Group	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$
a (Å)	52.41	52.56	52.77	52.70	52.51
b (Å)	149.00	148.87	148.83	148.91	149.57
c (Å)	78.97	80.01	79.34	79.59	78.25
R_{merge} (%)[a]	5.6 (30.9)	4.0 (30.8)	4.2 (27.0)	5.0 (31.8)	6.0 (33.6)
Completeness (%)[a]	94.7 (92.9)	99.3 (100.0)	98.9 (95.7)	99.3 (98.0)	97.4 (99.8)
Unique Reflections	21,791	27,916	40,426	32,876	26,085
Multiplicity	3.86	4.85	4.27	4.57	4.62
Mean I/σ(I)	10.4	16.2	17.8	13.8	11.5
Wilson B	33.6	33.7	24.9	26.8	27.4
Refinement					
Reflections used	20666	26509	38400	31214	24734
R_{free} reflections	1113	1407	2026	1662	1310
R_{work} (%)	21.8	17.7	18.7	17.3	18.1
R_{free} (%)	25.3	20.9	21.0	20.0	23.3
No. water	118	126	191	178	216
No. non-H atoms ligands	4 (Hg)	4 (Hg)	4 (Hg)	5 (Hg)	0
No. non-H atoms protein	2128	2130	2131	2200	2283
r.m.s. bond (Å)	0.017	0.019	0.023	0.021	0.019
r.m.s. angle (°)	1.8	1.8	2.3	2.1	1.9
Ramachandran outliers (%)	0	0	0	0.4	0.4
Clash score	2	2	4	6	5
Mean B-factor (Å^2)	41.0	38.0	26.0	30.0	30.0
PDB code	5CMF	5CMG	5CMH	5CMJ	5CMI

[a] Values in parentheses represent highest resolution shell.

* Asterisks indicate mercury derivative structures.

Reproduced from (Blackler et al. 2017) with permission.

Table 3. Data collection and refinement results for GTB/Glu303 mutants

	GTB/E303A*	**GTB/E303C***	**GTB/E303D***	**GTB/E303Q***	**GTB/E303C**
Resolution (Å)	20.00 – 1.69	20.00 – 1.65	20.00 – 1.69	20.00 – 1.83	20.00 – 1.61
Data collection wavelength (Å)	1.5418	1.5418	1.5418	1.5418	1.5418
Space Group	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$
a (Å)	52.77	52.67	52.87	52.72	52.57
b (Å)	149.01	149.53	149.55	149.22	149.82
c (Å)	79.30	79.26	79.39	79.82	78.24
R_{merge} (%)[a]	4.2 (28.2)	4.4 (30.9)	4.1 (23.2)	5.8 (30.5)	3.7 (31.8)
Completeness (%)[a]	99.3 (98.0)	98.1 (94.3)	96.4 (93.1)	99.1 (100.0)	96.4 (97.4)
Unique Reflections	39,257	37,346	34,359	27,957	38,967
Multiplicity	4.55	4.61	5.54	4.46	4.50
Mean I/σ(I)	16.5	14.9	19.3	12.0	18.5
Wilson B	26.6	24.5	25.7	29.3	23.4
Refinement					
Reflections used	32209	35473	32631	26545	37011
R_{free} reflections	1704	1871	1728	1412	1952
R_{work} (%)	17.4	17.3	17.7	17.7	18.4
R_{free} (%)	20.6	20.6	20.8	21.7	20.6
No. water	126	200	185	150	190
No. non-H atoms ligands	5 (Hg)	5 (Hg)	4 (Hg)	4 (Hg)	0
No. non-H atoms protein	2133	2150	2126	2165	2245
r.m.s. bond (Å)	0.020	0.021	0.021	0.018	0.024
r.m.s. angle (°)	1.9	2.0	2.1	1.9	2.2
Ramachandran outliers (%)	0	0.4	0	0.4	0
Clash score	2	3	2	2	4
Mean B-factor (Å^2)	32.0	27.0	28.0	33.0	25.0
PDB code	5CQL	5CQM	5CQO	5CQP	5CQN

Table 4. Michaelis-Menten constants, estimated modelled distances from residue 303 to select atoms, and electron density of the double-turn in GTA and GTB E303 mutants

Enzyme	Michaelis-Menten constants			Modelled distances[d] from 303 side chain to:							Electron density of double-turn[f]
				Donor			Acceptor O4[e]		Tyr 264	Gly 267	
	K_A (μM)[a]	K_B (μM)[a]	k_{cat} (s^{-1})	C1	O4	O5	A	B			295 300
AABB (2RJ7)[d]				4.1	4.1	4.2	2.7	3.2	2.65	2.78	**GIEAV WHDES**
GTA* (1LZ0)	9.9[b]	8.7	17.5[g]	4.1	4.0	4.1	2.8	3.3	2.56	2.89	**GIEAV WHDES**
GTA/E303C*	400 ± 80	51 ± 5	0.9 ± 0.01	4.9	2.7	4.4		5.3	6.36	3.18	**GI**E-- --D**CS**
GTA/E303D	ND[c]	86 ± 9	2.2 ± 0.4	4.2	2.1	4.0		4.0	5.26	2.89	**GIE**a**V WH**d**DS**
GTA/E303D*				4.2	2.5	4.0		4.2	5.73	2.89	**GI**E-- ---**DS**
GTA/E303Q*	270 ± 70	50 ± 7	0.027 ± 0.001								**GIEAV WHDQS**
GTA/E303A*	ND[c]	30 ± 10	0.0047 ± 0.0005								**GI**e-- -hd**AS**
GTB (2RIT)	88	27	5.1	4.1	4.0	4.1	2.9	3.4	2.65	2.81	**GIEAV WHDES**
GTB* (1LZ7)				4.2	4.0	4.1	2.9	3.3	2.67	2.79	**GIEAV WHDES**
GTB/E303C	190 ± 60	36 ± 6	1.3 ± 0.09	4.5	3.0	4.0	4.6	4.4	3.47	3.19	**GIEAV W**HD**CS**
GTB/E303C*				4.5	3.3	4.0		4.1	4.87	3.49	**GI**E-- whd**CS**
GTB/E303D*	2640 ± 280	211 ± 22	0.058 ± 0.015	4.7	3.0	4.3		3.1	5.73	2.89	**GI**--- ---**DS**
GTB/E303Q*	NA[c]	NA[c]	NA[c]	4.2	4.1	4.3	2.7	3.3	2.97	2.89	**GI**EA**V WH**D**QS**
GTB/E303A*	980 ± 100	74 ± 20	0.0004 ± 0.00004								**G**IE-v WHd**AS**

[a] The K_A and K_B Michaelis-Menten constants are the K_m for acceptor and donor, respectively
[b] Wild-type data are from (Lee et al. 2005)
[c] ND indicates not determined and NA indicates no activity
[d] Modelled distances are estimates calculated from an overlay with the structure of AABB + UDP-Gal + DA (PDB code 2RJ7).
[e] Distances A and B are given for bidentate interactions.
[f] One letter amino acid codes in bold capitals correspond to electron density for main chain and side chain atoms, capital letters correspond to electron density for main chain atoms only, and lower case letters correspond to weak electron density for main chain and/or side chain atoms. Residues replaced with hyphens are observed to be disordered and have not been included in the refined models.
[g] Measured k_{cat} corresponds to native GTA enzyme
[*] Asterisks indicate mercury derivative structure

Structures and kinetics of E303C and E303D mutants

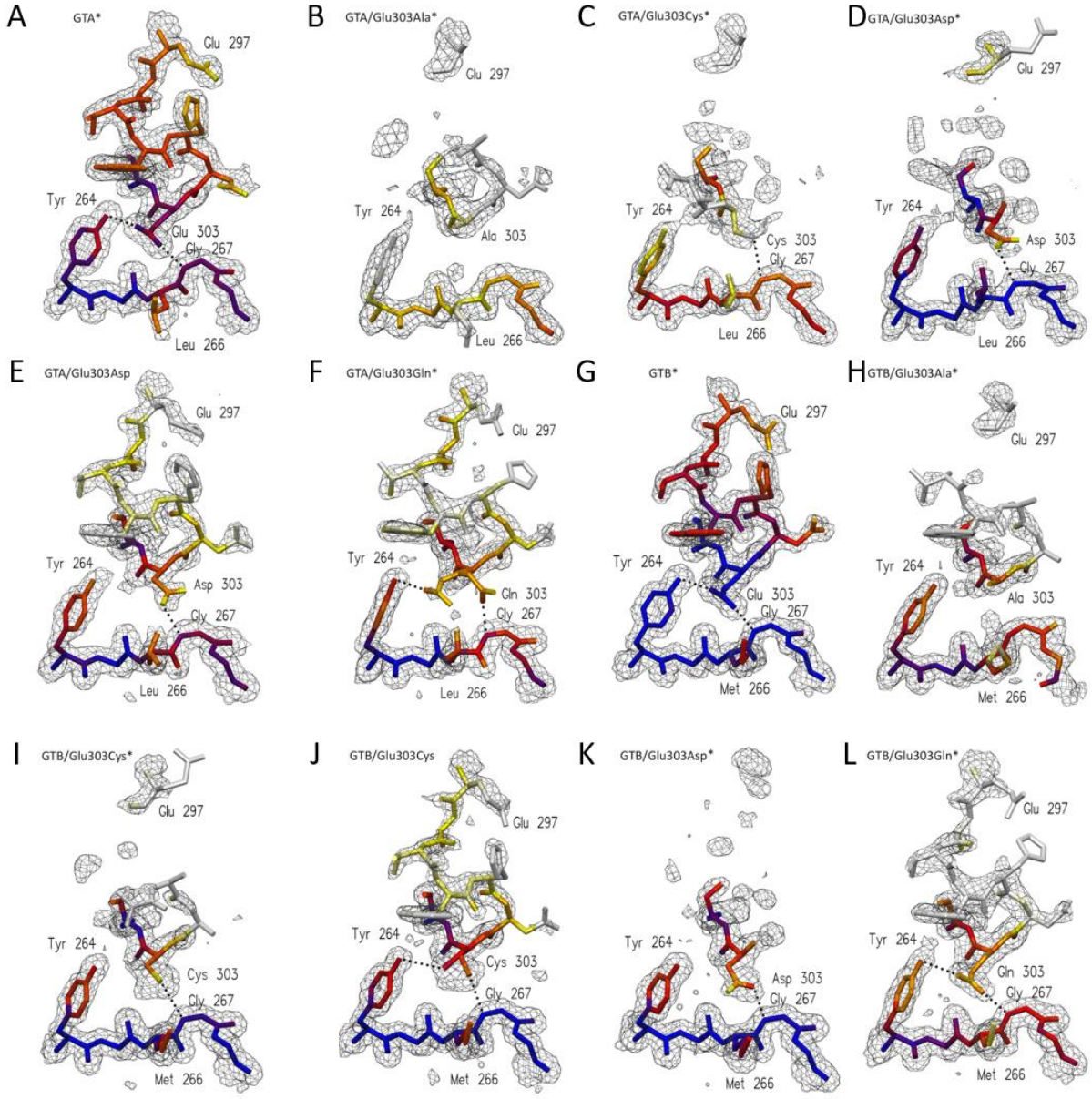
E303C mutation of GTA and GTB reduced the activity (turnover, k_{cat}) to 5% and 25%, respectively, of wild-type (**Table 4**). Again relative to wild-type, acceptor K_m is 40-fold higher for GTA/E303C but only 2-fold higher for GTB/E303C, and donor K_m is 6-fold higher for GTA/E303C and essentially unchanged for GTB/E303C.

The structure of GTA/E303C* residues 298-301 are almost entirely disordered (**Figure 8C**). Here, Asp302 is also observed in a different conformation than in the native wild-type enzyme, where its main chain carbonyl makes a hydrogen bond interaction with the Tyr264 side chain hydroxyl. Unambiguous density is observed for Cys303 oriented towards the Gly267 amine. GTB/E303C* residues 298-301 are disordered, and there is broadened density for Cys303, which is oriented toward Gly267 (**Figure 8I**).

The native GTB/E303C structure (**Figure 8J**) displays traceable electron density for the entire double-turn, but at lower levels and with higher temperature factors than wild-type. Cys303 has ambiguous electron density and is modeled in two conformations such that it is oriented towards either Tyr264 OH or Gly267 N.

For all mutants, models of liganded complexes, which allowed estimation of distances between residue 303 and select substrate atoms, were achieved by overlaying the unliganded structures reported here with AABB (PDB code 2RJ7), whose structure was collected in complex with UDP-Gal and DI acceptor (**Figure 9**). These distances are given in **Table 4**. For GTB/E303C, two Cys303 conformations are modelled, and the measured distances are given for the closest of the two conformations (**Table 4**).

Figure 8*. Double turn motif ordering in wild-type GTA/GTB and mutant*



(**A**) GTA*, (**B**) GTA/E303A*, (**C**) GTA/E303C*, (**D**) GTA/E303D*, (**E**) GTA/E303D, (**F**) GTA/E303Q*, (**G**) GTB*, (**H**) GTB/E303A*, (**I**) GTB/E303C*, (**J**) GTB/E303C, (**K**) GTB/E303D* and (**L**) GTB/E303Q*. Atoms are colored according to B-factor gradient with blue (<20 Å^2), red (20-30 Å^2), orange (30-40 Å^2), yellow (40-50 Å^2) and white (>50 Å^2). $2F_o$-F_c electron density map is contoured at 1.00 σ and cleaned around the displayed atoms or, in the absence of modelled atoms, the area corresponding to the double-turn. Asterisks (*) indicate mercury derivative structures. GTA* and GTB* are from PDBs 1LZ0 and 1LZ7, respectively. Reproduced from (Blackler et al., 2017) with permission.

For both GTA and GTB, and especially for the latter, E303D mutants had reduced activity. GTA/E303D and GTB/E303D had k_{cat} values of 13% and 1% that of wild-type GTA and GTB, respectively. Donor K_m values were elevated 10-fold for GTA/E303D and 8-fold for GTB/E303D, and acceptor K_m values also were elevated, 30-fold for GTB/E303D and 10-fold for GTA/E303D (**Table 4**).

In the GTA/E303D* structure, there is nearly complete disorder of residues 298-302 and moderate density for Asp303, which forms a hydrogen bond with the Gly267 amine (**Figure 8D**). The GTB/E303D* structure displays complete disorder of residues 297-302, and the Asp303 side chain forms a hydrogen bond to the Gly267 amine (**Figure 8K**). In GTA/E303D the double-turn has good electron density, though there is greater thermal motion relative to wild-type GTB – note that a comparison to wild-type GTB is necessary since there are no published structures of native, unliganded wild-type GTA. As we would expect, for GTA/E303D there is less disorder relative to the corresponding mutant derivative structures (**Figure 8E**). Here, the GTA/E303D Asp303 also forms a hydrogen bond contact with the Gly267 amine. The estimated distances from Asp303 to modelled substrate atoms in the mutant structures are given in **Table 4**.

Structures and kinetics of E303Q mutants

GTB/E303Q had no measurable activity, and GTA/E303Q had 0.15% of wild-type activity (**Table 4**). For GTA/E303Q, donor and acceptor *K*m values were elevated, consistent with the observations for the other mutants.

Both GTA/E303Q* and GTB/E303Q* structures showed traceable main chain and limited side chain electron density for the double-turn (**Figure 8F&L**). As with the other mutants, the double-turn had greater thermal motion relative to wild-type (**Figure 8F&L**).

Estimated distances from GTB/E303Q* Gln303 to the modelled substrates are given in **Table 4**. For GTA/E303Q*, Gln303 was too disordered to allow for an estimation.

Structures and kinetics of E303A mutants

The E303A mutation significantly reduced enzyme activity for both GTA and GTB (**Table 4**). In standard assays, the specific activities of GTA/E303A and GTB/E303A were barely measurable at 0.3 mU/mg and 0.06-0.6 mU/mg, respectively. This is substantially lower than wild-type GTA (10-15 U/mg) and GTB (4-7 U/mg). GTA/E303A and GTB/E303A had minimal activity: the k_{cat} for GTA/E303A dropped from 17.5 s^{-1} in wild-type GTA to 0.0047 s^{-1}, and the k_{cat} for GTB/E303A dropped from 5.1 s^{-1} in wild-type GTB to 0.0004 s^{-1}. Acceptor K_m increased by an order of magnitude from 88 μM in wild-type GTB to 980 μM in GTB/E303A. The UDP-GalNAc donor K_m increased from 27 μM in wild-type GTB to 74 μM in GTB/E303A, while that of GTA/E303A increased from 8.7 μM to 57 μM.

In GTA/E303A* and GTB/E303A* structures, amino acid 297-301 display significant disorder (**Figure 8B&H**), and so distances between residue 303 and modelled substrates could not be estimated.

4.3 Discussion & conclusions

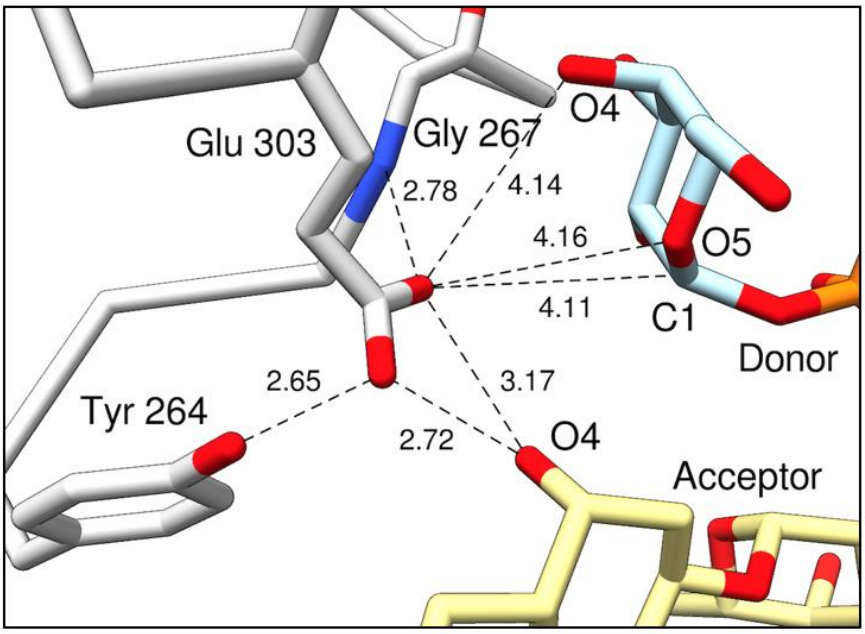
Hg derivatisation magnifies the disordering structural effect of 303 mutation

Only GTA/E303D and GTB/E303C crystallized under native conditions, while the remaining enzymes crystallized as mercury derivatives. In the native mutant structures the double-turn displays higher thermal motion than the native wild-type GTB structure, while for every mercury derivatized non-conservative mutant, the double-turn is completely disordered. In previous studies, the internal loop of native wild-type and mutant structures display a range of ordering (Alfaro et al. 2008; Johal et al. 2012; Johal et al. 2014), while for mercury derivatives it is completely disordered (Patenaude et al. 2002; Persson et al. 2007). The Glu303 mutant derivative structures follow this trend and display an increase in double-turn disorder relative to derivative wild-type greater than that observed for native mutants relative to native wild-type. This indicates that the internal loop plays a role in stabilizing the double-turn and supports the structural importance of Glu303 in the active site (**Table 4, Figure 8**).

Glu303 stabilizes the active site & influences substrate binding

In wild-type enzyme, several hydrogen bonds among active site residues stabilize the strained double-turn conformation. These include the Glu303 carboxyl's hydrogen bond interactions with the Tyr264 hydroxyl and Gly267 amide (**Figure 8 & 9**). In every published wild-type GTA/GTB structure, whether unliganded or liganded, Glu303 is observed to in the same conformation with the same hydrogen bond interactions, an observation that highlights the contribution of this residue toward maintaining active site architecture.

***Figure 9**. Distances between residue AABB Glu303 and select substrate atoms*



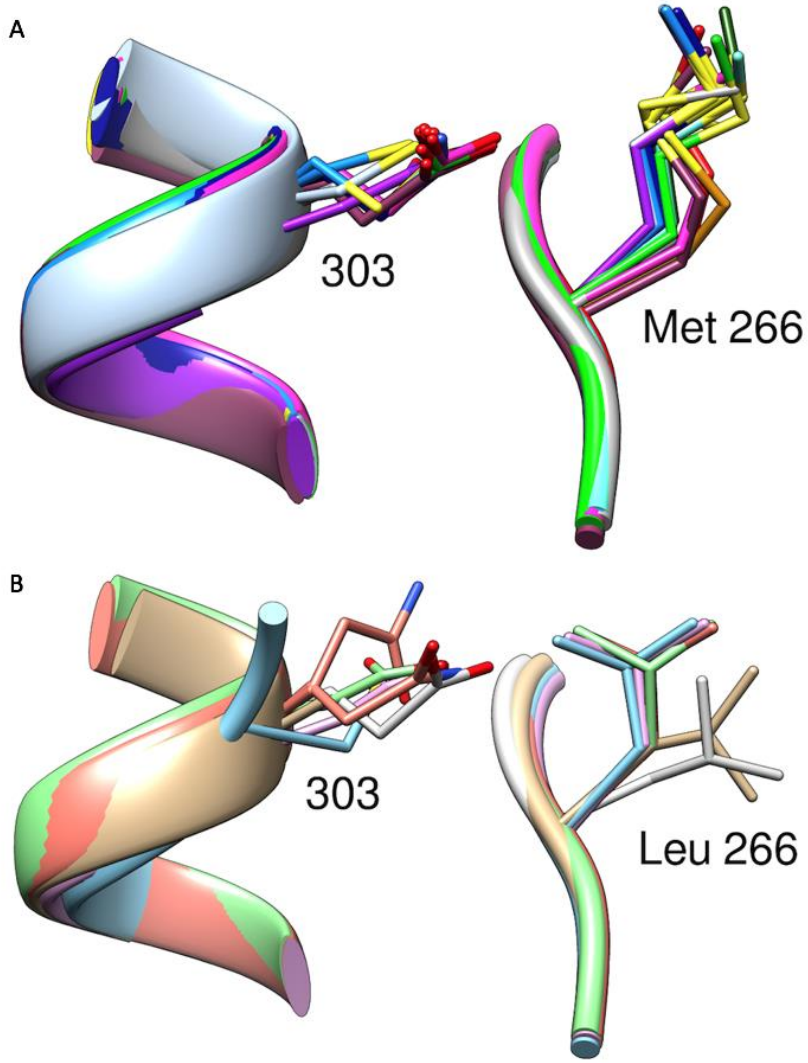
Distances from Glu303 of the chimeric enzyme AABB (PDB code 2RJ7) to select atoms (also provided in Table II). Interactions (with measured distances in Ångstroms) are indicated by dashed black lines. Adapted from (Blackler et al., 2017) with permission.

Glu303 mutants exhibited a range of decreased double-turn stability. For E303Q* mutants there was slightly increased thermal motion, while for GTA/E303A*, GTA/E303C* and both E303D* mutants the double-turn was completely disordered (**Table 4**, **Figure 8**). Even the structure of the conservative E303Q mutant exhibited elevated double-turn temperature factors and side chain disorder. The loss of Glu303's anchoring hydrogen bonds, which stabilize the strained double-turn conformation, likely explains these observations, which confirm the importance of Glu303's specific hydrogen bond character in active site stabilization.

Not only is Glu303 involved in double-turn stability, but it also influences the orientation of residue 266 (GTA/GTB: Leu/Met266), one of the 4 critical amino acids involved in donor substrate recognition. Prior to this study, in all published wild-type GTA structures Leu266 lies in the same position. For GTB, in contrast, Met266 lies in one of two conformations. In the 303

mutant structures described in this dissertation, Leu266 is observed in several unique conformations, while Met266 does not deviate from previous conformations (**Figure 10**).

Figure 10. *Conformations of residues 266 and 303*



Observed conformations of residue 303 and **(A)** Met266 or **(B)** Leu266. Unique conformations of Leu266 are seen for GTA/E303A* (tan), GTA/E303C* (blue), GTA/E303D* (pink), GTA/E303D (green), and GTA/E303Q* (salmon), whereas every structure with wild-type Glu303 shows electron density for Leu266 in the same position (white, PDB code 5BXC). In contrast, many structures of GTB and chimera with wild-type Glu303 show Met266 in various conformations with two general locations of the sulfur atom (example PDBs are colored as follows: 5C3B bright red, 5C3D dark green, 5C4D cyan, 4FRA yellow, 4FRB orange, 4FRQ bright green, 4GBP bright pink and 4KXO dark pink), and the GTB Glu303 mutants all show similar conformations (E303A* grey, E303C* light blue, E303C medium blue, Glul303Asp* purple and E303Q* dark blue). Asterisks (*) indicate mercury derivative structures. Reproduced from (Blackler et al., 2017) with permission.

Presumably, the double-turn disorder of the 303 mutants ought to impact substrate binding, given that residues within this region interact with donor and acceptor in the wild-type enzymes: there are stacking interactions between Trp300 and acceptor, and there are hydrogen bonds between His301 and donor 6-OH, Asp302 and 4-OH, and Glu303 and acceptor "key polar group" Gal-4-OH, which is critical for catalysis in GTA/GTB (Lemieux 1989; Lemieux 1996; Alfaro et al. 2008).

Additionally, another structural study (**Chapter 5**) revealed that GTA/GTB donor substrate binding is a stepwise process involving multiple donor conformations (Gagnon et al. 2015). In one intermediate donor conformation, there are hydrogen bond interactions between the donor sugar moiety and Asp302 and Glu303. In this conformation, Glu303 maintains its aforementioned two hydrogen bonds to the acceptor. Since Glu303 and the double-turn appear intimately involved in substrate recognition, it is unsurprising that the substrate K_m values of Glu303 mutants are higher than those of the wild-type (**Table 4**).

303 mutation impacts enzyme activity more than active site order

Despite disrupted substrate recognition sites, several mutants show significant activity (**Table 4**). There is a remarkable contrast between GTB/E303Q and the E303C/E303D mutants: in the former case, the active site retains its organization but there is no activity, while in the latter case, the active site is more disordered but the enzymes retain significant activity. For activity, a nucleophilic 303 residue is more critical for catalysis than active site organization.

The distances between the amino acid 303 side chain functional groups and the anomeric carbon C1 of modelled UDP-Gal donor (**Table 4**) are relatively consistent notwithstanding the varying lengths of the mutant R-groups. For the AABB complex structure (PDB code 2RJ7) this distance is 4.1 Å. Given the short length of the Cys303 side-chain compared to wild-type

Glu303, the E303C Sγ atoms might be expected to be several Ångstrom more distant from the donor sugar but are only 4.9 and 4.5 Å for GTA and GTB, respectively. Similarly, for Asp303 whose side chain also is shorter than Glu303, these distances are 4.2 and 4.7 Å.

Glu303 mutation had a significantly different effect on the catalytic constant of GTA *vs.* GTB (**Table 4**). This was surprising given the enzymes' overall homology. Mutation of Glu303 to Cys had a greater effect on GTA than GTB (5 *vs* 25% of wild-type k_{cat}; **Table 4**), while mutation of Glu303 to Asp showed the opposite trend (13 *vs* 1% of wild-type k_{cat}; **Table 4**).

Mechanistic implications

The differential effects of Glu303 mutation in GTA and GTB emphasize the need for discretion in making mechanistic deductions from such experiments or mechanistic comparisons among related retaining GTs. The residual activity of Glu303 mutants with different side chain chemistry, geometry, and disorder speaks to the mechanistic elasticity of GTA/GTB. These findings evince the prudence of Hurtado-Guerrero and Davies's caveat: elastic systems such as these, which may operate within a mechanistic continuum, could be guided toward a particular mechanism *via* the experimental conditions under which they are studied (Hurtado-Guerrero and Davies 2012). One potent example of such mechanistic sensitivity may be the MS detection of the covalent glycosyl-enzyme intermediate (CGEI) of GTA/GTB E303C mutants (Soya et al. 2011), while in contrast the wild-type GTA/GTB CGEI has not yet been detected. These observations are consistent with QM/MM calculations, which have demonstrated a CGEI that is remarkably stable in GTB/E303C but highly unstable in wild-type GTB (Bobovská, Tvaroška, and Kóňa 2015).

For all proposed mechanisms, whether double displacement, S_Ni-like, or orthogonal

associative, Glu303 has been assigned a role. As shown with Glu303 mutants, the generation of a covalent glycosyl-enzyme intermediate, which double displacement invokes, would disrupt the hydrogen-bonding network of the strained double-turn and thus destabilize the active site. Since some disordered 303 mutants can remain functional, it is not obvious how this increased lability influences catalysis. It is unclear how the structural change required for CGEI formation would occur, however it is evident that double displacement would necessitate a destabilized active site.

Most literature references of catalysis-associated structural strain address the enzyme-assisted transition state stabilization and reactant destabilization, however there are examples where strain within the enzyme itself is freed during catalysis, lowering the transition state activation energy barrier (Kagamiyama and Hayashi 2001; Gao 2003). Current GTA/GTB complex structures represent snapshots of the enzymes before or after catalysis, thus it is uncertain whether active site strain release contributes to the glycosyltransfer reaction. CGEI formation and coincident Glu303 hydrogen bonding changes also may impact active site opening and product release, though it is unclear how.

Other glycosyltransferases

Similar to GTA/GTB, homologous family 6 α-(1,3)-galactosyltransferase α3GalT has a strained double-turn motif, where the conserved Glu317 resides in a position that corresponds to Glu303 and also makes hydrogen bonds to acceptor O4 (PDB code 1GWW) (Boix et al. 2001; Gastinel et al. 2001; Boix et al. 2002; Zhang et al. 2003; Monegal and Planas 2006; Jamaluddin et al. 2007; Letts et al. 2007; Molina, Knegtel, and Macher 2007; Tumbale et al. 2008). The E317Q mutant of this enzyme displays low enzyme activity, and its crystal structures reveal no change in the overall structure or electron density of the active site (PDB codes 1O7O, 2VS5)

(Zhang et al. 2003; Monegal and Planas 2006; Tumbale et al. 2008), indicating that the loss of activity in this case is not due to active site deformation. The fact that activity is not lost despite active site disorganization is consistent with the observed behavior of GTA and GTB mutants.

Family 8 GT LgtC also possesses a strained double-turn motif (PDB code 1G9R, residues 184-189), where Gln189 is in the Glu303 position. Surprisingly, these Q189E mutants formed covalent glycosyl-enzyme intermediates at Asp190 (Lairson et al. 2004) ~9Å away from the anomeric carbon of the UDP-2F-Gal donor in the wild-type complex structure (PDB code 1GA8). Release of the double-turn upon mutation may explain this observation, however, NMR studies of the Q189E mutant did not identify additional dynamic properties that would account for such a conformational change, at least in the time ranges accessible to NMR (Chan et al. 2013).

The family 6 GT BoGT6a possesses a strained double-turn when in complex with 2'-fucosyllactose (PDB code 4AYJ). In the absence of substrate (PDB code 4AYL) the double-turn adopts an extended, unstrained conformation (PDB code 4AYL) (Thiyagarajan et al. 2012). For BoGT6a, Glu192 is in the same position as GTA/GTB Glu303. When in complex with GalNAc, the BoGT6a/E192Q mutant double-turn displays conformational variations. Only some of these are strained, each such conformation with Gln192 and bound GlcNAc in distinct positions (PDB code 4CJB) (Pham et al. 2014). The BoGT6a/E192Q mutant has near-wild-type donor and acceptor K_m values but significantly decreased k_{cat}. It is possible that since the double-turn of BoGT6a appears more mobile than that of GTA/GTB, the perturbation of active site structure by E192Q mutation has a lesser effect on substrate binding.

Conclusion

In GTA/GTB, the putative catalytic nucleophile Glu303 stabilizes the double-turn motif

in a strained conformation *via* several hydrogen bonds interactions. While even the most conservative Glu303 mutation destabilizes the active site, the effect on activity is much more dependent on mutant R-group nucleophilicity. Notably, active site geometry positions mutant functional groups at similar distances from modeled donor despite differences in R-group length. Further, these distances correlate with the relative activity of the mutant enzymes. In family 6 GTs α3GalT, LgtC, BoGT6a and GTA/GTB the structural and kinetic result of mutating residue 303 and its equivalents are inconsistent, underscoring the fine-tuned and complex nature of substrate binding and catalysis in these enzymes. Last, the residual activities of GTA/GTB mutants demonstrate catalytic elasticity, and it is plausible that these enzymes are sensitive to their precise environments, which may guide them to display one of multiple closely related, energetically-accessible mechanisms.

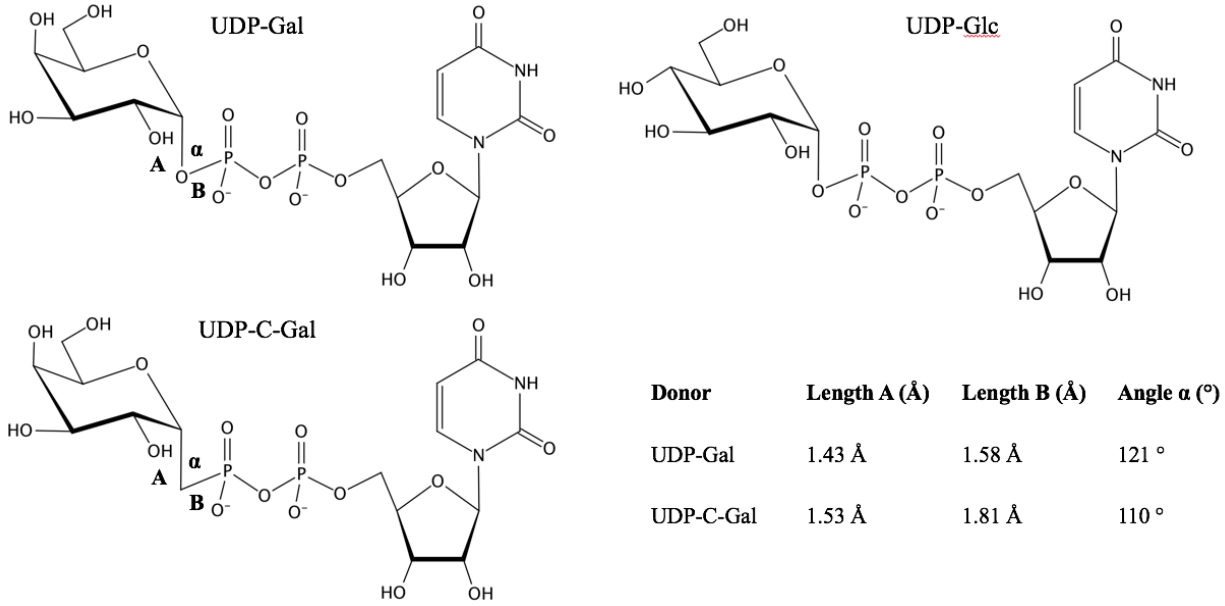
Chapter 5: Donor substrate conformational changes

5.1 Introduction to Chapter 5

This chapter addresses the second aim of this dissertation: to determine how donor substrate binds and achieves the constrained, high-energy conformation associated with catalysis. GTA and GTB are the near-homologous enzymes responsible for catalyzing the final step in ABO(H) blood group antigen synthesis. Since their initial structural characterization in 2002 (Patenaude et al. 2002), it has been a challenge to obtain complexes of GTA and GTB with intact donor substrate. Several factors long have obstructed these efforts: substrate hydrolysis, oxidation, and substrate and enzyme conformational changes, which can shatter crystal lattices (Boix, E., et al. 2001, Angulo, J., et al. 2006, Blume, A., et al. 2006, Alfaro, J.A., et al. 2008). Though in-solution approaches, including NMR and MS, have uncovered some limited aspects of GTA/GTB-substrate interactions (Soya, N., et al. 2009, Rademacher, C., et al. 2010, Sindhuwinata, N., et al. 2010), in the absence of liganded crystal structures, our understanding of substrate binding and enzyme conformational changes is inadequate.

Non-hydrolysable analogues can overcome one of the main obstacles to crystallization and further, such analogues potentially can act as competitive, mechanism-based GT inhibitors (Gordon, R.D., et al. 2006, Clarke, A.J., et al. 2008, Beaton, S.A., et al. 2009, Partha, S.K., et al. 2010). Replacement of the UDP-Gal anomeric oxygen with a methylene group (**Figure 11**) yields UDP-phosphono-galactose (UDP-C-Gal), a non-hydrolysable donor substrate that can act as a probe of the glycosyltransfer in GTA/GTB. Previously, UDP-C-Gal was used in structural studies probing the UDP galactopyranose mutase reaction but was found to bind the enzyme in an alternate conformation where the β-phosphate and galactosyl moiety of the analogue were shifted relative to the “U-shaped” natural donor conformation (Partha, S.K., et al. 2010).

Figure 11. *Donor substrates UDP-Gal, UDP-Glc, and UDP-C-Gal*



Donor	Length A (Å)	Length B (Å)	Angle α (°)
UDP-Gal	1.43 Å	1.58 Å	121 °
UDP-C-Gal	1.53 Å	1.81 Å	110 °

Key donor bond angles and lengths are displayed. The UDP-C-Gal analogue differs from UDP-Gal in replacement of the anomeric oxygen with a mythylene group, while UDP-Glc is an epimer of UDP-Gal and differs in the equatorial position of Gal-C4-OH. Labeled UDP-C-Gal carbon-carbon and carbon-phosphorous bonds are longer than UDP-Gal carbon-oxygen and oxygen phosphorous bonds, and the C-C-P bond angle of the analogue is much smaller than the C-O-P bond angle of UDP-Gal. Adapted from (Gagnon et al., 2015) with permission.

From structures of GTA/GTB chimeras, the donor substrate has been observed to bind in a high-energy catalytic conformation, where the enzyme's flexible internal loop and C-terminal tail become ordered about the active site (see **Chapter 1.5**). Notably, no wild-type GTA or GTB structures have been collected in complex with intact donor. It has been unclear how donor substrate binds and achieves the constrained, high-energy conformation associated with catalysis, and the contribution of loop ordering to this process is not yet fully understood.

Toward this second aim I undertook an examination of the X-ray crystal structures of GTA, GTB, and their chimeras in complex with natural donors UDP-Gal, UDP-Glc, and the non-hydrolysable donor analogue UDP-C-Gal. UDP-Glc, the C4 epimer of natural donor UDP-Gal, binds GTB with a similar affinity to UDP-Gal but shows much lower enzymatic activity: 0%

relative to GTA/UDP-GalNac and 0.02% relative to GTB/UDP-Gal. Accordingly, UDP-Glc, was included to provide insight into these enzymes' stereoselectivity. Toward the published data reported in Chapter 5, my contributions include processing, solving, and refining all X-ray diffraction data, analyzing and interpreting structural data, producing figures and tables, and writing much of the paper.

5.2 Results

GTA, GTB, and chimera structures

Data collection and refinement statistics are given in **Table 4** for structures in complex with UDP-C-Gal, **Table 5** for structures in complex with UDP-Gal, and **Table 6** for structures in complex with UDP-Glc. Overall, the resolution limit ranges from 1.55-1.39 Å, R_{work} from 17.2-20.0% and R_{free} from 18.6-21.8%. Observed electron density surrounding the internal loop and the C-terminal tail for all structures is summarized in **Table 7**. Apart from these disordered regions, the entire length of the polypeptide chains had excellent electron density.

GTA, GTB, and all chimeras, when in complex with UDP-C-Gal, have disordered C-terminal tails regardless of acceptor binding. In this set of structures, only GTA is in the semi-closed state with an organized internal loop. The remaining complexes are in the open state, although under longer soaking conditions ABBB becomes more ordered. Due to weak or ambiguous density, the UDP-C-Gal sugar moiety was not modeled for GTA and GTB (with HA acceptor), and the UDP-Glc sugar moiety was not modeled for GTB with HA acceptor.

Generally, structures collected in complex with native GTB donor UDP-Gal show much greater ordering than those collected in complex with UDP-C-Gal or UDP-Glc. GTA and ABBA are in the closed state with the internal loop and C-terminus ordered (**Table 7**) when in complex with UDP-Gal and DI, similar to the previously published AABB structure (Alfaro,

J.A., et al. 2008). ABBB is in the semi-closed state with a flexible C-terminus, and GTB is in the open state with both regions disordered.

Table 4. Data collection and refinement statistics for GTA, GTB, and their chimera in complex with UDP-C-Gal.

All crystal structures are in space group $C222_1$ with unit-cell dimensions *a*= 52.5-52.7, *b*= 149.0-149.9, *c*= 79.2-79.6 Å.

	DI							HA
	AAAA	**AABB**	**ABBA Short Soak**	**ABBA Long Soak**	**ABBB Short Soak**	**ABBB Long Soak**	**BBBB**	**BBBB**
Resolution (Å)	20.00-1.55	20.00-1.45	20.00-1.40	20.00-1.40	20.00-1.39	20.00-1.39	20.00-1.46	20.00-1.49
R_{merge} (%)[a,b]	4.2 (28.7)	4.3 (30.8)	3.2 (30.2)	3.3 (29.3)	3.0 (28.6)	3.2 (26.9)	4.1 (26.9)	3.7 (27.6)
Completeness (%)[b]	98.8 (100.0)	99.5 (100.0)	96.9 (85.0)	96.6 (82.4)	96.0 (87.0)	98.0 (91.1)	99.1 (98.0)	95.3 (91.1)
Unique Reflections	45,201	55,396	60,216	59,599	60,710	61,958	54,000	48,905
Multiplicity	4.53 (3.74)	4.26 (3.45)	3.89 (1.48)	3.97 (2.50)	4.05 (2.05)	4.05 (2.03)	4.50 (3.71)	5.38 (4.81)
I/σ(I) Highest resolution shell	3.9	3.2	1.9	2.9	3.0	3.0	4.0	4.8
Refinement								
R_{work} (%)[c]	18.2	18.1	20.0	18.4	19.7	17.8	18.2	17.8
R_{free} (%)[c,d]	21.7	19.5	21.8	20.4	21.3	19.9	19.0	19.1
No. Water	231	262	217	259	248	243	223	215
r.m.s. bond (Å)[e]	0.0105	0.0098	0.0097	0.0094	0.0092	0.0091	0.0097	0.0102
r.m.s. angle (°)[e]	1.516	1.504	1.532	1.553	1.437	1.528	1.540	1.542
PDB code	5C36	5C38	5C3A	5C3B	5C3D	5C47	5C48	5C49

[a] R-merge, $\sigma\ |I_{obs} - I_{avg}| / \sigma\ I_{avg}$
[b] Values in parentheses represent highest resolution shell.
[c] R-work, $\sigma\ ||F_o|-|F_c|| / S\ |F_o|$
[d] 10% of reflections were omitted for R-free calculations
[e] r.m.s. root-mean-square.
Reproduced from (Gagnon et al. 2015) with permission.

Table 5. Data collection and refinement statistics for GTA, GTB, and their chimera in complex with UDP-Gal

All structures are in space group $C222_1$ with unit-cell dimensions *a*=52.5-52.7, *b*=148.9-149.4, *c*=79.2-80.0 Å.

	AAAA	**ABBA**	**ABBB**	**BBBB**
Resolution (Å)	20.00-1.40	20.00-1.46	20.00-1.55	20.00-1.40
R_{merge} (%)[a,b]	3.3 (30.7)	3.1 (24.0)	3.4 (27.2)	3.2 (31.1)
Completeness (%)[b]	95.1 (81.4)	96.5 (97.0)	91.8 (91.4)	96.1 (81.7)
Unique Reflections	59,210	52,594	42,208	59,299
Multiplicity	4.07 (2.54)	4.28 (3.44)	3.71 (3.28)	4.52 (2.82)
I/σ(I) Highest resolution shell	3.1	4.4	3.9	3.2
Refinement				
R_{work} (%)[c]	18.0	17.5	17.8	17.2
R_{free} (%)[c,d]	19.9	19.2	21.4	19.6
No. Waters	308	224	238	274
r.m.s. bond (Å)[e]	0.0100	0.0097	0.0096	0.0095
r.m.s. angle (°)[e]	1.507	1.506	1.555	1.561
PDB code	5BXC	5C1G	5C1H	5C1L

[a] R-merge, $\sigma |I_{obs} - I_{avg}| / \sigma I_{avg}$

[b] Values in parentheses represent highest resolution shell

[c] R-work, $\sigma ||F_o|-|F_c|| / S |F_o|$

[d] 10% of reflections were omitted for R-free calculations

[e] r.m.s. root-mean-square

Reproduced from (Gagnon et al. 2015) with permission.

Table 6. Data collection and refinement statistics for GTA, GTB, and their chimera in complex with UDP-Glc

All crystal structures are in space group $C222_1$ with unit-cell dimensions *a*= 52.5-52.6, *b*= 149.3-149.7, *c*= 79.2-79.5 Å.

			DI			HA
	AAAA	**AABB**	**ABBA**	**ABBB**	**BBBB**	**BBBB**
Resolution (Å)	20.00-1.54	20.00-1.41	20.00-1.45	20.00-1.43	20.00-1.40	20.00-1.55
R_{merge} (%)[a,b]	4.2 (35.1)	3.6 (32.8)	3.0 (22.1)	3.2 (26.3)	3.3 (27.8)	4.6 (27.3)
Completeness (%)[b]	94.4 (89.1)	97.2 (82.0)	97.9 (98.8)	97.7 (96.3)	98.1 (87.1)	98.5 (99.7)
Unique Reflections	44,020	58,865	54,336	56,554	60,723	45,267
Multiplicity	4.48 (3.80)	4.42 (3.23)	4.25 (3.40)	4.25 (3.20)	4.16 (2.46)	4.20 (3.49)
I/σ(I) Highest resolution shell	3.0	3.2	4.6	4.0	3.3	3.7
Refinement						
R_{work} (%)[c]	18.2	18.8	17.6	18.5	17.5	17.3
R_{free} (%)[c,d]	21.0	20.5	18.6	20.9	18.9	19.8
No. Water	235	252	217	220	272	275
r.m.s. bond (Å)[e]	0.0101	0.0095	0.0091	0.0097	0.0095	0.0114
r.m.s. angle (°)[e]	1.552	1.503	1.509	1.501	1.532	1.612
PDB code	5C4B	5C4F	5C8R	5C4C	5C4D	5C4E

[a] R-merge, $\sigma\ |I_{obs} - I_{avg}| / \sigma\ I_{avg}$

[b] Values in parentheses represent highest resolution shell.

[c] R-work, $\sigma\ ||F_o|-|F_c|| / S\ |F_o|$

[d] 10% of reflections were omitted for R-free calculations

[e] r.m.s. root-mean-square.

Reproduced from (Gagnon et al. 2015) with permission.

All UDP-Glc complexes have more disordered C-terminal tails and internal loops except for GTA and GTB. GTA is in the semi-closed state with UDP-Glc and DI bound, while AABB, ABBA, ABBB, and BBBB (GTB) are in the open state. The UDP-Glc sugar moiety was modeled for all complexes except BBBB (GTB), where electron density was weak.

Four distinct donor conformations

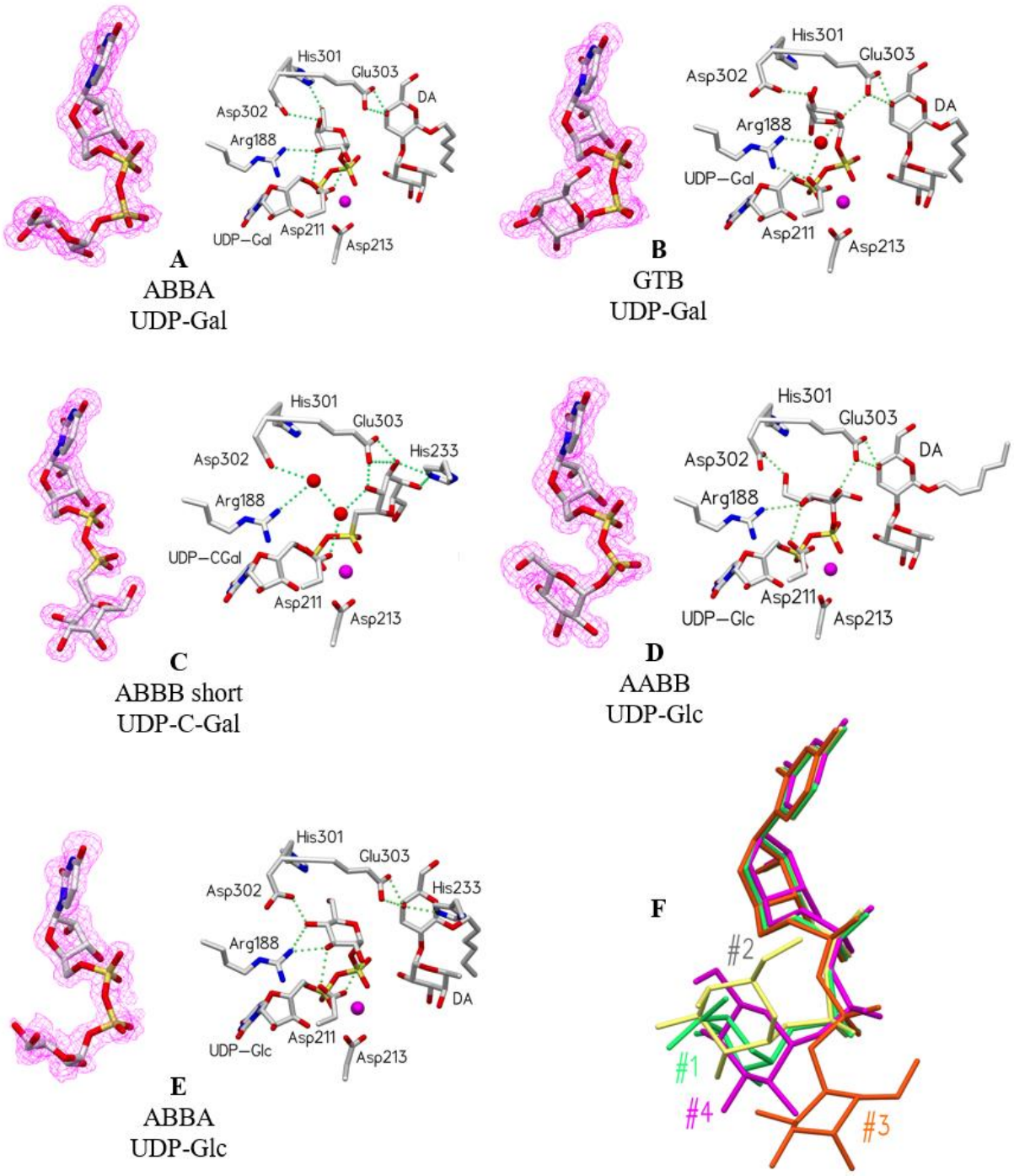
Among the structures with an ordered donor sugar there are a total of four distinct donor conformations, each supported with excellent electron density (**Figure 12**). In all cases, it is the donor sugar moiety that is observed to shift with respect to the enzyme, with some movement about the β-phosphate while the uracil maintains its position. Notably, in each of the four donor conformations there is a distinct set of substrate-enzyme hydrogen bond interactions, shown in **Figures 12 & 13**.

The so-called "tucked-under" UDP-Gal conformation (conformation #1; **Figures 12A & 13A**) associated with catalysis in GTs has the epimeric carbon positioned adjacent to the phosphate groups. This conformation was first reported for GTA and GTB by Alfaro *et al.* (2008) for UDP-Gal bound to AABB. The data presented in this chapter shows that UDP-Glc also can become tucked-under when in complex with ABBA.

A second conformation (conformation #2) is observed for GTB (**Figure 12B**) in complex with DI and UDP-Gal or UDP-C-Gal. Over longer soaking times with DI acceptor bound, UDP-C-Gal adopts this conformation. In conformation #1, Gal is oriented such that there are hydrogen bond interactions between Arg188 and Asp211 and Gal-3-OH, Asp302 and Gal-4-OH, His301 and Gal-O-6, Asp211 and the UDP β-phosphate oxygen, and His233 and Glu303 and DI acceptor (**Figures 12A & 13A**). Here, internal loop residue Trp181 lies far from Gal-O-6

(**Figure 15**). In conformation #2 (**Figures 12B & 13B**) the galactosyl group shifts so that Arg188 and Asp211 no longer contact Gal-O-3 directly and instead interact, *via* a water molecule with Gal-O-2.

Figure 12. *Four donor conformations and enzyme-substrate interactions*



(**A**) ABBA+UDP-Gal+DI with tucked under donor (conformation #1), (**B**) GTB+UDP-Gal+DI with donor in an alternative conformation (#2). In (**C**) ABBB+UDP-Glc+DI (short soak), UDP-C-Gal is in an extended conformation (#3). Here acceptor is absent, and donor extends into the acceptor binding site. In (**D**) AABB+UDP-Glc+DI, UDP-Glc is in conformation #4. In (**E**) ABBA+UDP-Glc+DI, Glc is in the same conformation as (**A**), but with the sugar 4-OH in the equatorial position. (**F**) Overlapped donor conformations show movement of the saccharide moiety and the β-phosphate. ABBA+UDP-Gal is in

green, GTB+UDP-Gal is in *yellow*, ABBB short+UDP-C-Gal is in *orange*, and AABB+UDP-Glc is in *magenta*, with conformation numbers indicated in the image. Electron density diagrams are $2F_O$-F_c maps contoured at 1.00 σ. Protein atoms are colored by element with oxygen *red*, phosphorous *yellow*, carbon *grey*, and nitrogen *blue*. Waters are shown as red spheres and Mn^{2+} is shown as a magenta sphere. Green dotted spheres represent hydrogen bonds. Reproduced from (Gagnon et al., 2015) with permission.

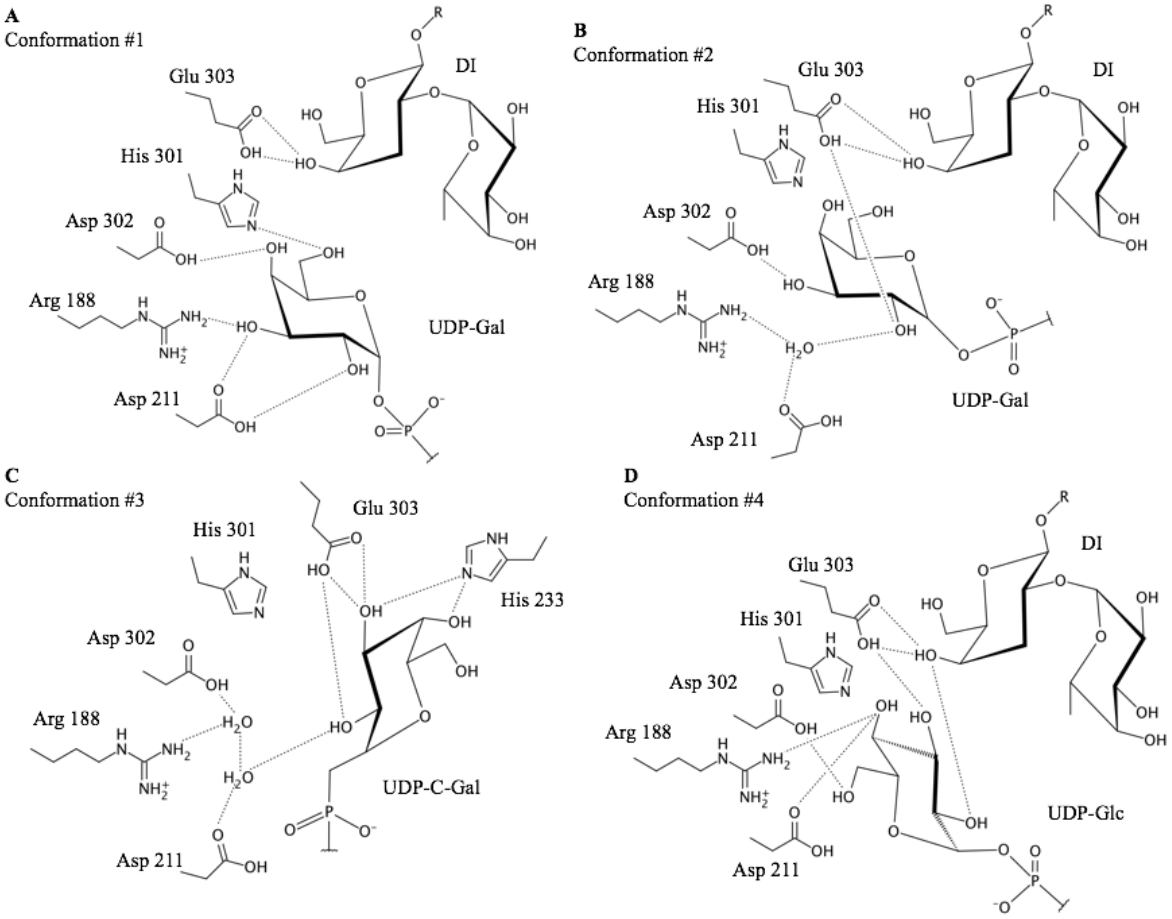
Here, Asp302 binds Gal-O-3, Glu303 interacts with Gal-O-2, His301 still forms a hydrogen bond to Gal-O-6, and Glu303 still contacts DI. Gal-O-6 is now in a position distinct from conformation #1, and here in conformation #2 it prevents Trp181 from approaching the active site as required to form the closed conformation (**Figure 14**). With UDP-Gal in position #2, GTB is in the semi-closed state with a partially disordered internal loop and a highly disordered C-terminal tail (**Table 7**).

UDP-C-Gal binds ABBB in a third, "extended" conformation (#3) under short soaking time conditions (**Figure 12C**). In these structures, two water molecules are introduced into the active site and DI is absent. C-galactosyl moiety extends into the acceptor-binding site such that there are contacts between Glu303 and Gal-O-2 and O-3 as well as His233 and Gal-O-3 and Gal-O-4. Unlike in conformations #1 and #2, active site residues Arg188, Asp211, Asp302, and His301 are no longer involved in binding the C-Gal group. Carbon-phosphorous bond lengths of the modified donor are longer than corresponding oxygen-phosphorous bonds in UDP-Gal structures, and key differences in C-C-P (110°) and C-O-P (121°) bond angles contribute to UDP-C-Gal's extended conformation (**Figure 11**).

UDP-Glc occupies the donor-binding sites of GTA, AABB, ABBB, and GTB in a fourth distinct conformation (#4), where the glucose moiety is rotated so that each hydroxyl group changes binding partners relative to the first UDP-Gal conformation (**Figure 12D**). Glc-O2 no longer contacts with Asp211 and instead interacts with DI, Glu303 interacts with Glc-O3, Arg188 and Asp211 make contacts to Glc-O4, and Asp302 forms a hydrogen bond to Glc-O6.

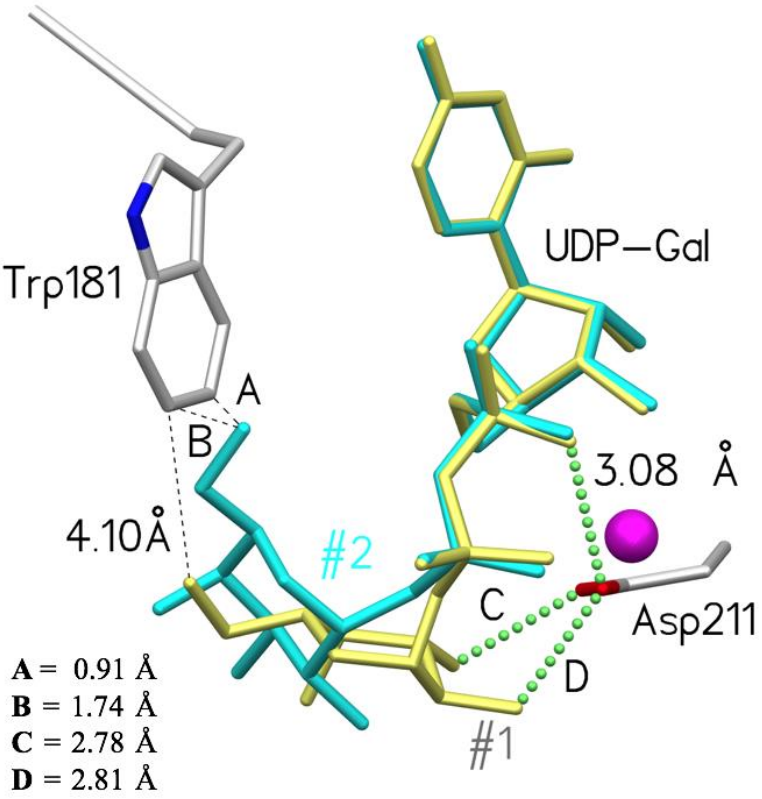
When in complex with ABBA, Glc is in conformation #1 but with Glc-O4 in the equatorial position (**Figure 12E**), which permits an additional interaction with Arg188 and causes a shift in Asp302 to maintain the hydrogen bond to Glc-O4.

Figure 13. *Hydrogen bond schematic for substrate-enzyme interactions in four donor conformations*



Hydrogen bond schematic for donor- and acceptor-enzyme interactions in the four observed conformations. (**A**)-(**E**) correspond to Figure 13 where (**A**) is tucked under UDP-Gal (#1) bound to ABBA, (**B**) is UDP-Gal in conformation #2 bound to GTB, (**C**) is UDP-C-Gal in conformation #3 bound to ABBB (short soak), and (**D**) is UDP-Glc in conformation #4 bound to AABB. Enzyme residues that participate in hydrogen bonding with donor and acceptor are indicated. Hydrogen bonds are shown as dashed lines. Adapted from (Gagnon et al., 2015) with permission.

Figure 14. *UDP-Gal conformations influence enzyme ordering* via *Trp181*



UDP-Gal is in conformation #1 (*yellow*) and the enzyme is in the closed state in ABBA+UDP-Gal+DI. Here, Trp181 is distant from Gal-O-6 and Lys346 interacts with the β-phosphate oxygen. When UDP-Gal is in conformation #2 (*cyan*), as seen in complex with GTB, the enzyme is in the semi-closed state with a less-ordered internal loop and a disordered C-terminal tail. Here, Trp181 is near Gal-O-6, and Asp211 interacts with the β-phosphate oxygen. Mn^{2+} is shown as a magenta sphere. Green dotted spheres represent hydrogen bonds, and black dashed lines indicate measured distances that are not hydrogen bond interactions. Reproduced from (Gagnon et al., 2015) with permission.

Table 7 Loop ordering in AAAA, BBBB, and chimeric enzyme complexes.

Black single letter amino acid codes correspond to electron density for main and side chain atoms; blue letters correspond to electron density for main chain atoms only; red letters correspond to weak electron density for both main and side chain atoms; lower case letters represent amino acids omitted from refined models.

Enzyme+substrates	Donor sugar e^- density	Donor conformation	Acceptor e- density	**Internal loop** 176 181 186	**C-terminus** 346 351
UDP-C-Gal complexes					
AAAA+UDP-C-Gal+DI	None	n/a	Good	EV RAYKR WQDVS MR	VP knhqa vrnp
AABB+UDP-C-Gal+DI	Good	#2	Good	EV RAykr wqdvs MR	VP KNhqa vrnp
ABBA+UDP-C-Gal short soak+DI[a]	Good	#3 (#2)	Poor	EV Raykr wqdvS MR	VP Knhqa vrnp
ABBA+UDP-C-Gal long soak+DI	Good	#2 (#3)	Good	EV Raykr wqdvS MR	VP Knhqa vrnp
ABBB+UDP-C-Gal short soak+DI	Good	#3	Poor	EV RAYkr wqdvS MR	VP knhqa vrnp
ABBB+UDP-C-Gal long soak+DI	Partial	#2	Good	EV RAYkr wqdvS MR	VP KNhqa vrnp
BBBB+UDP-C-Gal +DI	Good	#2	Good	EV GAYkr wqdVS MR	VP Knhqa vrnp
BBBB+UDP-C-Gal +HA	Poor	n/a	Good	EV Gaykr wqdVS MR	VP Knhqa vrnp
UDP-Gal complexes					
AAAA+UDP-Gal+DI	None	n/a	Good	EV RAYKR WQDVS MR	VP KNHQA VRNP
AABB+UDP-Gal+DI[b]	Good	#1	Good	EV RAYKR WQDVS MR	VP KNHQA VRNP
ABBA+UDP-Gal+DI	Good	#1 (#2)	Good	EV RAYKR WQDVS MR	VP KNHQa Vrnp
ABBB+UDP-Gal+DI	Good	#2 (#1)	Good	EV RAYKR WQDVS MR	VP KNhqa vrnp
BBBB+UDP-Gal+DI	Good	#2	Good	EV GAykr wqdVS MR	VP Knhqa vrnp
UDP-Glc complexes					
AAAA+UDP-Glc+DI	Good	#4	Good	EV RAYKR WQDVS MR	VP knhqa vrnp
AABB+UDP-Glc+DI	Good	#4	Good	EV RAYkr wqdvS MR	VP Knhqa vrnp
ABBA+UDP-Glc+DI	Good	#1	Good	EV Raykr wqdvs MR	VP Knhqa vrnp
ABBB+UDP-Glc+DI	Good	#4	Good	EV RAYkr wqdvS MR	VP Knhqa vrnp
BBBB+UDP-Glc+DI	Good	#4	Good	EV GAYkr wqdvS MR	VP Knhqa vrnp
BBBB+UDP-Glc+HA	None	n/a	Good	EV GAYKR wqDVS MR	VP KNHQA VRNp

[a] value in parentheses indicates a lower occupancy alternate conformation

[b] as reported in Alfaro *et al*., 2008 PDB code: 2RJ7

Reproduced from (Gagnon et al. 2015) with permission.

Multiple donor sugar conformations in single crystal structures

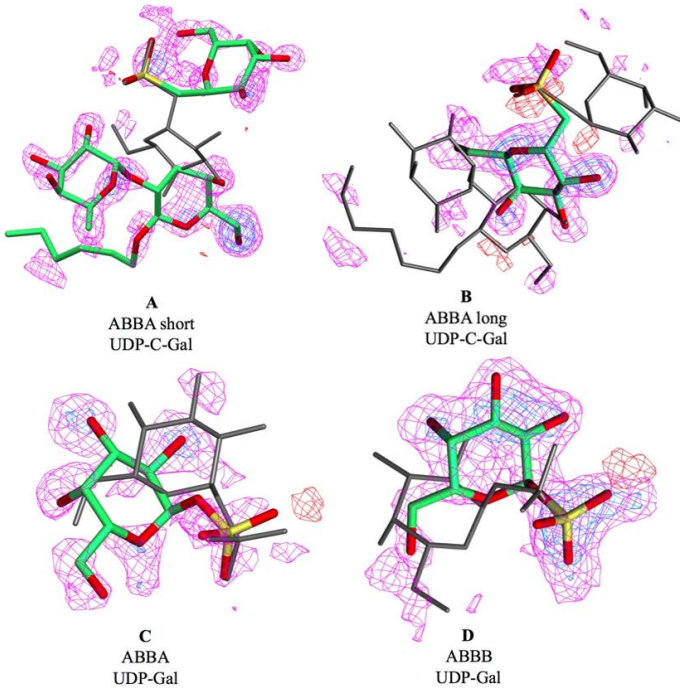
In both ABBA+UDP-C-Gal structures, the donor was disordered over conformations #2 and #3, (**Figure 15A&B**). In the short soak structure, extended conformation #3 was dominant over #2 in an 80:20 ratio, while in the long soak structure conformation #2 was dominant over #3 in a 70:30 ratio. In complex with AABB, ABBB (short soak), ABBB (long soak), and GTB (DI), UDP-C-Gal had 100% occupancy, while it was completely disordered in GTA and GTB (HA) structures. Like UDP-C-Gal, UDP-Gal was disordered over conformations #1 and #2 in an approximate ratio of 70:30 in the active site of ABBA and with an approximate ratio of 40:60 in the active site of ABBB (**Figure 15C&D**), as determined *via* examination of electron density and B-factors (of the donor sugar moiety relative to nearby enzyme residues) following refinement with difference occupancy values set for substrate. UDP-Gal was 100% occupied when in complex with GTB and completely disordered when in complex with GTA. Since all enzymes were soaked with the same batch of donor, the poor density for the donor sugar observed in the GTA+UDP-Gal structure could be due to hydrolysis or disorder of the donor Gal moiety. UDP-Glc did not exhibit any intra-structure conformational variety: it was ordered when in complex with GTA, AABB, ABBA, ABBB, and GTB (DA) and disordered when in complex with GTB (HA)

5.3 Discussion & conclusions

In total, the donors and donor analogues were observed in four distinct conformations (**Figure 12F, Table 7**). Previously, intact donor has been observed in only two studies: Alfaro *et al.* (2008), where UDP-Gal is tucked-under in complex with AABB and DI (Alfaro, J.A., et al. 2008), and Schuman *et al.* (2010), where UDP-Gal is in conformation #2 in complex with GTB

Cys/Ser mutants. In neither case was there a wild-type structure with intact donor bound. The observation here of several new structures, including wild-type GTB, with bound donor intact represents a significant advance in understanding donor recognition in this class of enzymes.

Figure 15. *Difference maps for structures with multiple donor conformations in the active site*



In (**A**) ABBA+UDP-C-Gal+DI (short soak), donor is primarily in conformation #3 (0.8 occupancy) but is also in conformation #2. In (**B**) ABBA+UDP-C-Gal+DI (long soak), the donor is in conformations #2 (0.7 occupancy) and #3. In (**C**) ABBA+UDP-Gal+DI, the donor is in conformation #1 (0.7 occupancy) and conformation #2. In (**D**) ABBB+UDP-Gal+DI the donor is in conformations #1 (0.6 occupancy) and #2. In each case the dominant conformation (for which $2F_O$-F_C maps have been previously provided in Figure 13) is in *gray*, and the alternate conformation is colored by element with carbon *green*, oxygen *red*, and phosphorous *yellow*. Electron density diagrams are F_O-F_C maps contoured at -3.00 (*red*), 2.00 (*magenta*), and 4.00 σ (*blue*). Alternate conformations are modeled into the difference maps and were not included during structure refinement for map generation.

UDP-C-Gal is a flawed UDP-Gal analogue

There is a scarcity of GT structures determined in complex with donor substrates. Partly, this results from the enzymes' ability to hydrolyse donor in the absence of acceptor as well as the

enzymes' tendency to undergo substrate binding-induced conformational rearrangement, which can shatter the crystal lattice (Boix, E., et al. 2001, Angulo, J., et al. 2006, Blume, A., et al. 2006, Alfaro, J.A., et al. 2008, Soya, N., et al. 2009).

The use of the non-hydrolysable substrate UDP-C-Gal negates the first barrier however, the analogue contains structural differences from the natural donors that preclude complete mimicry (**Figure 11**): the anomeric C-O bond length of 1.4 Å changes to 1.55 Å for the C-C bond, and the C-O-P bond angle is close to a trigonal 123° ± 11° (averaged from UDP-sugar structures in the RCSD PDB), since the anomeric oxygen can assume partial sp^2 character, while the purely sp^3 methylene group of UDP-C-Gal would be expected to be closer to a tetrahedral 109°.

UDP-C-Gal adopts two distinct conformations in the active sites of six of the structures obtained from crystals in which it was soaked. The analogue is most ordered in the chimeric enzyme structures (**Table 7**), while in GTA and GTB the galactosyl residue is completely disordered. In four chimeras UDP-C-Gal is in a single conformation, either #2 or #3, while in two chimeras there is clear evidence that it is disordered over conformations #2 and #3 (**Figure 15A&B**). UDP-C-Gal does not assume the tucked-under conformation associated with catalysis in any of these structures. This inability to assume the catalytically competent conformation likely results from bond angle and length differences described above (**Figure 11**).

In conformation #3 (**Figure 12C & 13C**), UDP-C-Gal utilizes almost none of the stabilizing donor-enzyme interactions seen in the tucked-under conformation. Acceptor is absent and the Gal moiety makes enzyme contacts normally reserved for acceptor: His233 and Glu303 side chains make hydrogen bond interactions with C-Gal-O2, -O3, and -O4 hydroxyls. The rotated donor sugar (#3) accommodates two new water molecules in the active site, and these

form bridging hydrogen bonds with the side chains of Asp302, Arg188, and Asp211, residues normally in direct contact with donor sugar hydroxyls in conformations #1 and #2 (**Figures 12C & 13C**).

UDP-C-Gal promotes the open conformation

Generally, the GTA/GTB internal loop has been observed to become more ordered in the presence of both donor and acceptor (Alfaro, J.A., et al. 2008), and enzymes with greater GTA character order more readily than enzymes with greater GTB character (Alfaro, J.A., et al. 2008, Johal, A.R., et al. 2012, Johal, A.R., et al. 2014). This trend is consistent with the UDP-Gal-bound structures presented reported here (**Table 7**, middle): GTA+UDP-Gal is closed, ABBA+UDP-Gal is nearly closed, ABBB+UDP-Gal is semi-closed, and GTB+UDP-Gal is open.

UDP-C-Gal/DI complexes deviate from this trend and are biased toward the open form (**Table 7**, top). GTA is known to be the best-ordered or least labile of all wild-type and chimeric enzymes and, while it is in the closed state when in complex with UDP-Gal, with the analogue bound it only achieves the semi-closed state. All other UDP-C-Gal structures are in the open conformation, which can be attributed to UDP-C-Gal's inability to tuck under. Without tucking under, there is no resolution to the steric clash between internal loop residue Trp181 of and the acceptor, which in turn impedes organization of the C-terminal tail.

Intact UDP-Gal bound to wild-type GTB

Here we report of the first crystal structure of wild-type GTB in which unambiguous electron density is seen for intact natural donor UDP-Gal. Interestingly, the substrate adopts the same conformation (#2) in the wild-type GTB structure that was observed for GTB Cys/Ser

mutants (Schuman, B., et al. 2010) (**Figure 12B**), but in this case the enzyme is not in the closed state. While apparently similar, the donor-enzyme contacts in conformations #1 and #2 are almost entirely shifted: relative to tucked-under donor (**Figure 12A**), the enzyme no longer interacts with Gal-O6, and Arg188 and Asp211 no longer interact with the galactosyl moiety directly but make bridging contacts through a newly introduced water molecule (**Figure 13B**).

UDP-Glc can adopt conformation #1

In these structures, UDP-Glc is observed in a fourth distinct conformation (#4, **Figure 12D**), so far unique to UDP-Glc, where five hydrogen bonds stabilize the glucosyl moiety (**Figures 12D & 13D**), and there is an additional contact between DI and Glc-O-2. However, UDP-Glc is observed to assume the tucked under conformation (#1, **Figure 12E**) as seen in complex with ABBA. UDP-Glc is a poor donor for these enzymes: with GTA, the observed k_{cat} is too low to measure and for GTB the rate of transfer is low but measurable (k_{cat} = 0.0010 s^{-1}) (Seto, N.O.L., et al. 2000). NMR studies have demonstrated that UDP-Gal and UDP-Glc bind GTB with the same affinity. In the latter case binding is unproductive (little to no product is formed), since transfer to acceptor does not occur readily (Angulo, J., et al. 2006, Blume, A., et al. 2006). Based on NMR data corresponding to GTB-bound UDP-Glc, Angulo *et al.* (2006) proposed a "tweezers" mechanism whereby Asp302 and Glu303 lock the galactosyl moiety (of the natural donor UDP-Gal) in the tucked-under conformation to facilitate formation of the transition state. The authors suggest that UDP-Glc, in contrast to natural donor, is incapable of undergoing such a conformational transition on the basis that Glc-O4 and Asp302 are unable to interact. However, the structure of ABBA described here shows clearly that Glc-O4 in fact forms a strong hydrogen bond (2.62 Å) to Asp302 (**Figure 12E**), and that UDP-Glc is quite capable of

entering the tucked-under conformation. The enzymes' selectivity between UDP-Gal and UDP-Glc more likely depends on internal and C-terminal loop organization.

The structure of ABBA in complex with tucked-under UDP-Glc displays considerably more disorder than the corresponding complex with UDP-Gal (**Table 7**). In this structure shows residues Met186 and Ser185 are offset significantly compared to their position in the ordered AABB+UDP-Gal structure (PDB code 2RJ7). The resulting internal loop and C-terminal tail disorder displaces Trp181, Lys346, and Arg352, residues key in stabilizing substrate prior to glycosyltransfer. Although the orientation of C4-OH seems compatible with conformation #1, nevertheless it has a significant effect on mobile loop organization and the enzymes' ability to adopt the catalytically active state. This may explain why the UDP-Gal epimer, UDP-Glc, does not undergo transfer as readily as the natural GTB donor.

Substrate-binding is a multi-step process

Using NMR methods, Angulo *et al.* (2006) observed conformation #1 in the active site of GTB and hypothesized that the enzymes could select for the catalytically competent conformation despite its low abundance in solution. However, bound UDP-Gal conformation #2 is similar to the dominant low energy conformer observed in solution (Angulo, J., et al. 2006). Thus, our data suggest that the enzymes can bind abundant low-energy UDP-Gal conformers and effect a shift to the catalytically competent conformation. Indeed, the selection of the extended confomer #3 from solution may explain the kinetically–observed phenomenon that donor binding precedes acceptor binding, as the structures with the bound donor in conformation #3 clearly show that this conformation sterically precludes acceptor binding (Kamath, V.P., et al. 1999).

In four structures, there is clear evidence that the donor sugar is disordered over two conformations (**Figure 15A-D**), indicating that the enzymes do in fact bind the substrates in the predominant extended conformation #3 and shift it in a stepwise fashion to the tucked-under conformation #1. Further, the relative conformational distributions in these structures (70:30, 60:40, 80:20) listed in **Figure 15** suggest that conformations #1-3 are of similar free energy. The two ABBA/UDP-C-Gal structures (short and long soaks) provide the strongest evidence. The short soak structure shows the donor analogue predominantly in the extended conformation (#3), with a small proportion in conformation #2, while for the long soak structure the relative proportion of each conformation is reversed. Further evidence for a stepwise process comes from ABBA and ABBB structures in complex with UDP-Gal. Here, donor is disordered over conformations #1 and #2, with ABBA being predominantly #1 and ABBB being predominately #2.

Evaluation of donor analogue

The UDP-C-Gal donor was designed to circumvent the problem of hydrolysis and by extension to allow visualization of the donor sugar residue in the GTA/GTB active site. Thus, it is somewhat paradoxical that the sugar moiety is not observed in two of the structures (**Table 7**). In both wild-type enzymes, there is unambiguous density for UDP bound to the donor-binding site but weak or absent density for the C-Gal moiety, indicating that the moiety is disordered. For GTA this is understandable, since its active site accommodates the much larger *N*-acetylgalatosamine residue, and in structures collected in complex with UDP-Gal, the donor sugar also appears disordered. The wild-type GTB structures are interesting, since in the complex that includes both UDP-C-Gal and DI, the active site is organized with the donor

analogue in conformation #2, while in the UDP-C-Gal and HA complex, the donor analogue sugar group is completely disordered. HA only differs from DI in the addition of a 3-OH group on the galactosyl residue. This hydroxyl group clearly results in large steric clashes, as a superposition of the two structures would put the 3-OH group only ~2.5 Å away from C1 of the ring, ~3.3 Å from the methylene carbon, and ~2.89 Å from one of the phosphate oxygen atoms.

Conclusions

In complex with GTA, GTB, and chimeras, the UDP-C-Gal donor analogue approaches but cannot achieve the tucked-under catalytic conformation, likely due to geometric constraints. The distinct conformations of the C-Gal moiety do not permit formation of the semi-closed or closed states, and its utility as a probe of the GTA/GTB transfer mechanism may be limited. Overall, the donors and donor analogues were seen to bind these enzymes in four conformations in the active site. UDP-Glc binds in a distinct conformation but like UDP-Gal, it may adopt the tucked-under position. It is likely that the enzymes distinguish between epimers *via* the effect of the 4-OH group on loop ordering. Together these structures provide insight into GT inhibitor development and call into question the potential utility of C-phosphonate analogues as biologically relevant sugar nucleotide mimics. In addition, the multiple observed conformers suggest that donor binding is a multi-stage process that implements significant stepwise shifts to achieve the catalytically active conformation.

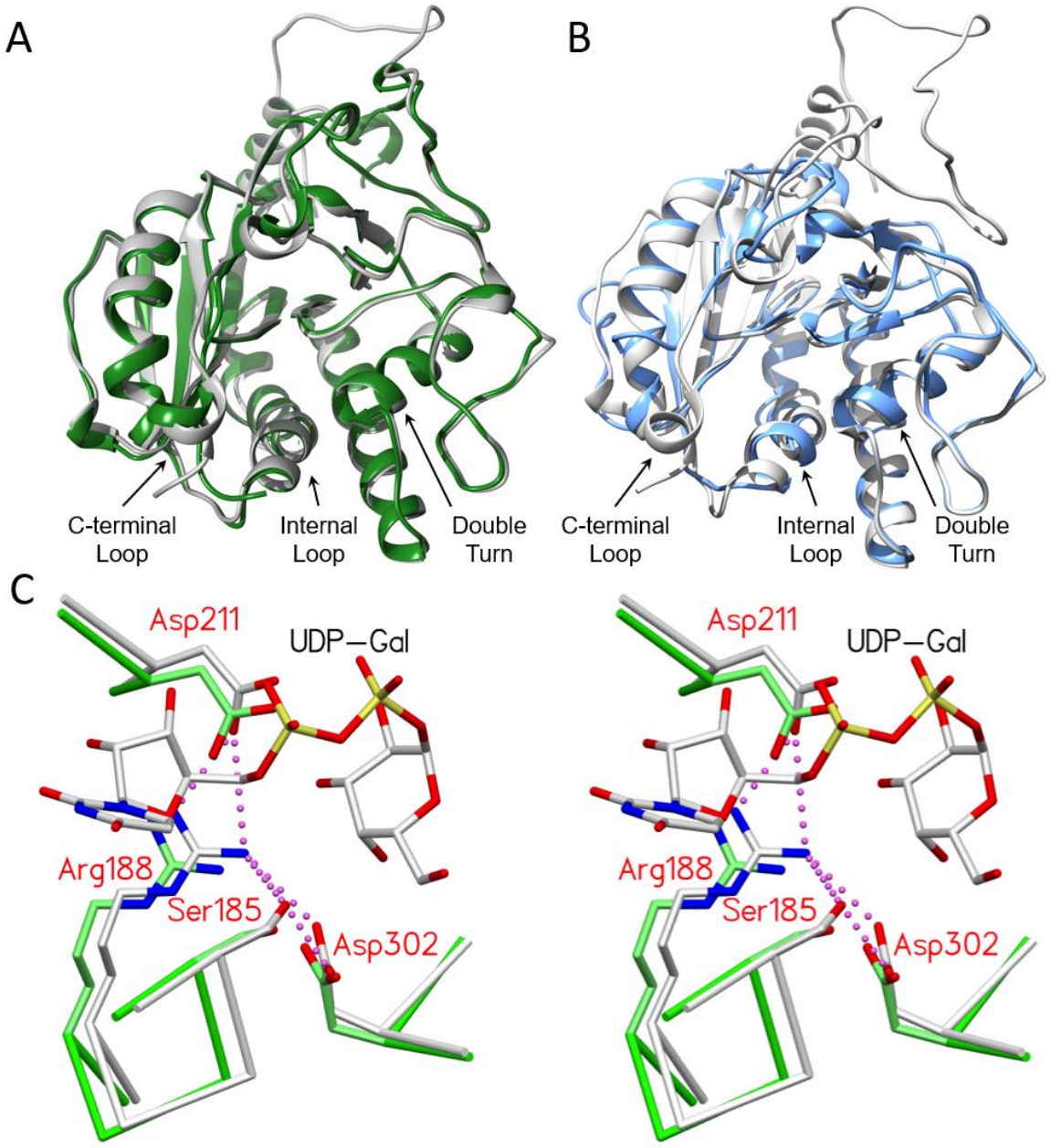
Chapter 6: Critical role of Asp302 in active site organization

6.1 Introduction to Chapter 6

As with Chapter 5, this chapter also covers my progress toward the second aim of this dissertation: to determine how donor substrate binds and achieves the constrained, high-energy conformation associated with catalysis. As discussed in the previous chapters, GTA and GTB are the highly homologous blood group A- and B-synthesizing enzymes, and they serve as convenient model for other GTs, since they are easily produced and well-characterized. While **Chapter 5** focused on substrate-binding in GTA/GTB by examining structures collected in complex with three distinct donors, the current chapter focuses on another set of experiments involving mutants of residues Arg188 and Asp302. In family 6 GTs, there are several conserved amino acids within the substrate-binding pockets that make critical hydrogen bond contacts to the donor sugar in multiple conformations, including Arg188 and Asp302 in GTA/GTB, as observed in previous structural studies (Alfaro, J.A., et al. 2008) as well as the structures discussed in **Chapter 5**. The family 6 GT, *Bos taurus* β-galactosyl α-1,3-galactosyltransferase (α3GT; EC 2.4.1.87), has active site residues Ser199, Arg202, Asp225, and Asp316, which correspond to GTA/GTB residues Ser185, Arg188, Asp211, and Asp302, respectively (Zhang, Y.N., et al. 2003). In both GTA/GTB and α3GT, these spatially equivalent residues form hydrogen bond interactions with the donor galactosyl moiety (**Figure 16A&C**).

Bacteroides ovatus BoGT6a (EC 2.4.1.40), a metal-independent GT with an NXN rather than the DXD motif typically observed for family 6 GTs (Tumbale, P. and Brew, K. 2009), possesses corresponding residues Thr70, Arg73, Asn95, and Asp191 that interact with the donor sugar in a manner similar to GTA/GTB and α3GT (Tumbale, P. and Brew, K. 2009, Thiyagarajan, N., et al. 2012, Pham, T.T.K., et al. 2014).

Figure 16. *Salt bridge network in GTA/GTB, α3GT, and BoGT6a*



(**A**) Overlap of α3GalT (green, PDB code 1K4V) with the catalytically competent closed state of AABB (grey, PDB code 2RJ7; 0.77 Å RMSD). (**B**) Overlap of BoGT6a (blue, PDB code 4AYJ) with AABB (grey, PDB code 2RJ7; 1.08 Å RMSD). (**C**) Superposition of the semi-open GTB (gray carbons) and α3GT (green carbons, PDB code 1O7Q) showing the equivalent positions of active site residues Ser185/199, Arg188/202, Asp211/225, and Asp302/316 (GTB side chains labeled) with UDP-Gal bound in the tucked under conformation (also called conformation #1). Pink dashed spheres represent salt bridge interactions. Reproduced from (Gagnon et al., 2018) with permission.

These residues participate in donor sugar binding and are conserved in CAZy family 6 GTs, yet prior to this investigation their roles were unknown. To uncover the function of two of these residues, Arg188 and Asp302, we generated mutant enzymes GTB/R188K, GTA/D302A, GTA/D302C, GTA/D302L, GTB/D302A, GTB/D302C, GTB/D302E, and GTB/D302L for structural and kinetic characterization.

Toward the published data reported in Chapter 6, my contributions include processing, solving, and refining X-ray diffraction data, analyzing and interpreting structural and kinetic data, performing a thorough literature review on which the main theme of the paper relied, producing figures and tables, and writing much of the paper.

6.2 Results

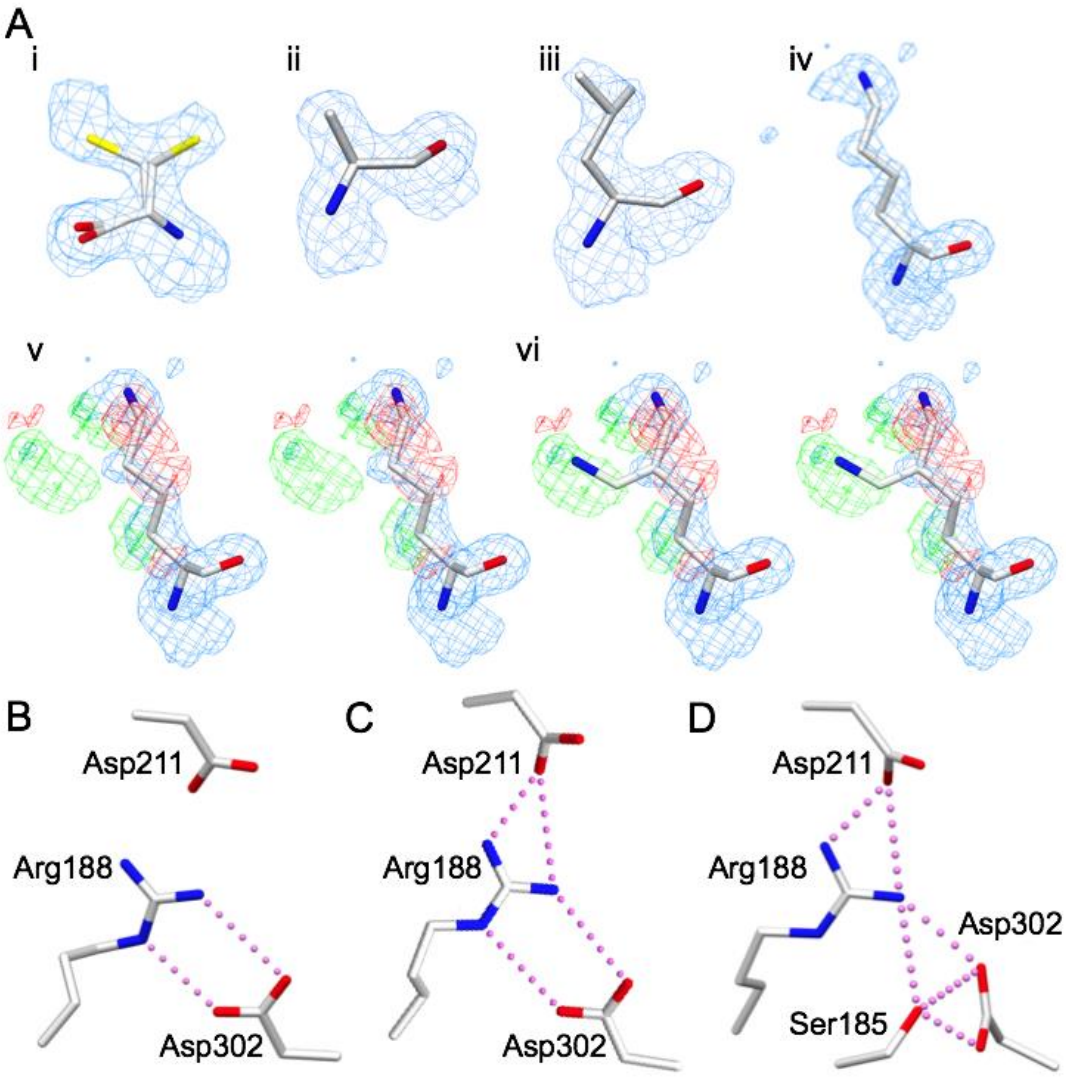
Kinetics

All kinetics were done by Prof. Monica Palcic's research groups and are described in Gagnon *et al.* (2018). Kinetic constants were determined for each purified mutant at a high concentration of the alternate substrate and are presented in **Table 8** alongside kinetic data for bovine α3GT (Zhang, Y., et al. 2003). All 188 and 302 mutants had lower k_{cat} values than the wild-type enzymes. For GTA/GTB R188K, D302A, and D302L mutants exhibited nearly complete loss of activity. GTB/D302A and GTB/R188K had elevated K_m values for UDP-Gal donor. The GTB/D302E mutant had the highest activity of the 302 mutants with a k_{cat} that was 47% that of wild-type GTB. The K_m for UDP-Gal donor was elevated with no change in the K_m for acceptor. GTA/D302C and GTB/D302C k_{cat} values were, respectively, 9.1%, and 9.0% of wild-type (**Table 8**). For both D302C mutants, the donor and acceptor K_m values were elevated.

Crystallography

Only GTA/D302C, GTB/D302A, GTB/D302C, GTB/D302L, and GTB/R188K mutant enzymes formed diffraction-quality crystals. Data collection and refinement results for these mutants are summarized in **Table 9**. Diffraction data were collected to a maximum resolution of 2.18-1.45 Å with a final R_{work} of 18.4-20.1% and R_{free} of 20.2-24.6%. Apart from the disorder between R188K Cε and Nζ atoms there is unambiguous electron density about the mutation sites (**Figure 17A**). In the case of the D302C mutants, Cys302 is disordered over two conformations (**Figure 17A)** While the internal loop and C-terminal tail were largely disordered, there was good electron density for the remaining main chain and side chain polypeptide. Internal loop organization varied across all structures (**Table 10**), while the C-terminal tail was uniformly disordered. The degree of loop disorder in the mutant structures is similar to unliganded wild-type GTB (PDB code 1LZ7). Superposition of the ordered residues of 1LZ7 Cα atoms with any of the mutant structures reported here gives <0.30 Å RMSD displacement values.

Figure 17. *Electron density for 188/302 mutants and salt bridge interactions*



(**A**) Electron density for mutated key active site residues (i) D302C, with the sulfhydryl group modeled in two positions each with 50% occupancy, (ii) D302A, (iii) D302L, (iv) R188K modeled in one conformation, (v) the stereo view of R188K modeled in one conformation with the difference map included, and (vi) the stereo view of R188K modeled in two conformations with the same maps as (v) demonstrating the flexibility of the Lys188 side chain. Electron density diagrams are $2F_O$-F_C maps contoured at 1.0 σ (*blue*) and F_O-F_C difference maps contoured at -3.00 (*red*) and 3.00 σ (*green*). (**B**,**C**,**D**) Salt bridge interactions among Ser185, Arg188, Asp211, and Asp302 in (**B**) unliganded GTB in the open state (PDB code 2RIT), (**C**) GTB in the liganded intermediate state (conformation #2, PDB code 5C1L), (**D**) chimeric enzyme AABB in the liganded closed state (conformation #1, PDB code 2RJ7). Atoms are colored by element with oxygen *red*, nitrogen *blue*, and sulfur *yellow*. Reproduced from (Gagnon et al., 2018) with permission.

Table 8. Kinetic constants for GTA/GTB wild-type and Asp302 & Arg188 mutants and comparison to chimeric enzyme AABB & bovine α3GT

Enzyme / Donor	**GTA / UDP-GalNAc**			**GTB / UDP-Gal**			**AABB / UDP-GalNAc**			**AABB / UDP-Gal**			**α3GT / UDP-Gal**		
	K_{Acc} (μM)[a]	K_{Don} (μM)[a]	k_{cat} (s^{-1})	K_{Acc} (μM)[a]	K_{Don} (μM)[a]	k_{cat} (s^{-1})	K_{Acc} (μM)[b]	K_{Don} (μM)[b]	k_{cat} (s^{-1})	K_{Acc} (μM)[b]	K_{Don} (μM)[b]	k_{cat} (s^{-1})	K_{Acc} (mM)	K_{Don} (μM)	k_{cat} (s^{-1})
Wild-type	9.9	8.7	17.5	88	27	5.1	13	37	0.60	1.6	1.8	2.2	19.9[c]	430[c] 12.5[d]	6.4[c] 1.3[d]
D302A/D316A	LA[f]	LA	LA	117±6[g]	170±40	0.002± 0.0001	-	-	-	-	-	-	-	70[d]	0.0013[d]
D302E/D316E	-	-	-	100±20	112±18	2.4±0.5	-	-	-	-	-	-	6.4[e]	160[e]	2.48[e]
D316N	-	-	-	-	-	-	-	-	-	-	-	-		inactive	
D302C	45±5	22±2	1.6±0.4	332±24	53±4	0.46±0.02	-	-	-	-	-	-	-	-	-
D302L	LA	LA	LA	LA	LA	LA	-	-	-	-	-	-	-	-	-
R188K	-	-	-	101±20	140± 30	0.0006± 0.0006	-	-	-	-	-	-	-	-	-
R188S	-	-	-	-	-	LA[h]	-	-	-	-	-	-	-	-	-

[a] The K_{Acc} and K_{Don} Michaelis-Menten constants are the K_m for acceptor and donor, respectively
[b] As reported in Alfaro *et al.* (2008).
[c] As reported in Zhang *et al.* (2003).
[d] As reported in Monegal *et al*. (2006).
[e] As reported in Tumbale *et al*. (2008).
[f] LA = low activity enzymes with activities measured as less than 1/10 000 of wild-type enzyme
[g] K_m errors are from curve fitting while for k_{cat} they are the average of k_{cats} from donor and acceptor kinetics.
[h] As reported in Yazer & Palcic (2005)
*Boxes contain dashes when the specified mutation does not apply to a given enzyme. For example, Asp316 is a residue belonging to α3GT (not GTA/GTB or AABB), and so the kinetics for the D316N mutant are displayed appropriately for α3GT but not GTA, GTB, or AABB.
Reproduced from (Gagnon et al. 2018) with permission.

Table 9. Data collection and refinement statistics for GTA/GTB 302 and 188 mutants

All crystals belong to space group $C222_1$ with approximate unit cell dimensions of a=52.6, b=150.2, and c=78.9 Å.

	GTA/D302C	**GTB/D302A**	**GTB/D302C**	**GTB/D302L**	**GTB/R188K**
Resolution (Å)	20 – 1.54	20 – 1.45	20 – 2.18	20 – 1.69	20 – 1.45
R_{sym} (%)[a,b]	3.8 (32.0)	2.9 (30.3)	8.8 (29.4)	4.6 (32.0)	3.1 (28.3)
Completeness (%)[b]	98.7 (100.0)	99.5 (100.0)	95.3 (98.8)	98.3 (100.0)	96.6 (96.2)
Unique Reflections	46171	55649	17135	34748	53903
Multiplicity	4.44	4.28	4.68	4.71	3.84
$\langle I/\sigma(I) \rangle$	17.2	21.1	9.1	15.1	21
Refinement					
Reflections used in refinement	43837	52822	16261	32996	51167
R_{free} reflections	2329	2825	874	1749	2733
Wilson B-factor (Å^2)	20.4	21.7	32.1	22.1	20.4
R_{work} (%)[c]	19.5	18.6	20.1	18.4	18.9
R_{free} (%)[c,d]	21.9	20.2	24.6	22.2	20.5
No. non-hydrogen protein atoms	2247	2154	2139	2229	2269
No. water molecules	270	266	122	229	218
r.m.s. angle (°)[e]	1.41	1.43	1.33	1.44	1.47
r.m.s. bond (Å)[e]	0.0092	0.0093	0.0078	0.0093	0.0098
Ramachandran outliers (%)	0	0	0	0	0
PDB code	6BJI	6BJJ	6BJK	6BJL	6BJM

[a] $R_{sym} = \sum_{hkl}\sum_i |I_{hkl,\,i} - [I_{hkl}]| \,/\, \sum_{hkl}\sum_i I_{hkl,\,i}$ where $[I_{hkl}]$ is the average of Friedel-related observations (i) of a unique reflection (hkl)

[b] Values in parentheses represent highest resolution shell.

[c] R-work, $\Sigma\, ||F_o| - |F_c|| \,/\, \Sigma\, |F_o|$.

[d] 5% of reflections were omitted for R-free calculations.

[e] r.m.s. root-mean-square.

Reproduced from (Gagnon et al. 2018) with permission.

Table 10. Internal loop disorder in wild-type GTB and Arg188/Asp302 mutant structures.

All residues in these internal loops lie in conformations corresponding to the 'open' state described in Alfaro et al. 2008.

Enzyme	Internal loop 175 180 185 190
GTB[1]	VGayk rWQDV SMRRM EMISD
GTA/D302C	VRayk rwqDV SMRRM EMISD
GTB/D302A	VGayk rwqdv smrrm emisd
GTB/D302C	Vgayk rwqdv smrrm emiSD
GTB/D302L	Vgayk rwqdV SMRRM EMISD
GTB/R188K	VGayk rwqdV SMRKM EMISD

[1] PDB code: 2RIT (Alfaro, J.A., et al. 2008)

One-letter amino acid codes with upper case black letters correspond to unambiguous electron density for main chain and side chain atoms, blue corresponds to unambiguous electron density for main chain atoms only; red letters correspond to weak or ambiguous electron density for main chain and side chain atoms. Residues with one-letter amino acid codes in lower case have not been included in the refined models.

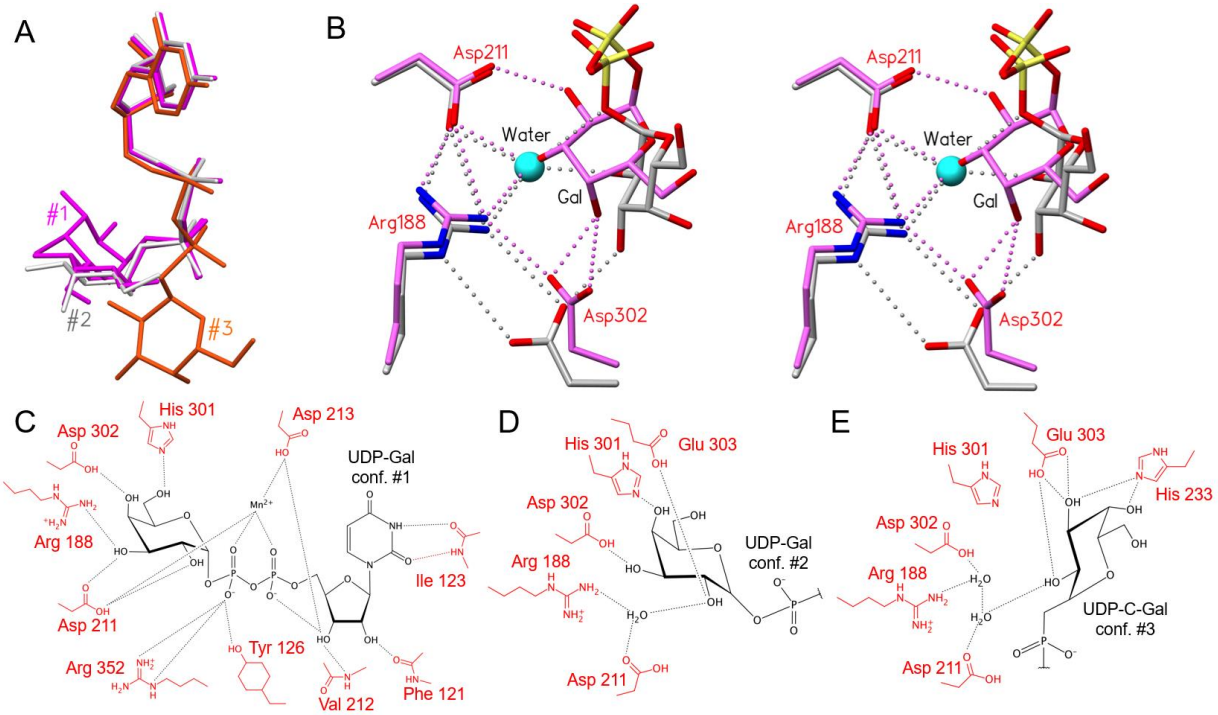
Reproduced from (Gagnon et al. 2018) with permission.

6.3 Discussion & conclusions

A salt bridge network drives formation of the catalytically competent state

As outlined in **Chapter** 5, previous GTA/GTB structures demonstrate that binding of the donor sugar *vi*a Arg188, Asp211, and Asp302 is crucial to formation of the catalytically active closed conformation (also called conformation #1) (Alfaro, J.A., et al. 2008, Schuman, B., et al. 2010). These residues form a hydrogen bond network, including salt bridges, with each other and with donor substrate (**Table 11, Figures 17B-D & 18**). In wild-type GTA/GTB, Asp302 forms a bidentate salt bridge to Arg188 both in unliganded open structures (**Figure 17B**) and in structures with bound donor in conformation #2, the alternate UDP-Gal conformation observed in Gagnon *et al.* (2015) and discussed at length in **Chapter 5** (**Figures 17C & 18D**). In conformation #2, Asp302 also forms a hydrogen bond to donor Gal-O3 and a water molecule bridges an interaction between Arg188 and donor Gal-O2. When the donor shifts into the tucked-under conformation (**Figure 17D**) and enzyme is in the closed state, the Asp302 side chain loses the Nε hydrogen bond with Arg188 and switches from binding Gal-O3 to Gal-O4 (**Table 11**; **Figure 17C&D; Figure 18B&C**) (Alfaro, J.A., et al. 2008, Schuman, B., et al. 2010, Gagnon, S.M., et al. 2015). During the shift from conformation #2 to #1, the donor hydroxyls move by as much as 4.0 Å, and both Arg188 and Asp211 side chains make direct hydrogen bonds to Gal-O3 (**Figure 18B&C**) (Alfaro, J.A., et al. 2008, Gagnon, S.M., et al. 2015). The donor Gal-O6 also is displaced by 2.7 Å, preventing a steric clash from occurring between internal loop residue Trp181 and promoting a stabilizing interaction between Trp181 and Arg352 of the C-terminal tail. When UDP-Gal shifts to conformation #1, the internal loop becomes more organized, and Arg188 makes anchoring contacts with the donor sugar while participating in salt bridge interactions with Asp211 and Asp302.

Figure 18. *GTA/GTB donors adopt a series of conformations facilitated by Arg188, Asp211, and Asp302*



(**A**) Overlap of three donor conformations with tucked under conformation #1 in *magenta* (PDB code 2RJ7), intermediate conformation #2 in *light grey* (PDB code 5C1L) and "extended" conformation #3 in *orange* (PDB code 5C3D) (Alfaro, J.A., et al. 2008, Gagnon, S.M., et al. 2015). (**B**) Stereo view of the superposition of AABB (PDB code 2RJ7; conformation #1; *magenta*) and GTB (PDB code 5C1L; conformation #2; carbon atoms colored *light grey*) crystal structures depicting salt bridge and hydrogen bond interactions among Arg188, Asp211, Asp302, a water molecule (*cyan* sphere), and donor galactose. Interactions include a single/bidentate salt bridge between Arg188 and Asp302. In donor conformation #2 a water molecule occupies the same position as donor Gal-O3 in conformation #1. Except where otherwise indicated, atoms are colored by element with oxygen *red*, nitrogen *blue*, and phosphorous *yellow*. (**C**,**D**,**E**) Schematic depictions of hydrogen bond interactions (black lines) between (**C**) UDP-Gal in tucked under conformation #1 and chimeric enzyme AABB (PDB code 2RJ7), (**D**) the substrate Gal moiety in the intermediate conformation #2 (PDB code 5C1L; GTB) and (**E**) in the extended conformation #3 (PDB code 5C3D; ABBB). Key active site resides are labeled in *red* and the donor sugar in *black*. Reproduced from (Gagnon et al., 2018) with permission.

Glycosyltransfer is thought to follow internal loop and C-terminal tail organization about the substrates to achieve the closed state (Alfaro, J.A., et al. 2008), a common theme in GTs (Rini, J., et al. 2009). The Arg188, Asp211, and Asp302 salt bridges may not only mediate the stepwise shift of donor from the predominant solution conformation to the tucked-under

catalytically competent conformation but may also drive enzyme closure. As well, it is clear that the formation of the closed state itself is promoted by the new and reordered active site interactions that form as the donor binds.

One of the salt bridge residues, Asp302, resides in the active site within the conformationally strained double-turn motif (see **Chapter 4**), which is conserved in family 6 enzymes α3GT, BoGT6a, and GTA/GTB and family 8 α-1,4-galactosyltransferase LgtC, and which has a known catalytic role in GTA/GTB (**Chapter 4**) (Tumbale, P., et al. 2008, Blackler, R.J., et al. 2017), where it makes hydrogen bond interactions with the donor sugar moiety. In GTA/GTB the hydrogen bond/salt bridge network residues link the double-turn with the flexible internal loop region, and there is a corresponding set of interactions in unliganded and liganded structures of homologous CAZy family 6 GTs, α3GT and BoGT6a (**Table 11**).

In the unliganded α3GT structure (PDB code: 5NR9) (Albesa-Jove, D., et al. 2017) only one of the two molecules in the asymmetric unit has comparable salt bridge interactions: here, glycerol is bound in the position that would normally be occupied by the donor sugar moiety. In the second molecule in the asymmetric unit, where neither glycerol nor donor is bound, there is no corresponding salt bridge network.

Family 6 GTs α3GT and BoGT6a share only ~43% and ~33% sequence identity with GTA/GTB, respectively, yet structurally they are highly similar as shown *via* overlap with the AABB chimera. AABB/α3GT residues Ser185/Ser199, Arg188/Arg73, Asp211/Asn95 and Asp302/Asp191 are in near-identical orientations (**Figure 9A&C, Table 11&12**). In the same manner, BoGT6a achieves the closed enzyme state and there are interactions among corresponding residues Thr70, Arg73, Asn95, Asp191 (**Figure 9B**).

Salt-bridge network conservation in GT-A fold glycosyltransferases

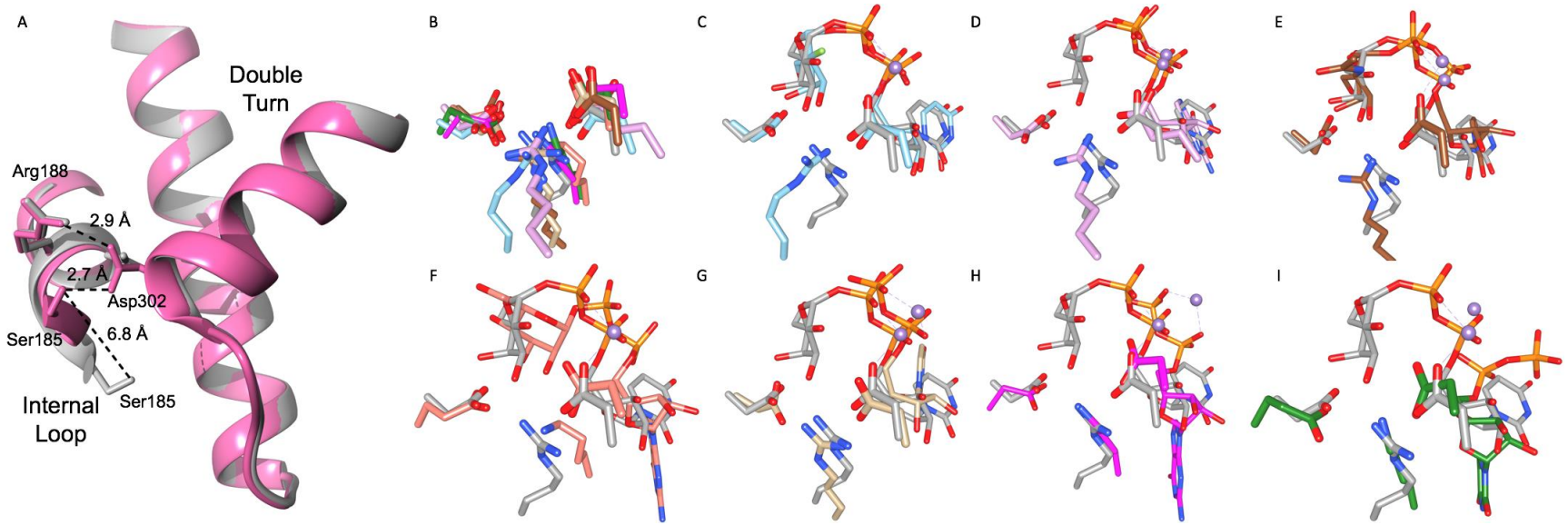
The GTA/GTB Arg188, Asp211, and Asp302 salt bridge network may represent a conserved feature among other retaining GT-A fold-type GTs. Overlap of chimeric AABB (PDB code: 2RJ7) with published retaining GT-A fold structures shows that each family contains at least one representative structure with equivalent interacting salt bridge residues that participate in donor binding (**Table 13**, **Figure 19**). All published structures within retaining GT-A fold CAZy family 7, 13, 15, 27, 55, 64, 78, and 81 show salt bridge conservation. Conservation is not total: there are some exceptions within families 2, 8, and 43, where either there is insufficient structural data to allow for comparison or else the active site suggests the salt bridge is absent. For some family 2 GTs there is ambiguous electron density or a lack of intact donor-bound complexes as with *Pyrococcus* furiosus PF0058, *Bacteroides fragilis* BF2801, and *Oryza sativa* Os07g0208500. Family 2 SpsA does not display salt bridge conservation, however the cognate substrate for this enzyme is unknown, and as a consequence no natural donor-complexed structures have been published. The structural data for family 8 GTs *Anaerococcus prevotii* Apre_0416, *Streptococcus gordonii* Gtf2, and *Caenorhabditis elegans* glycogenin α-glucosyltransferase is insufficient to evaluate salt bridge residue conservation unambiguously, though the remaining family 8 GTs are conserved. The final exception to salt bridge conservation is family 43 GT β-1,3-glucuronosyltransferase 3, where Leu280 occupies the position corresponding to GTA/GTB residue Asp302. This trend also extends to inverting GT-A fold GTs, where residues equivalent to GTA/GTB Arg188, Asp211, and Asp302 are present in representative structures of each family (**Table 13**).

In contrast to Arg188, Asp211, and Asp302, Ser185 is not strictly conserved, and its presence may depend on internal loop length. For example, the relatively shorter internal loop of

retaining GT-A fold family 27 polypeptide GalNAc-transferase 2 (GalNAc-T2) may require fewer contacts to achieve the ordered, closed catalytic state, which means that a Leu204 may suffice in the position occupied by a hydrogen-bonding Ser185 in GTA/GTB.

Internal loop residue Ser185, as with Arg188, undergoes a major shift as the donor sugar transitions from conformation #2 into tucked-under conformation #1 (see **Chapter 5**) (Gagnon, S.M., et al. 2015). The side chain of Ser185 moves over 6 Å and forms a hydrogen bond interaction with double-turn residue Asp302, stabilizing the closed enzyme state (**Figures 17B-D & 18C-E**). Underscoring the importance of Ser185 in catalytic turnover, Nakahara *et al.* (2006) observed that GTB S185N and S185C mutants has decreased k_{cat} values of 0.52 s^{-1} and 0.092 s^{-1}, respectively, relative to their reported value for wild-type GTB, 5.2 s^{-1} (Nakahara, T., et al. 2006).

***Figure 19**. Salt bridge conservation in GT-A fold-type GTs*



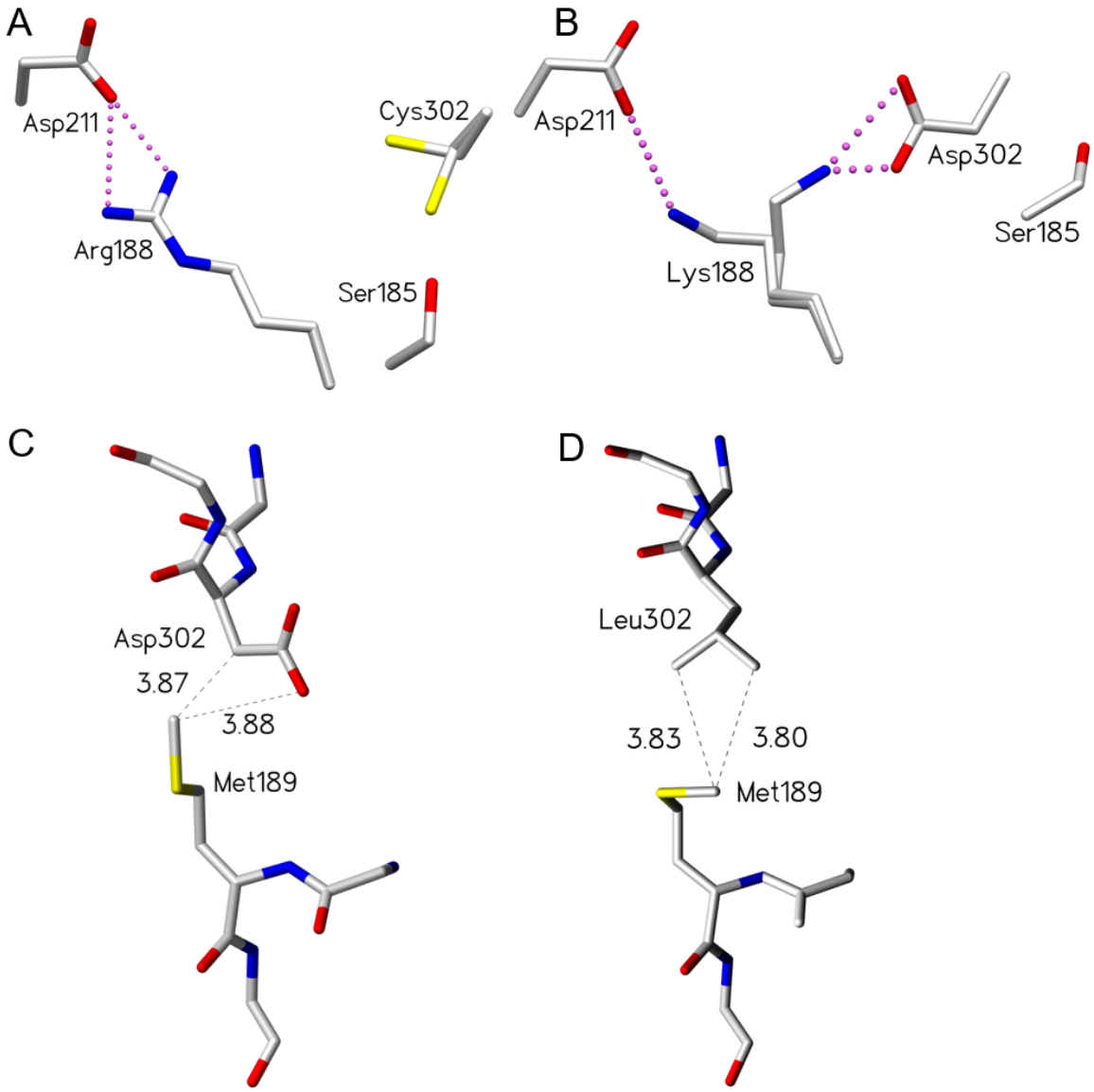
The salt bridge interactions formed in the closed enzyme state represent a structurally conserved feature found in representative structures from each retaining GT-A fold family. (**A**) Overlap of salt bridge residues Arg188, Asp211, and Asp302 from chimeric AABB (*grey*, PDB code 2RJ7) with corresponding residues (listed in Table VIII) in representative structures from each retaining GT-A fold family using the following color scheme: GT-8 in *sky blue* (PDB code 1GA8), GT-15 in *plum* (PDB code 1SP4), GT-27 in *brown* (PDB code 4D0T), GT-55 in *salmon* (PDB code 2ZU8), GT-64 in *tan* (PDB code 1ON8), GT-78 in deep *pink* (PDB code 2Y4L), and GT-81 in *forest green* (PDB code 4DEC). In (**B**)-(**H**) AABB again is shown in *grey* (PDB code 2RJ7) and is separately overlapped with each GT-A fold family representative. Overlaps are between AABB and a representative structure (listed in Table VIII) of (**B**) GT-8, (**C**) GT-15, (**D**) GT-27, (**E**) GT-55, (**F**) GT-64, (**G**) GT-78, and (**H**) GT-81. Atoms are colored by element with oxygen red, nitrogen *blue*, fluorine *light green*, carbon as indicated above, phosphorous *orange*, and manganese *medium-purple*. Color naming is consistent with the program USCF Chimera. (**I**) Overlap of AABB (*pink*, PDB code 2RJ7) with the semi-open GTB (*grey*, PDB code 5C1L) showing the >6Å shift of internal loop residue Ser185 that occurs upon substrate binding and salt bridge reorganization. Reproduced from (Gagnon et al., 2018) with permission.

Loop ordering and kinetics of 188/302 mutants

Among all mutant structures, only GTA/D302C and GTB/R188K showed electron density corresponding to the residue 188 side chain (**Table 10**). For GTB/D302A and GTBD302L this is unsurprising given that the 302 side chains of these mutants cannot form hydrogen bonds to residue 188. They cannot take part in the salt bridge network or stabilize the closed conformation, which is consistent with their near-zero activities (**Table 8**). Though for both GTA/D302C and GTB/D302C, residue 302 is disordered over two conformations (**Figure 17A**) and sulfur is a poor hydrogen bond acceptor, the Cys side chain may still be able to interact with the donor and participate in the hydrogen bond network interactions that stabilize the closed conformation. This may explain the retention of ~10% of wild-type activity for these mutants (**Table 8**).

Arg/Gly176, the first of the four "critical" GTA/GTB residues, has been associated with internal loop stability. Compared to Gly176, Arg176 imparts greater loop order due to more restricted main chain dihedral angles (Patenaude, S.I., et al. 2002, Lee, H.J., et al. 2005, Letts, J.A., et al. 2006, Alfaro, J.A., et al. 2008, Johal, A.R., et al. 2012). This trend is evident in the D302C mutant structures, where GTA/D302C displays unambiguous electron density for eleven additional internal loop residues over GTB/D302C (**Table 10**). An exception to the trend is the GTB/D302L mutant. Leu302, which cannot form a salt bridge interaction with Asp211. Despite this, the D302L mutant manages to achieve internal loop organization in the open position *via* van der Waals contacts formed between Leu302 and Met189, as observed in the crystal structure (**Figure 20**).

Figure 20. *Stabilizing van der Waals contacts in 188/302 mutants*



Certain mutants of residue 302 and 188 compensate for the loss of critical hydrogen bonds within the salt bridge network by establishing alternative interactions. D302C and R188K mutants maintain the salt bridge network by adopting multiple conformations, while the D302L mutant maintains partial loop order *via* van der Waal's contacts. (**A**) Cys302 side chain observed in two conformations in unliganded GTA/D302C. (**B**) Lys188 side chain observed in two conformations in unliganded GTB/R188K. Stereo view of (**C**) wild-type GTB residues Asp302 and Met189 (PDB code 2RIT) and (**D**) GTB/D302L residues Leu302 and Met189 with van der Waal's interactions shown as dashed lines (distances in Ångstrom) between the Leu302 side chain carbon atoms and the Met189 side chain methyl group. The dashed lines between the Met189 methyl group and the Asp302 carbonyl group represent a measured distance, not a van der Waal's interaction. Atoms are colored by element with carbon *gray*, oxygen *red*, nitrogen *blue*, and sulfur *yellow*. Salt bridge interactions are depicted as dashed spheres (*pink*).

Among GTB mutant structures, GTB/R188K has the greatest internal loop order and partially maintains the salt bridge network (**Table 11**), albeit with a pattern of active site hydrogen bonds distinct from the wild-type. Compared to Lys, the multivalency and higher pKa of the Arg side chain guanidinium group (Sokolowski, T., et al. 1998) permit stronger and more numerous electrostatic interactions. This is consistent with the kinetics of GTB/R188K, which has higher donor and acceptor K_m values relative to wild-type and displays a reduction in k_{cat} of nearly four orders of magnitude (**Table 8**). The structure of GTB/R188K suggests that Lys188 may form a salt bridge with both Asp211 and Asp302 (**Figure 17**) through its side-chain flexibility, which could allow it to be disordered over two conformations. In one conformation, Lys188 interacts with Asp211, while in the other it interacts with Asp302. This represents a weaker set of interactions than seen for the wild-type, indicating that both the length and multivalent nature of the Arg188 side chain are catalytically important.

Acidic amino acid residue 302 is catalytically important

Kinetic and structural data demonstrate that Asp302 strongly influences enzyme catalysis (**Table 8**). Mutation of Asp302 to non-polar alanine and leucine nearly abolishes enzyme activity. In contrast, the more conservative mutation to glutamate preserves almost 50% of wild-type activity, and mutation to cysteine preserves ~10% (**Table 8**) – in the case of these D302C mutants, at pH 7.4, which is the pH used for these kinetic studies, a comparable percentage (~10%) of cysteine side chains would be expected to be in the ionized state. These results support the notion that a negatively-charged residue 302 is catalytically important. Our data is also consistent with published findings for α3GT, which retains 40% of wild-type activity when residue Asp316 (Asp302 equivalent) is mutated to glutamate (Tumbale, P., et al. 2008) but loses

nearly all detectable activity when mutated to alanine (Monegal, A. and Planas, A. 2006) or asparagine (Tumbale, P., et al. 2008) (**Table 8**).

Conclusions

Conserved active site residues Arg188 and Asp302 are catalytically vital in GTA/GTB, and disruption of their hydrogen bond network *via* mutagenesis can dramatically decrease k_{cat}. Taken together, these data suggest that Arg188 and Asp302, in concert with Ser185 and Asp211, are critical to the generation of the catalytic state. Asp302 and Arg188 mutant enzymes are unable to achieve this state due to a weakened or abolished hydrogen bond network, where Ser185 cannot anchor the internal loop to the double-turn motif, impeding loop organization and by extension, catalysis.

Table 11. Hydrogen bond interactions between UDP-Gal donor substrate and key residues of GTA/GTB, α3GalT, and BoGT6a

Hydrogen bond distances between select GTA/GTB residues and UDP-Gal donor is in the catalytically competent conformation.

GTA/GTB residue[1]	**UDP-Gal**	**α3GalT residue[2]**	**UDP-Gal**	**BoGT6a residue[3]**	**UDP-Gal**
R188 [NH1]	O3’	R202 [NH1]	O3’	R73 [NH2[	O3’
R188 [NH2]	None	R202 [NH2]	None	R73 [NH1]	None
D211 [OD1]	O3’	D225 [OD1]	O3’	N95 [ND2]	O3’
D211 [OD2]	O2’	D225 [OD2]	O2’	N95 [ND2]	O2’
H301 [ND1]	O6’	H315 [ND1]	O6’	H190 [NE2]	O6’
D302 [OD1]	O4’	D316 [OD2]	O4’	D191 [OD1]	O4’
D302 [OD2]	None	D316 [OD1]	None	D191 [OD2]	None
E303	Acceptor	E317	O4’	E192	Acceptor

[1] Liganded AABB, PDB code: 2RJ7 (Alfaro, J.A., et al. 2008)
[2] Liganded α3GalT, PDB code: 1O7Q (Zhang, Y., et al. 2003)
[3] Acceptor-bound BoGT6a, PDB code: 4AYJ (Thiyagarajan, N., et al. 2012), superposed with liganded AABB, PDB code: 2RJ7, where UDP-Gal substrate is oriented as described in Thiyagarajan *et al.* (Thiyagarajan, N., et al. 2012)
Reproduced from (Gagnon et al. 2018) with permission.

Table 12. Salt bridge interactions in GTB, AABB, α3GalT, and BoGT6a

Salt bridge interactions in unliganded and liganded GTB with UDP-Gal in alternate conformation #2, and liganded AABB with UDP-Gal in the catalytically competent conformation (#1) (black). Interactions are shown alongside those observed for α3GalT (blue) and BoGT6a (red) with relevant residues grouped with their GTA/GTB equivalents.

Atoms from conserved salt-bridge residues	[NH1] from R188, R202, R73	[NH2] from R188, R202, R73	[NE] from R188, R202, R73	[O] from S185, S199, T70
Partnering atom(s) from GTB unliganded[1]		D302 [OD1], 3.4 Å D302 [OD2], 3.5 Å	D302 [OD1], 3.1 Å	
Partnering atom(s) from GTB + UDP-Gal #2[2]	D211 [OD1], 3.6 Å D302 [OD1], 3.5 Å D302 [OD2], 3.8 Å	D211 [OD1], 2.8 Å	D302 [OD2], 3.7 Å	
Partnering atom(s) from GTB + UDP-Gal #1[3]	D211 [OD1], 3.4 Å D302 [OD1], 2.9 Å	D211 [OD1], 2.8 Å	S185 [O], 3.0 Å	D302 [OD1], 4.0 Å D302 [OD2], 3.4 Å
Partnering atom(s) from α3GalT unliganded[4]	D225 [OD1], 2.6 Å	S199 [O], 4.0 Å D225 [OD1], 3.1 Å D316 [OD2], 3.6 Å	S199 [O], 3.1 Å	R202 [NH2], 4.0 Å R202 [NE], 3.1 Å D316 [OD2], 3.7 Å
Partnering atom(s) from α3GalT + UDP + GalNAc[5]	D225 [OD1], 2.8 Å	S199 [O], 4.0 Å D225 [OD1], 3.2 Å D316 [OD2], 2.9 Å	S199 [O], 3.0 Å	R202 [NH2], 4.0 Å R202 [NE], 3.0 Å D316 [OD2], 3.7 Å
Partnering atom(s) from BoGT6a unliganded[6]	D191 [OD1], 3.0 Å	N95 [OD1], 3.9 Å D191 [OD1], 3.9 Å		
Partnering atom(s) from BoGT6a + FAL[7]	N95 [OD1], 4.0 Å D191 [OD1], 3.2 Å D191 [OD2], 3.5 Å	N95 [OD1], 2.6 Å D191 [OD2], 3.9 Å	T70 [O], 3.8 Å D191 [OD2], 3.4 Å	R73 [NE], 3.8 Å D191 [OD1], 3.2 Å

[1] Unliganded GTB, PDB code: 2RIT (Alfaro, J.A., et al. 2008)
[2] Liganded GTB, PDB code: 5C1L (Gagnon, S.M., et al. 2015)
[3] Liganded AABB, PDB code: 2RJ7 (Alfaro, J.A., et al. 2008)
[4] Unliganded α3GalT, PDB code: 5NR9 (Albesa-Jove, D., et al. 2017)
[5] Liganded α3GalT, PDB code: 1O7Q (Zhang, Y., et al. 2003)
[6] Unliganded BoGT6a, PDB code: 4AYL (Thiyagarajan, N., et al. 2012)
[7] Acceptor-bound BoGT6a, PDB code: 4AYJ (Thiyagarajan, N., et al. 2012), superposed with liganded AABB, PDB code: 2RJ7, where UDP-Gal is oriented as described in (Thiyagarajan, N., et al. 2012)
Reproduced from (Gagnon et al. 2018) with permission.

Table 13. Representative PDB structures of GT-A fold GTs with salt bridge residues corresponding to GTA/GTB Arg188, Asp211, Asp302

Reaction stereochemistry	Family	Example enzyme	Donor analogue	Acceptor analogue	PDB	Arg188 equivalent	Asp211 equivalent	Asp302 equivalent
Retaining	GT-8	α-1,4-galactosyltransferase, LgtC[1]	UDP 2-deoxy-2-fluoro-galactose	4'deoxylactose	1GA8	Arg86	Asp103	Asp188
	GT-15	α1,2-mannosyltransferase, Kre2p/Mnt1p[2]	GDP	O1-methyl-mannose	1S4P	Arg245*	Glu247	Asp361
	GT-27	polypeptide GalNAc-transferase 2, GalNAc-T2[3]	UDP-N-acetylgalactosamine	EA2 peptide	4D0T	Arg208	Asp224	Glu334
	GT-55	mannosyl-3-phosphoglycerate synthase[4]	GDP-α-D-mannose	-	2ZU8	Lys145	Asp168	Glu271
	GT-64	α-1,4-*N*-acetylhexosaminyltransferase, EXTL2[5]	UDP	[glucuronic acid]β1–3[galactose]β1-*O*-naphthalenemethanol	1ON8	Arg135	Asp151	Asp245
	GT-78	mannosylglycerate synthase, MGS[6]	GDP	Malonate ion	2Y4L	Lys76	Asp100	Asp192
	GT-81	glucosyl-3-phosphoglycerate synthase[7]	UDP	3-phosphoglyceric acid	4DEC	Lys114	Asp134	Glu232
Inverting	GT-2	Teichoic acid β-glycosyltransferase[8]	UDP-N-acetylgalactosamine	-	5TZJ	Arg75	Asp91	Glu177
	GT-7	β1–4-galactosyltransferase-I, β4Gal-T1[9]	UDP	[N-acetylglucosamine]β1–6[galactose]β1-4[glucose]	4EEA	Arg224	Asp248	Glu313
	GT-13	O-linked mannose β1,2-N-acetylglucosaminyltransferase 1, POMGnT1[10]	UDP	*O*-mannosylated peptide	5GGI	Tyr372	Glu393	Trp475
	GT-43	Glucuronyltransferase, GlcAT-P[11]	UDP	*N*-acetyllactosamine	1V84	Arg170	Asp195	Gln283

[1] Persson *et al.* (Persson, K., et al. 2001), [2] Lobsanov *et al.* (Lobsanov, Y.D., et al. 2004), [3] Lira-Navarrete *et al.* (Lira-Navarrete, E., et al. 2014), [4] Kawamura *et al.* to be published, [5] Pedersen *et al.* (Pedersen, L.C., et al. 2003), [6] Nielsen *et al.* (Nielsen, M.M., et al. 2011), [7] Urresti *et al.* (Urresti, S., et al. 2012), [8] Sobhanifar *et al.* (Sobhanifar, S., et al. 2016), [9] Ramakrishnan *et al.* (Ramakrishnan, B., et al. 2012),[10] Kuwabara *et al.* (Kuwabara, N., et al. 2016),[11] Kakuda *et al.* (Kakuda, S., et al. 2004), * Unlike GTA/GTB Arg188 and the corresponding residues of the other GT-A fold GTs, Arg160 in this family 15 enzyme does not reside within an ordered alpha helix, though it appears to play the same or similar role in salt bridge formation

Chapter 7: Trisaccharide product formation and release

7.1 Introduction to Chapter 7

This chapter addresses the third and final aim of this dissertation: to characterize trisaccharide product formation and release from the active site following glycosyltransfer. GTA/GTB, the enzymes that catalyze the final step of ABO(H) blood group A and B antigen synthesis, were first characterized structurally in 2002 (Patenaude, S.I., et al. 2002). Since then, they have been examined in complex with nucleotide donor, acceptor, and various substrate analogues *via* X-ray diffraction crystallography. As reviewed in **Chapter 5** of this dissertation, our recent findings have demonstrated that the donor sugar undergoes a series of conformational shifts concurrent with internal loop and C-terminal tail ordering (Gagnon, S.M., et al. 2015). However, many aspects of the glycosyltransfer reaction are still unknown, including the release mechanism for trisaccharide and UDP reaction products.

There is a need for this mechanistic information given the scarcity of GT structures collected in complex with their reaction products. Any forthcoming insights from such structures could inform the development of product analogues for use as GT inhibitors. Recently, BshA, a CAZy family 4, ret-GT with a GT-B fold-type, was structurally characterized in complex with its reaction product, *N*-acetylglucosaminyl-malate (Winchell, K.R., et al. 2016), and based on the authors' interpretation of the data, this provided support for the proposed S_Ni mechanism (Winchell, K.R., et al. 2016). Given the plasticity of GTA/GTB and GTs in general, recently reviewed by Albesa-Jové and Guerin (Albesa-Jove, D. and Guerin, M.E. 2016), it is entirely possible that retaining GTs belonging to other families and fold-types exhibit different mechanistic behavior, and characterization of their product-bound states could prove crucial.

To advance understanding of product release in these enzymes, we crystallized GTA and GTB with A and B trisaccharides analogues, respectively, under native and mercury derivatized conditions.

Toward the published data reported in Chapter 7, my contributions include processing, solving, and refining X-ray diffraction data, analyzing and interpreting structural data, synthesizing the X-ray diffraction and STD NMR results, drawing conclusions based on literature knowledge, developing the story of the paper, producing figures and tables, and writing much of the paper.

7.2 Results

Trisaccharide antigen binding epitopes

Data collection and refinement statistics for native and derivative enzyme complexes, four structures in total, are shown in **Table 14**. GTA and GTB were both observed bound to the trisaccharide product analogues, α-L-Fuc*p*-(1→2)[α-D-GalNAc*p*-(1→3)]-β-D-Gal*p*-O$(CH_2)_7CH_3$ and α-L-Fuc*p*-(1→2)[α-D-Gal*p*-(1→3)]-β-D-Gal*p*-O$(CH_2)_7CH_3$, respectively (**Figure 21**). Asterisks indicate mercury derivative structures. For GTA+A and GTB+B diffraction data were collected to a maximum resolution of 1.45 and 1.47 Å with a final R_{work} of 18.2 and 18.3% and R_{free} of 19.3 and 20.4%, respectively. For GTA*+A and GTB*+B data were collected to a maximum resolution of 1.89 and 1.57 Å, with a final R_{work} of 18.8 and 17.4% and R_{free} of 22.4 and 19.6%, respectively. In each case crystals were soaked with the trisaccharide product prior to data collection, in a manner similar to that outlined in **Chapter 3.3.4**. All structures showed excellent density along the entire length of the polypeptide chain except for the two characteristically flexible regions, the internal loop and the C-terminal tail (**Table 15**). For the

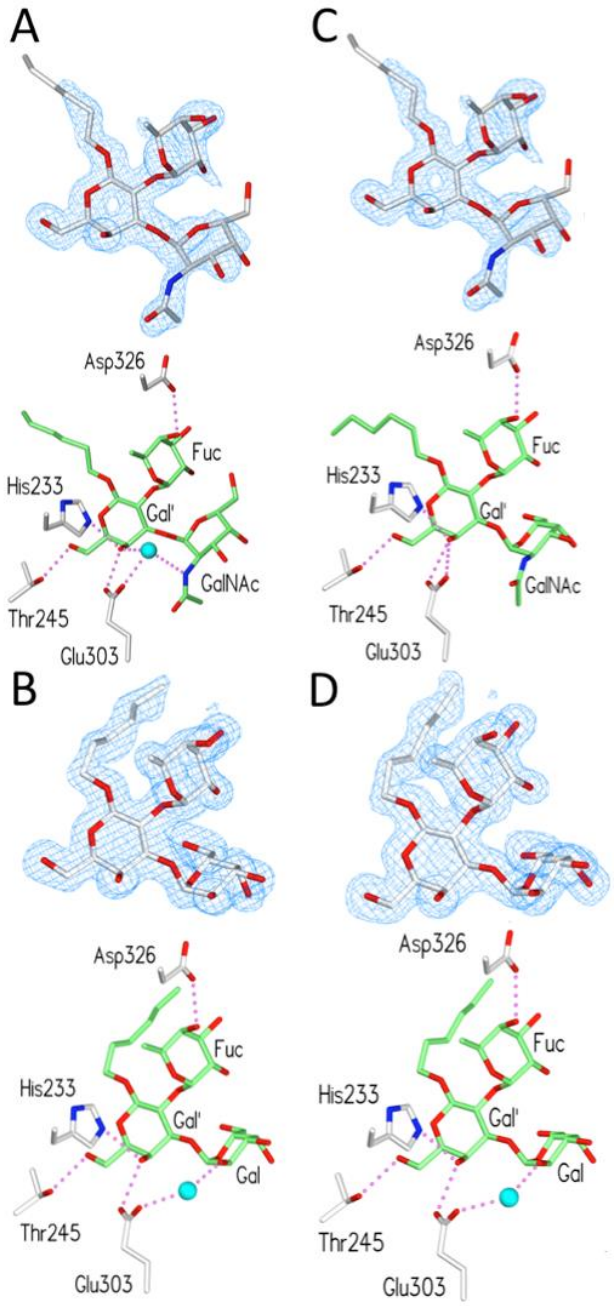
mercury derivative structures, the degree of disorder is greater than in the native structures, consistent with other published data, including the 303 mutants discussed in **Chapter 4**.

Table 14. Data collection and refinement statistics for Hg derivative and non-derivative GTA/GTB structures bound by their respective product trisaccharides*

	GTA*+A *	**GTB*+B ***	**GTA+A**	**GTB+B**
Resolution (Å)	20 - 1.89	20 - 1.57	20 – 1.45	20 – 1.47
Space Group	$C222_1$	$C222_1$	$C222_1$	$C222_1$
a (Å)	52.58	52.71	52.55	52.67
b (Å)	150.66	150.65	148.66	150.03
c (Å)	79.30	79.18	79.79	79.25
R_{merge} (%)[a,b]	4.7 (31.7)	4.0 (31.7)	3.6 (31.0)	4.3 (34.6)
Completeness (%)[b]	99.3 (99.9)	98.6 (98.1)	99.1 (97.8)	99.3 (97.7)
Total unique reflections	25,471	43,816	55,231	53,342
Multiplicity	4.74	4.46	4.27	4.01
Mean I/σ(I)	16.3	17.3	17.7	15.1
Non-hydrogen trisaccharide atoms	42	39	42	39
Non-hydrogen protein atoms	2094	2121	2326	2290
Refinement				
R_{work} (%)[c]	18.8	17.4	18.2	18.3
R_{free} (%)[c,d]	22.4	19.6	19.3	20.4
Reflections used in refinement	24171	41593	52424	50642
R_{free} reflections	1299	2213	2805	2706
Modeled protein residues	260	263	284	278
Wilson B-factor (Å^2)	29.8	22.5	20.5	20.9
Overall <B> (Å^2)	31.0	24.0	21.0	21.0
Trisaccharide <B > (Å^2)	50.1	29.7	40.8	26.9
Solvent <B> (Å^2)	57.6	56.1	44.5	41.7
No. water	86	200	202	243
r.m.s. bond (Å)[e]	0.0083	0.0101	0.0096	0.0102
r.m.s. angle (°)[e]	1.3535	1.5000	1.4481	1.5050
Ramachandran outliers (%)	0	0	0	0
Rotamer outliers	2	2	2	3
Clashscore	1	2	1	2
PDB code	5TJL	5TJO	5TJK	5TJN

[a] R-merge, $\sum |I_{obs} - I_{ave}| / \sum I_{ave}$.
[b] Values in parentheses represent highest resolution shell.
[c] R-work, $\sum||F_o| - |F_c|| / \sum |F_o|$.
[d] 10% of reflections were omitted for R-free calculations.
[e] r.m.s. root-mean-square.
Reproduced from (Gagnon et al. 2017) with permission.

Figure 21. *GTA/GTB bound to their respective product trisaccharides*



Active site interactions of GTA (upper panels) and GTB (lower panels) in complex with A and B trisaccharide antigens. Electron density (from the $2F_O$-F_C maps contoured at 1.00 s) corresponds to: (**A**) GTA complexed with product trisaccharide A (the GTA+A structure) and refined to 1.45 Å, (**B**) GTB complexed with product trisaccharide B (the GTB+B structure) and refined to 1.47 Å, (**C**) GTA derivative structure complexed with product trisaccharide A in the presence of mercury (GTA*+A) and refined to 1.89 Å, (**D**) GTB derivative structure complexed with product trisaccharide B in the presence of mercury (GTB*+B) and refined to 1.57 Å. Protein atoms are colored by element with carbon *light gray*, oxygen *red*, nitrogen *blue*, and sulfur *yellow*, while product analog carbon atoms are colored *green*. Water molecules are shown as *cyan* spheres, manganese ions as *magenta* spheres, and hydrogen bond interactions as *pink* dashed spheres. Reproduced from (Gagnon et al., 2017) with permission.

Table 15. Internal loop, C-terminal loop, and product ordering for the trisaccharide complexes

Enzyme	Internal loop 176 181 186 191 196	C-terminus 346 351	Terminal saccharide
GTA+A	EV RaykR WQDVS MRRME MISdF CERR	VP knhqa vrnp	Ordered
GTB+B	EV Gaykr wqDVS MRRME MISDF CERR	VP knhqa vrnp	Ordered
GTA+A+Hg	EV raykr wqdvs mrrme misdf ceRR	VP knhqa vrnp	Ordered
GTB+B+Hg	EV gaykr wqdvs mrrme misdf cERR	VP knhqa vrnp	Ordered

Black one letter amino acid codes correspond to electron density for main and side chain atoms, blue corresponds to electron density for main chain atoms only, and red correspond to weak electron density for main and side chain atoms. Residues with one letter amino acid codes in lower case have not been included in the refined models. Reproduced from (Gagnon et al. 2017) with permission.

Table 16. Family I and II conformation dihedral angle ranges for A- and B-trisaccharides

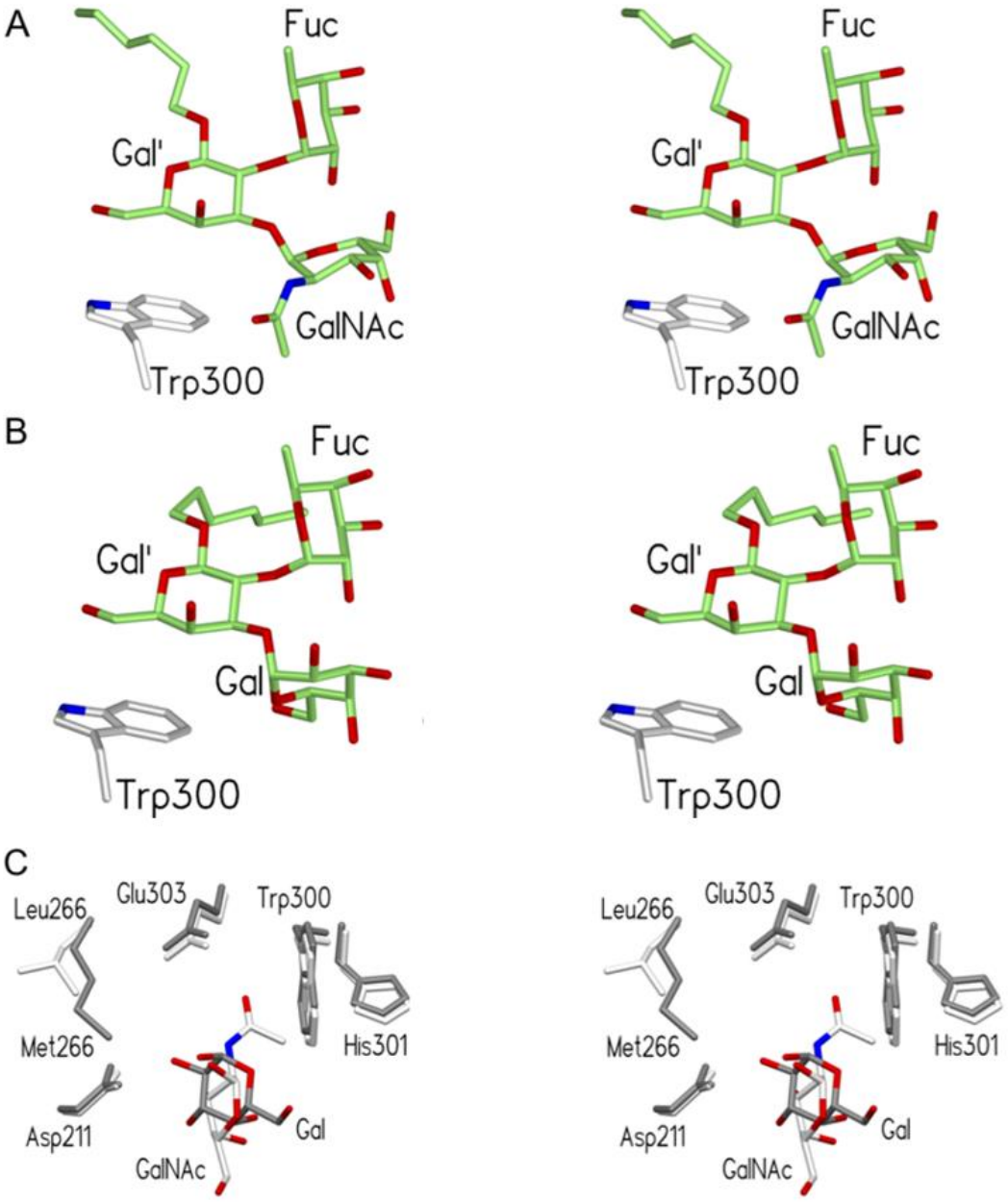
	Dihedral angle	A-Trisaccaride				B-Trisaccharide			
		fam I	***fam II***	***observed***		***fam I***	***fam II***	***observed***	
Φ_1 (º)	$O5_{\alpha GN/G}$-$C1_{\alpha GN/G}$-$O1_{\alpha GN/G}$-$C3_{\beta G}$	62 to 82	64 to 79	82	57*	63 to 88	68 to 89	53	55*
Ψ_1 (º)	$C1_{\alpha GN/G}$-$O1_{\alpha GN/G}$-$C3_{\beta G}$-$C4_{\beta G}$	61 to 74	69 to 81	111	75*	60 to 72	65 to 81	72	70*
Φ_2 (º)	$O5_{\alpha F}$-$C1_{\alpha F}$-$O1_{\alpha F}$-$C3_{\beta G}$	-77 to -67	-102 to -91	-60	-62*	-78 to -65	-101 to -91	-62	-64*
Ψ_2 (º)	$C1_{\alpha F}$-$O1_{\alpha F}$-$C3_{\beta G}$-$C4_{\beta G}$	-109 to -86	-173 to -164	-93	-94*	-110 to -86	-173 to -164	-98	-98*
	Computed abundance	7.00%	92.5%	n/a		5.40%	93.6%	n/a	

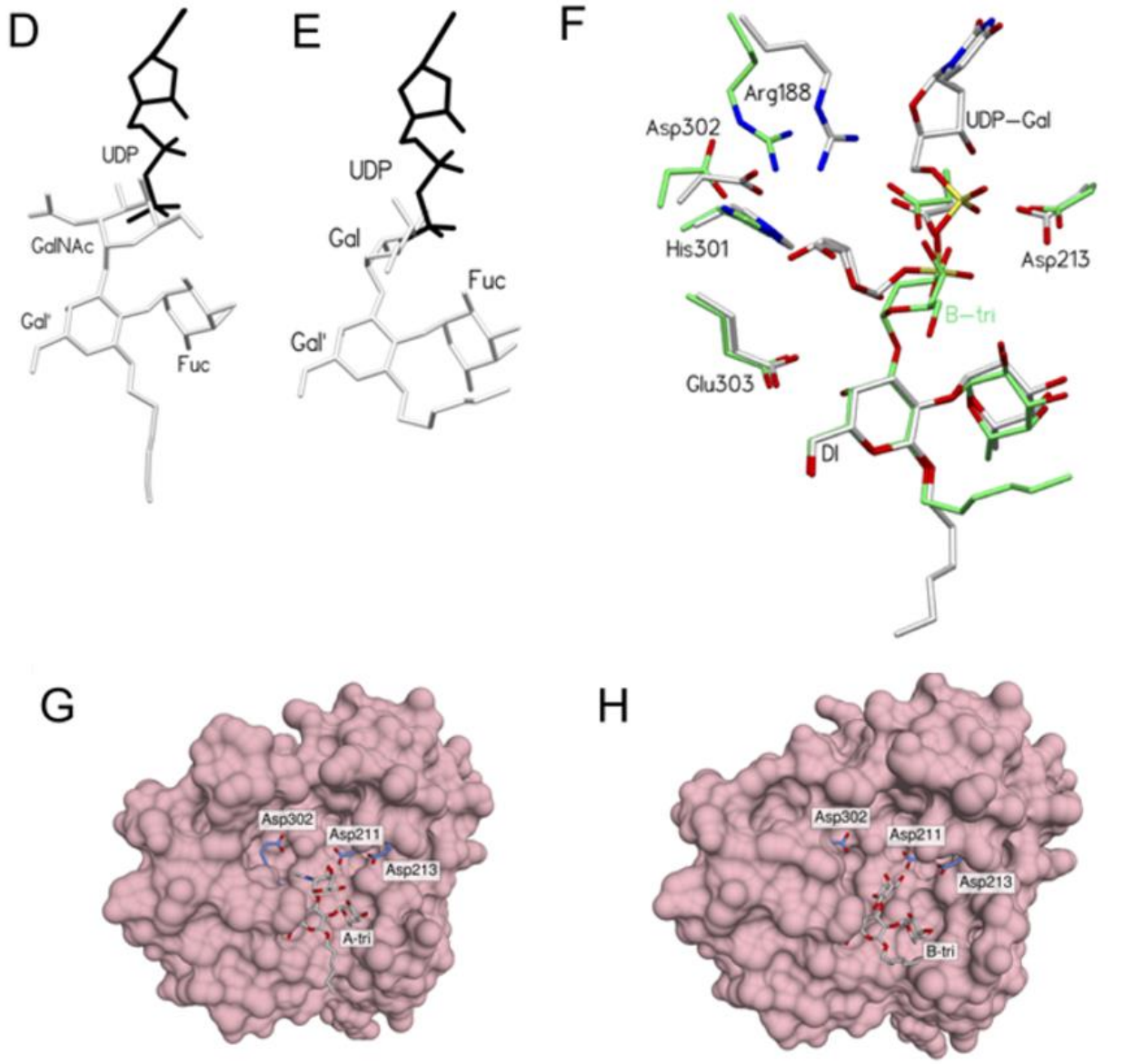
*value for derivative structure

Dihedral angle ranges for A and B trisaccharides are those reported by Imberty et al. (1995). Reproduced from (Gagnon et al. 2017) with permission.

The trisaccharide analogues, in both native and derivative structures, exhibit good electron density (**Figure 21**). In GTA+A and GTA*+A, the A-trisaccharide occupies the same relative position within the active site (**Figure 21A&C**), as does the B-trisaccharide in GTB+B and GTB*+B (**Figure 21B&D**). Contacts between the enzyme active site residues and the trisaccharide product β-Gal and α-Fuc moieties (*i.e.* the H-antigen disaccharide backbone) are consistent with previous acceptor-bound complexes (Alfaro, J.A., et al. 2008). In these, the β-Gal C4-hydroxyl hydrogen bonds with the His233 and Glu303 side chains, the β-Gal C6-hydroxyl hydrogen bonds with the Thr245 side chain (**Figure 21**), and there is a hydrophobic interaction between the acceptor/product β-Gal and Trp300 rings (**Figure 22A&B**). In the published acceptor-bound structures (Alfaro, J.A., et al. 2008), the α-Fuc C2- and C3-hydroxyls make contacts to C-terminal residue His348 (ordering the C-terminal tail in the process), and the C4-hydroxyl forms a hydrogen bond with the side chain of Asp326. However, in the structures reported here, the C-terminal tail is largely disordered, and though the trisaccharide α-Fuc C4-hydroxyl interacts with the Asp326 side chain, it does not make contacts to His348.

Figure 22. Hydrophobic interactions between GTA/GTB and product trisaccharides



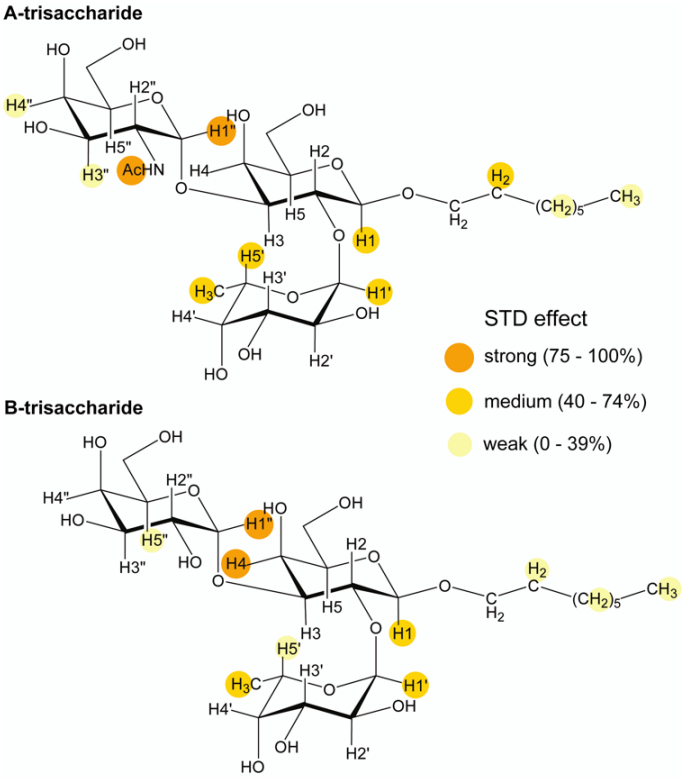


Hydrophobic stacking interactions between Trp300 and trisaccharide products in (**A**) GTA+A and (**B**) GTB+B. (**C**) Overlap of trisaccharide product α-GalNAc from GTA+A (white) and α-Gal from GTB+B (*gray*). Superposition of (**D**) GTA+A A-trisaccharide and (**E**) GTB+B B-trisaccharide products (*black*) with the tucked under donor substrate UDP moiety (*gray*) as it appears in Alfaro *et al.* (2008; PDB code 2RJ7). (**F**) B-trisaccharide (*green* carbon atoms) bound in the active site of GTB+B superimposed with the UDP and deoxy inhibitor (DI) acceptor (*grey* carbon atoms) substrates bound in the active site of chimeric GTA/GTB enzyme AABB (PDB code 2RJ7) as reported by Alfaro *et al.* (2008). Active site residue side chains of GTB (*green* carbon atoms) and AABB (*grey* carbon atoms) are displayed, amongst them Arg188, Asp211, Asp213, His301, Asp302, and Glu303. Surface representation of (**G**) GTA+A and (**H**) GTB+B with product trisaccharides (*grey* carbon atoms) and active site residue side chains of Asp211, Asp213, and Asp302 displayed (*blue* carbon atoms). The remaining protein surfaces are colored *pink*. With the exception of the carbon atoms indicated, atom coloring is by element, with nitrogen *blue*, oxygen *red*, and phosphorous *yellow*. Reproduced from (Gagnon et al., 2018) with permission.

STD NMR data

STD NMR spectroscopy allowed in-solution characterization of A- and B-trisaccharide binding *via* the determination of binding epitopes, which reflect the distances from ligand protons to the protons of binding site residues. Binding epitopes for A- and B-trisaccharide are shown in **Figure 23**. For both products, the hydrogen atoms around the glycosidic linkage between α-D-GalNAc/α-D-Gal and β-D-Gal received the largest amount of saturation, namely $H1_{\alpha\text{-D-GalNAc}}$ and $NAc_{\alpha\text{-D-GalNAc}}$ of the A product and $H1_{\alpha\text{-D-Gal}}$ and $H4_{\beta\text{-D-Gal}}$ of the B product.

Figure 23. *STD NMR data of product trisaccharides in the presence of GTA/GTB*



(**A**) Binding epitopes of A- and B-trisaccharide as determined by STD NMR spectroscopy. Circles illustrate the relative size of saturation transfer and reflect the proximity to the binding site, where a closer contact shows larger STD effects. The epitopes were derived from STD build-up curves and the highest value was set to 100%. The absence of circles indicates resonances that were not fully resolved, and for

which the STD effect cannot be stated unambiguously. Reproduced from (Gagnon et al., 2018) with permission.

7.3 Discussion & conclusions

Trisaccharide β-Gal moiety dominates binding

As with HA and DI in previous structures (Patenaude, S.I., et al. 2002, Nguyen, H.P., et al. 2003, Letts, J.A., et al. 2006), the structures reported here show that the trisaccharide β-Gal residue dominates binding through three hydrogen bond interactions with active site residues Thr245, His233, and Glu303. The STD NMR data agrees with this: STD effects are largest around the glycosidic linkage between α-Gal/α-GalNAc and β-Gal (**Figure 23**), and the strongest saturation transfers correspond to H-4 of β-Gal (B-trisaccharide) and H-1 of α-Gal/α-GalNAc (**Figure 23**). Based on the crystal structures, the β-Gal 4-OH is the only hydroxyl to form three-centered hydrogen bonds, in this case with amino acids His233 and Glu303 (**Figure 22**). There is a hydrophobic interaction between Trp300 and the β-Gal H4 (**Figure 22**), which renders the product trisaccharide surface inaccessible to the bulk solvent about β-Gal C3, C4, and C5 and α-Gal/α-GalNAc C1.

Product binds via *hydrogen bond and van der Waals interactions*

Carbohydrates possess weakly polar hydroxyl groups and have regions that lack hydrophobicity. Thus, generally carbohydrate-protein interactions have low affinities and tend to be dynamic, relying on an assortment of hydrogen bond and van der Waals contacts (Kadirvelraj, R., et al. 2008). Further, ordered water molecules proximal to a carbohydrate-protein interface are common in crystallography, and these molecules can have a significant impact on the stability and affinity of carbohydrate-protein interactions (Kadirvelraj, R., et al. 2008). Published crystal structures of GTA/GTB in complex with acceptor analogues universally show a water

molecule bridging the β-Gal 3-OH and 4-OH and the Glu303 Oε (Letts, J.A., et al. 2006, Alfaro, J.A., et al. 2008, Schuman, B., et al. 2010). In the structures reported here, the β-Gal O3 is covalently linked to C1 of the transferred monosaccharide (which is either GalNAc or Gal), occluding this instance of water-mediated hydrogen bonding. As with the donor- and acceptor analogue-bound structures, a single water molecule forms the only observed bridging hydrogen bond interaction between enzyme and the product terminal α-GalNAc/α-Gal. In the trisaccharide-bound GTA structures, this water molecule bridges the β-Gal O4-hydroxyl and the α-GalNAc N2 to the Glu303 Oε (**Figure 21**), while in the corresponding GTB structures it bridges the α-Gal O2-hydroxyl to the Glu303 Oε (**Figure 21**). As well, the GTB+B structure features a unique van der Waals contact between the terminal carbohydrate residue and the Cε of critical amino acid Met266. Consistent with these observations, a medium STD effect is observed for H5 of the B-trisaccharide α-Gal moiety (**Figure 23**). In the case of A-trisaccharide the corresponding STD effect to H5 of α-GalNAc has not been quantified because the signals of H5 of α-GalNAc and H4 of β-Gal overlap (**Figure 23**). According to the crystal structures in GTB+B residue Met266 is ~1 Å closer to the corresponding α-Gal moiety C5 than is Leu266 to the α-GalNAc moiety C5 in GTA+A.

Though the product trisaccharide β-Gal and Fuc moieties are positioned similarly in the GTA/GTB active sites, the terminal α-Gal/α-GalNAc orientations differ. Relative to the B-trisaccharide α-Gal moiety, the A-trisaccharide α-GalNAc is shifted ~1 Å away, presumably to avoid a steric clash between Asp211 and the bulky GalNAc acetamido group (**Figure 22**).

Spatial conflict between UDP and product

The substrate-complexed AABB chimera was the first published structure that showed any of the wild-type, chimeric or mutant enzymes in the fully closed conformation with an organized internal loop and C-terminal tail (Alfaro, J.A., et al. 2008). Over nearly all GTA, GTB, and other chimeric (including AABB) and mutant enzyme structures, the donor UDP binds within the Rossmann fold-like donor binding site in a virtually invariant conformation (Patenaude, S.I., et al. 2002, Marcus, S.L., et al. 2003, Nguyen, H.P., et al. 2003, Lee, H.J., et al. 2005, Letts, J.A., et al. 2006, Persson, M., et al. 2007, Alfaro, J.A., et al. 2008). Overlap of the trisaccharide product-bound structures with the corresponding UDP-bound structures deposited in the PDB reveals an obvious conflict between the UDP-donor β-phosphate and the product analogue α-GalNAc and α-Gal moieties (**Figure 22D&E**). Binding of the A- and B-trisaccharide products are within the GTA/GTB active site is depicted in surface representation in **Figure 22G&H**. In GTA, the α-GalNAc O3-hydroxyl occupies the same space as would the UDP β-phosphate (**Figure 22**), and this is also true for GTB, where the clash is between the α-Gal O3-hydroxyl and C4-carbon and the UDP β-phosphate (**Figure 22**). This suggests that neither trisaccharide can occupy the active site simultaneously with UDP.

Trisaccharide binds in a higher-energy conformation

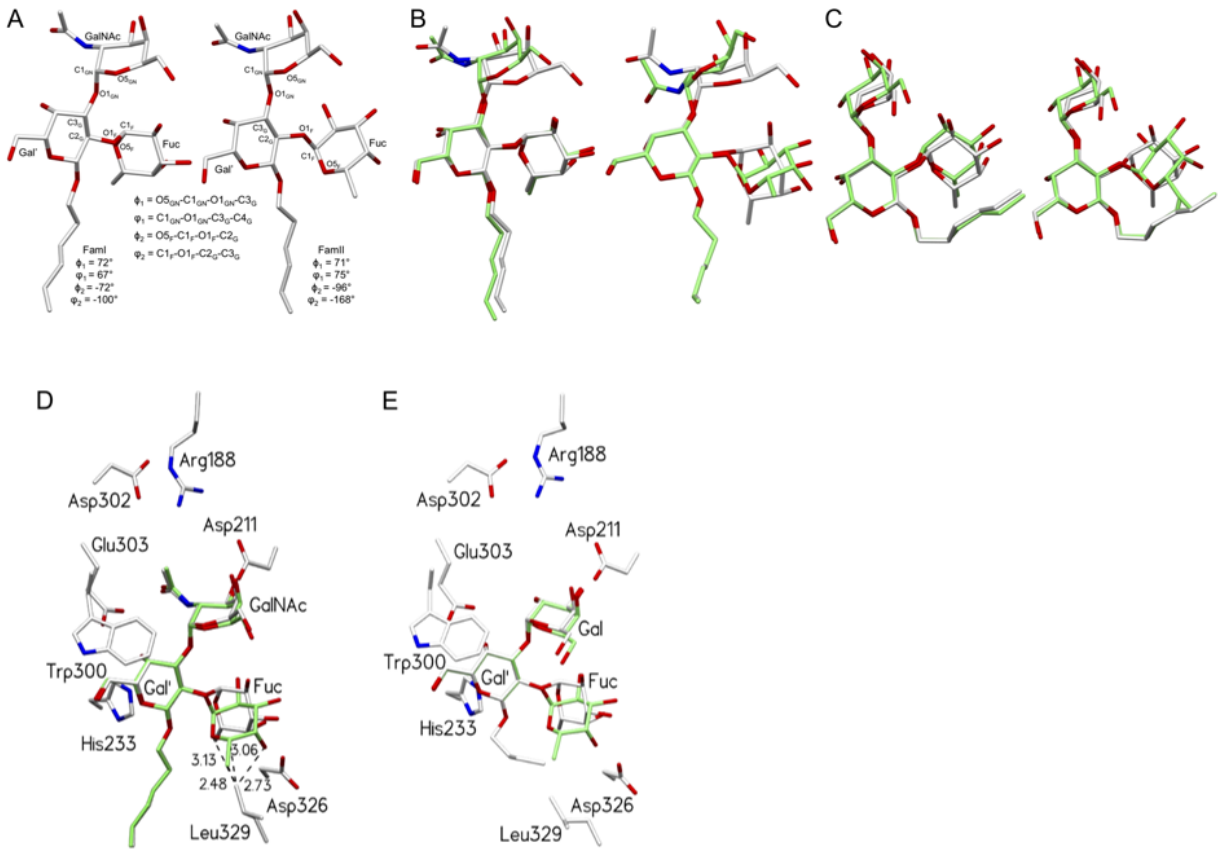
Initial conformational studies involving the nuclear Overhauser effect (NOE) and molecular modeling report that the A- and B-trisaccharides assume only one conformation (Lemieux, R.U., et al. 1980, Bush, C.A., et al. 1986). Later, a more flexible two-conformation model emerged (Imberty, A., et al. 1995, Casset, F., et al. 1996), though the one-conformation model retained support (Otter, A., et al. 1999, Azurmendi, H.F. and Bush, C.A. 2002). Imberty *et al.*'s (1995) two predicted conformational included a high-abundance family II (Fam II)

conformation (~93% of the total solution population) and a low-abundance family I (Fam I) conformation (~6% of the total solution population) (**Table 16**). Based on their data, Fam I had a higher energy relative to Fam II. The dihedral angle ranges reported for Fam I and Fam II α-GalNAc-(1→3)-β-Gal and α-Gal-(1→3)-β-Gal were similar. The geometric variance lay with the α-Fuc-(1→2)-β-Gal glycosidic linkage (**Table 16** and **Figure 24A-C)**. Theoretically, either conformation would be expected to fit within the GTA/GTB active site, however in the crystal structures we observe only the Fam I trisaccharide (**Table 16, Figure 24B&C**). For the A-trisaccharide, this is comparable to the results of another earlier NMR study (Casset, F., et al. 1996), where the antigen was observed in the Fam I conformation in complex with a lectin. Similarly, the B-trisaccharide conformation seen here is analogous to another NMR structure reported by Otter *et al.*, which revealed the B-trisaccharide in the Fam I conformation (Otter, A., et al. 1999).

It is not obvious why the higher-energy Fam I conformation would be preferred. Contacts between C-terminal residue His348 and the acceptor Fuc O2 and O3 hydroxyls are known key polar groups essential for GTA/GTB catalytic activity (Mukherjee, A., et al. 2000). Fuc O2 acts as both a hydrogen bond donor and acceptor and participates in acceptor recognition, while Fuc O3 acts as a hydrogen bond acceptor (Alfaro, J.A., et al. 2008). In all structures, the C-terminal tail is disordered; with the trisaccharides in the Fam I conformation, the β-Gal and α-Fuc moieties approximate the position the acceptor disaccharide occupies in other structures – here, the key polar groups are estimated to lie within hydrogen bonding distance of His348. The Fam II trisaccharide conformation would shift the Fuc residue away from His348 and impede the formation of these interactions. Another feature that may explain the Fam I preference is the presence of a stabilizing hydrophobic interaction between Leu329 and Fuc C4 (**Figure 24D&E**),

though this only occurs in GTA+A. Probably due to the dynamic nature of Leu329, in GTB+B the leucine side chain is rotated out of the active site, resulting in no observable explanation for Fam I preference in this case.

Figure 24. *GTA/GTB product trisaccharide conformations*



(**A**) Representation of family I (l*eft*) and family II (*right*) A-trisaccharide product conformations based on dihedral angles given in Table X and reported by Imberty *et al.* (1995). The B-trisaccharide family I and II conformers are highly similar to those of the A-trisaccharide. (**B**) Overlap of the trisaccharide product in GTA+A with family I (*left*) and family II A-tri (*right*). (**C**) Overlap of the trisaccharide product in GTB+B with family I (*left*) and family II B-tri (*right*). β-Gal was used as the orientation point for all overlaps. (**D**) Overlap of family I and II A-trisaccharide within the active site of GTA+A. Black dashed lines represent steric conflicts. Distances are in Ångströms. (**E**) Overlap of family I and II B-tri within the active site of GTB+B. In (**B**) and (**C**) family I and family II A- and B-tri carbon atoms are colored green, while the products as observed in GTA+A and GTB+B structures have gray carbon atoms. In (**D**) and (**E**) the family I A- and B-tri carbon atoms are colored *grey*, while the family II A- and B-tri carbon atoms are colored *green*. With the exception of the specified carbon atoms, all other atoms in (**A**)-(**E**) are colored by element with carbon *grey*, nitrogen *blue*, and oxygen *red*. Reproduced from (Gagnon et al., 2018) with permission.

Post-glycosyltransfer active site disorder

Based on superposition with published structures of GTA/GTB with UDP in the donor-binding site, following transfer of the monosaccharide to the acceptor molecule, there would be a clash between the resultant trisaccharide product and UDP leaving group β-phosphate. UDP must either leave the active site or else undergo a significant change in conformation or position. Alternatively, or in addition to these events, product release must occur. The double displacement mechanism circumvents the steric clash problem, since there is an initial monosaccharide transfer to the enzyme that results in formation of a glycosyl-enzyme intermediate. Here, UDP can shift position or leave the active site prior to or concomitant with the subsequent nucleophilic attack by the acceptor. In contrast, the S_Ni mechanism suggests that the UDP-donor remains present briefly after sugar transfer to assist in deprotonating acceptor or to stabilize the partial positive charge of the detached monosaccharide. Based on our crystal structure and STD NMR data, this would result in an unfavorable conformation: direct nucleophilic attack by the acceptor upon the donor sugar would yield a product trapped under the β-phosphate of the UDP. To examine the order of product formation and departure in family 8 lipopolysaccharyl α-galactosyltransferase C, Ly *et al.* (2002) used a combination of kinetic data and structural results (Ly, H.D., et al. 2002). Their approach may prove useful in the study of GTA/GTB and may provide a necessary complement to our structural data.

The Theorell-Chance kinetic mechanism proposed for GTB predicts the formation of only binary complexes (Kamath, V.P., et al. 1999). According to this model, UDP is released from the enzyme after the donated sugar residue forms a glycosyl-enzyme intermediate with acceptor binding and product formation to follow. This two-step mechanism resolves any potential conflict that might otherwise arise between the β-phosphate of the UDP and the α-Gal

residue of the product (Kamath, V.P., et al. 1999). This is consistent with the ESI-MS findings of 2009: GTA can weakly bind the A-trisaccharide product in the presence (K_a = 1.2 x 10^3 M^{-1}) or absence (K_a = 1.4 x 10^3 M^{-1}) of UDP and Mn^{2+}, with association favored slightly in the latter case, while GTB weakly associates with the B-trisaccharide product alone (K_a = 2.7 x 10^3 M^{-1}) but exhibits no appreciable association in the presence of UDP and Mn^{2+} (K_a < 10^3 M^{-1}) (Soya, N., et al. 2009).

To unravel the mechanism of retaining GTs, Schuman *et al.* (2013) proposed the OA mechanism, which is subtly distinct from S_Ni in that it does not invoke dissociative donor substrate character. While this has credence for certain ret-GTs, particularly those lacking an appropriately-positioned nucleophile (Sinnott, M.L. 1990, Lairson, L.L., et al. 2008, Schuman, B., et al. 2013), our structural data is most consistent with double displacement, which would permit UDP departure prior to product formation. Double displacement features an acyl-enzyme intermediate, which arises *via* an enzyme-initiated nucleophilic attack on the donor; in a subsequent step, the acceptor performs a second nucleophilic attack, collapsing the covalent adduct to yield product. In this mechanism, the first step produces the UDP leaving group, which may then leave the active site or shift position prior to product formation in the next step. This would circumvent the clash between UDP and trisaccharide product (**Figure 22D&E**).

Conclusions

The GTA/GTB crystal structures collected in complex with their trisaccharide reaction products agree with the solution binding epitopes uncovered through STD NMR. Together, these reveal a spatial conflict between the product α-GalNAc/α-Gal moiety and the donor UDP β-phosphate, suggesting a mode of product release where this clash disorders the active site and prevents UDP from being stably bound. Future work should include kinetic investigation to

define the order of UDP and product release as a complement to ongoing efforts to address the mechanism of glycosyltransfer in retaining GTs.

Chapter 8: Summary & future work

The investigations described herein have been fruitful in explaining the processes of substrate recognition, mobile loop organization, and product release in GTA and GTB. The study of Glu303 mutants crystallized with and without mercury provided an important lesson regarding the fine-tuned and complex nature of substrate binding and catalysis in GTA and GTB, which can display anomalously distinct behavior despite their homology. In terms of kinetics, the residual activity of the 303 mutants suggests that these enzymes may not depend on a strong nucleophile at position 303 for catalysis. As suggested in recent QM/MM studies of the enzymes, GTA/GTB may possess mechanistic elasticity, a property that may confound our efforts to reach mechanistic conclusions.

Next, the massive set of eighteen donor- and donor analogue-bound structures supplied evidence for at least four distinct donor conformations. A handful of these complexes exhibited dynamic information concerning substrate binding, perhaps a surprising outcome from a methodology often viewed as static. We were able to infer that multi-stage nature of donor binding, where *via* stepwise adjustment, the donor sugar moves toward the catalytically active conformation.

Surprisingly, this conformational data proved vital to the interpretation of the subsequent study of Arg188 and Asp302 mutants. These residues, together with Ser185 and Asp211, help to achieve the catalytic state by making contacts that stabilize the internal loop as well as the donor as it undergoes conformational shifts. Notably, we saw that the Arg188/Asp302 salt bridge is a conserved feature, to varying degrees, among GT-A fold-type glycosyltransferases.

Last, the STD-NMR results and the structures of GTA/GTB in complex with their trisaccharide reaction products suggested steric conflict between the formed product and the donor UDP β-phosphate. This clash disorders the active site and prevents stable binding of UDP. As well, through the crystal structure data and in reference to earlier solution NMR work, we demonstrated stable binding of a higher energy product conformation.

Though the question of the retaining glycosyltransfer mechanism remains, efforts to crystallize the trapped covalent intermediate observed by Soya *et al.* (2011) using recombinantly expressed GTA/GTB E303C are underway. The ideal next step in this mechanistic line of inquiry is to obtain structure of the acyl-enzyme adduct. This structure would confirm the double displacement mechanism as a realized mechanistic pathway for GTA and GTB and perhaps other retaining GTs and may inform the rational design of selective glycosyltransferase inhibitors. Provided the active site is sufficiently well-ordered, an X-ray crystal structure should reveal the stereochemistry of the covalent adduct. A uniformly inverted stereochemistry would provide convincing evidence of double displacement, at least for the 303 mutant enzymes. If crystallizing the adduct remains an obstacle, it may be possible to use NMR to make stereochemical conclusions.

References

Albesa-Jove D, Guerin ME. 2016. The conformational plasticity of glycosyltransferases. *Curr Opin Struct Biol*, 40:23-32.

Albesa-Jove D, Sainz-Polo MA, Marina A, Guerin ME. 2017. Structural snapshots of alpha-1,3-galactosyltransferase with native substrates: insight into the catalytic mechanism of retaining glycosyltransferases. *Angew Chem Int Ed Engl*, 56:14853-14857.

Alfaro JA, Zheng RB, Persson M, Letts JA, Polakowski R, Bai Y, Borisova SN, Seto NOL, Lowary TL, Palcic MM, *et al.* 2008. ABO(H) blood group A and B glycosyltransferases recognize substrate via specific conformational changes. *J Biol Chem*, 283:10097-10108.

Angulo J, Langpap B, Blume A, Biet T, Meyer B, Krishna NR, Peters H, Palcic MM, Peters T. 2006. Blood group B galactosyltransferase: Insights into substrate binding from NMR experiments. *J Am Chem Soc*, 128:13529-13538.

Ardèvol A, Rovira C. 2015. Reaction mechanisms in carbohydrate-active Enzymes: glycoside hydrolases and glycosyltransferases. insights from ab initio quantum mechanics/molecular mechanics dynamic simulations. *J Am Chem Soc*, 137:7528-7547.

Azurmendi HF, Bush CA. 2002. Conformational studies of blood group A and blood group B oligosaccharides using NMR residual dipolar couplings. *Carbohyd Res*, 337:905-915.

Beaton SA, Huestis MP, Sadeghi-Khomami A, Thomas NR, Jakeman DL. 2009. Enzyme-catalyzed synthesis of isosteric phosphono-analogues of sugar nucleotides. *Chem Commun*:238-240.

Bella M, Koos M, Lin CH. 2015. Towards inhibitors of glycosyltransferases: A novel approach to the synthesis of 3-acetamido-3-deoxy-D-psicofuranose derivatives. *Beilstein J Org Chem*, 11:1547-1552.

Blackler RJ, Gagnon SM, Polakowski R, Rose NL, Zheng RB, Letts JA, Johal AR, Schuman B, Borisova SN, Palcic MM, *et al.* 2017. Glycosyltransfer in mutants of putative catalytic residue Glu303 of the human ABO(H) A and B blood group glycosyltransferases GTA and GTB proceeds through a labile active site. *Glycobiology*, 27:370-380.

Blume A, Angulo J, Biet T, Peters H, Benie AJ, Palcic M, Peters T. 2006. Fragment-based screening of the donor substrate specificity of human blood group B galactosyltransferase using saturation transfer difference NMR. *J Biol Chem*, 281:32728-32740.

Bobovska A, Tvaroska I, Kona J. 2014. A theoretical study on the catalytic mechanism of the retaining [small alpha]-1,2-mannosyltransferase Kre2p/Mnt1p: the impact of different metal ions on catalysis. *Organic & Biomolecular Chemistry*, 12:4201-4210.

Bobovská A, Tvaroška I, Kóňa J. 2015. Theoretical study of enzymatic catalysis explains why the trapped covalent intermediate in the E303C mutant of glycosyltransferase GTB was not detected in the wild-type enzyme. *Glycobiology*, 25:3-7.

Boix E, Swaminathan GJ, Zhang YN, Natesh R, Brew K, Acharya KR. 2001. Structure of UDP complex of UDP-galactose :beta-galactoside-alpha-1,3-galactosyltransferase at 1.53-angstrom resolution reveals a conformational change in the catalytically important C terminus. *J Biol Chem*, 276:48608-48614.

Bush CA, Yan ZY, Rao BNN. 1986. Conformational energy calculations and proton nuclear Overhauser enhancements reveal a unique conformation for blood group A oligosaccharides. *J Am Chem Soc*, 108:6168-6173.

Campbell JA, Davies GJ, Bulone V, Henrissat B. 1997. A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem J*, 326:929-939.

Casset F, Peters T, Etzler M, Korchagina E, Nifantev N, Perez S, Imberty A. 1996. Conformational analysis of blood group A trisaccharide in solution and in the binding site of Dolichos biflorus lectin using transient and transferred nuclear Overhauser enhancement (NOE) and rotating-frame NOE experiments. *Eur J Biochem*, 239:710-719.

Clarke AJ, Hurtado-Guerrero R, Pathak S, Schuttelkopf AW, Borodkin V, Shepherd SM, Ibrahim AFM, van Aalten DMF. 2008. Structural insights into mechanism and specificity of O-GlcNAc transferase. *Embo J*, 27:2780-2788.

Clausen H, Hakomori S. 1989. ABH and related histo-blood group antigens; immunochemical differences in carrier isotypes and their distribution. *Vox Sang*, 56:1-20.

Compain P, Martin OR. 2001. Carbohydrate mimetics-based glycosyltransferase inhibitors. *Bioorg Med Chem*, 9:3077-3092.

Compain P, Martin OR. 2003. Design, synthesis and biological evaluation of iminosugar-based glycosyltransferase inhibitors. *Curr Top Med Chem*, 3:541-560.

Cooling L. 2015. Blood groups in infection and host susceptibility. *Clin Microbiol Rev*, 28:801-870.

Coutinho PM, Deleury E, Davies GJ, Henrissat B. 2003. An evolving hierarchical family classification for glycosyltransferases. *J Mol Biol*, 328:307-317.

Creighton TE. 1993. *Proteins: Structures and molecular properties*. 2 ed. W. H. Freeman and Company:New York.

Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography*, 66:486-501.

Evans SV. 1993. Setor - hardware-lighted 3-dimensional solid model representations of macromolecules. *J Mol Graphics*, 11:134-&.

Ewald DR, Sumner SC. 2016. Blood type biochemistry and human disease. *Wiley Interdiscip Rev Syst Biol Med*, 8:517-535.

Fastag E, Varon J, Sternbach G. 2013. Richard Lower: the origins of blood transfusion. *The Journal of emergency medicine*, 44:1146-1150.

Ferguson-Smith MA, Aitken DA, Turleau C, de Grouchy J. 1976. Localisation of the human ABO: Np-1: AK-1 linkage group by regional assignment of AK-1 to 9q34. *Human genetics*, 34:35-43.

Franchini M, Liumbruno GM, Lippi G. 2016. The prognostic value of ABO blood group in cancer patients. *Blood transfusion = Trasfusione del sangue*, 14:434-440.

Gagnon SM, Meloncelli PJ, Zheng RB, Haji-Ghassemi O, Johal AR, Borisova SN, Lowary TL, Evans SV. 2015. High resolution structures of the human ABO(H) blood group enzymes in complex with donor analogs reveal that the enzymes utilize multiple donor conformations to bind substrates in a stepwise manner. *J Biol Chem*, 290:27040-27052.

Gagnon SML, Legg MSG, Polakowski R, Letts JA, Persson M, Lin S, Zheng RB, Rempel B, Schuman B, Haji-Ghassemi O, *et al.* 2018. Conserved residues Arg188 and Asp302 are critical for active site organization and catalysis in human ABO(H) blood group A and B glycosyltransferases. *Glycobiology*, 28:624-636.

Gagnon SML, Legg MSG, Sindhuwinata N, Letts JA, Johal AR, Schuman B, Borisova SN, Palcic MM, Peters T, Evans SV. 2017. High-resolution crystal structures and STD NMR mapping of human ABO(H) blood group glycosyltransferases in complex with trisaccharide reaction products suggest a molecular basis for product release. *Glycobiology*, 27:966-977.

Gastinel LN, Bignon C, Misra AK, Hindsgaul O, Shaper JH, Joziasse DH. 2001. Bovine alpha 1,3-galactosyltransferase catalytic domain structure and its relationship with ABO histo-blood group and glycosphingolipid glycosyltransferases. *Embo J*, 20:638-649.

Gloster TM. 2014. Advances in understanding glycosyltransferases from a structural perspective. *Curr Opin Struc Biol*, 28:131-141.

Gomez H, Polyak I, Thiel W, Lluch JM, Masgrau L. 2012. Retaining glycosyltransferase mechanism studied by QM/MM methods: lipopolysaccharyl-alpha-1,4-galactosyltransferase C transfers alpha-galactose via an oxocarbenium ion-like transition state. *J Am Chem Soc*, 134:4743-4752.

Gordon RD, Sivarajah P, Satkunarajah M, Ma D, Tarling CA, Vizitiu D, Withers SG, Rini JM. 2006. X-ray crystal structures of rabbit N-acetylglucosaminyltransferase I (GnT I) in complex with donor substrate analogues. *J Mol Biol*, 360:67-79.

Hakomori S. 1981. Blood group ABH and Ii antigens of human erythrocytes: chemistry, polymorphism, and their developmental change. *Semin Hematol*, 18:39-62.

Hamasaki N, Yamamoto M. 2000. Red blood cell function and blood storage. *Vox Sang*, 79:191-197.

Heggelund JE, Varrot A, Imberty A, Krengel U. 2017. Histo-blood group antigens as mediators of infections. *Curr Opin Struc Biol*, 44:190-200.

Hollingsworth SA, Karplus PA. 2010. A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins. *Biomol Concepts*, 1:271-283.

Huang JY, Wang R, Gao YT, Yuan JM. 2017. ABO blood type and the risk of cancer - Findings from the Shanghai Cohort Study. *Plos One*, 12:e0184295.

Hurtado-Guerrero R, Davies GJ. 2012. Recent structural and mechanistic insights into post-translational enzymatic glycosylation. *Current Opinion in Chemical Biology*, 16:479-487.

Imberty A, Mikros E, Koca J, Mollicone R, Oriol R, Pérez S. 1995. Computer simulation of histo-blood group oligosaccharides: energy maps of all constituting disaccharides and potential energy surfaces of 14 ABH and Lewis carbohydrate antigens. *Glycoconjugate J*, 12:331-349.

Johal AR, Blackler RJ, Alfaro JA, Schuman B, Borisova S, Evans SV. 2014. pH-induced conformational changes in human ABO(H) blood group glycosyltransferases confirm the importance of electrostatic interactions in the formation of the semi-closed state. *Glycobiology*, 24:237-246.

Johal AR, Schuman B, Alfaro JA, Borisova S, Seto NOL, Evans SV. 2012. Sequence-dependent effects of cryoprotectants on the active sites of the human ABO(H) blood group A and B glycosyltransferases. *Acta Crystallogr D*, 68:268-276.

Kabat EA, Bassett EW, Pryzwansky K, Lloyd KO, Kaplan ME, Layug EJ. 1965. Immunochemical studies on blood groups. XXXIII. The effects. of alkaline borohydride and of alkali on blood group A, B, and H substances. *Biochemistry-Us*, 4:1632-1638.

Kadirvelraj R, Foley BL, Dyekjær JD, Woods RJ. 2008. Involvement of water in carbohydrate-protein binding: concanavalin A revisited. *J Am Chem Soc*, 130:16933-16942.

Kakuda S, Shiba T, Ishiguro M, Tagawa H, Oka S, Kajihara Y, Kawasaki T, Wakatsuki S, Kato R. 2004. Structural basis for acceptor substrate recognition of a human glucuronyltransferase, GlcAT-P, an enzyme critical in the biosynthesis of the carbohydrate epitope HNK-1. *J Biol Chem*, 279:22693-22703.

Kamath VP, Seto NOL, Compston CA, Hindsgaul O, Palcic MM. 1999. Synthesis of the acceptor analog alpha Fuc(1 -> 2)alpha Gal-O(CH2)(7) CH3: A probe for the kinetic mechanism of recombinant human blood group B glycosyltransferase. *Glycoconjugate J*, 16:599-606.

Khalili H, Wolpin BM, Huang ES, Giovannucci EL, Kraft P, Fuchs CS, Chan AT. 2011. ABO blood group and risk of colorectal cancer. *Cancer Epidemiol Biomarkers Prev*, 20:1017-1020.

Kuwabara N, Manya H, Yamada T, Tateno H, Kanagawa M, Kobayashi K, Akasaka-Manya K, Hirose Y, Mizuno M, Ikeguchi M, *et al.* 2016. Carbohydrate-binding domain of the POMGnT1 stem region modulates O-mannosylation sites of alpha-dystroglycan. *Proc Natl Acad Sci U S A*, 113:9280-9285.

Lairson LL, Henrissat B, Davies GJ, Withers SG. 2008. Glycosyltransferases: Structures, Functions, and Mechanisms. *Annual Review of Biochemistry*, 77:521-555.

Lee HJ, Barry CH, Borisova SN, Seto NOL, Zheng RXB, Blancher A, Evans SV, Palcic MM. 2005. Structural basis for the inactivity of human blood group O-2 glycosyltransferase. *J Biol Chem*, 280:525-529.

Leloir LF. 1983. Far away and long ago. *Annu Rev Biochem*, 52:1-15.

Lemieux RU, Bock K, Delbaere LTJ, Koto S, Rao VS. 1980. The conformations of oligosaccharides related to the ABH and Lewis human blood group determinants. *Canadian Journal of Chemistry*, 58:631-653.

Letts JA, Persson M, Schuman B, Borisova SN, Palcic MM, Evans SV. 2007. The effect of heavy atoms on the conformation of the active-site polypeptide loop in human ABO(H) blood-group glycosyltransferase B. *Acta Crystallogr D*, 63:860-865.

Letts JA, Rose NL, Fang YR, Barry CH, Borisova SN, Seto NOL, Palcic MM, Evans SV. 2006. Differential recognition of the type I and II H antigen acceptors by the human ABO(H) blood group A and B glycosyltransferases. *J Biol Chem*, 281:3625-3632.

Lira-Navarrete E, Iglesias-Fernandez J, Zandberg WF, Companon I, Kong Y, Corzana F, Pinto BM, Clausen H, Peregrina JM, Vocadlo DJ, *et al.* 2014. Substrate-guided front-face reaction revealed by combined structural snapshots and metadynamics for the polypeptide N-acetylgalactosaminyltransferase 2. *Angew Chem Int Ed Engl*, 53:8206-8210.

Liumbruno GM, Franchini M. 2013. Beyond immunohaematology: the role of the ABO blood group in human diseases. *Blood transfusion = Trasfusione del sangue*, 11:491-499.

Lobsanov YD, Romero PA, Sleno B, Yu B, Yip P, Herscovics A, Howell PL. 2004. Structure of Kre2p/Mnt1p: a yeast alpha1,2-mannosyltransferase involved in mannoprotein biosynthesis. *J Biol Chem*, 279:17921-17931.

Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*, 42:D490-D495.

Lowary TL, Hindsgaul O. 1993. Recognition of synthetic deoxy and deoxyfluoro analogs of the acceptor α-l-Fucp-(1 → 2)-β-d-Galp-OR by the blood-group A and B gene-specified glycosyltransferases. *Carbohyd Res*, 249:163-195.

Lowe JB. 1993. The blood group-specific human glycosyltransferases. *Bailliere's clinical haematology*, 6:465-492.

Ly HD, Lougheed B, Wakarchuk WW, Withers SG. 2002. Mechanistic studies of a retaining α-galactosyltransferase from Neisseria meningitidis. *Biochemistry-Us*, 41:5075-5085.

Marcus SL, Polakowski R, Seto NOL, Leinala E, Borisova S, Blancher A, Roubinet F, Evans SV, Palcic MM. 2003. A single point mutation reverses the donor specificity of human blood group B-synthesizing galactosyltransferase. *J Biol Chem*, 278:12403-12405.

Meo SA, Suraya F, Jamil B, Rouq FA, Meo AS, Sattar K, Ansari MJ, Alasiri SA. 2017. Association of ABO and Rh blood groups with breast cancer. *Saudi J Biol Sci*, 24:1609-1613.

Monegal A, Planas A. 2006. Chemical rescue of alpha 3-galactosyltransferase. Implications in the mechanism of retaining glycosyltransferases. *J Am Chem Soc*, 128:16030-16031.

Morgan WTJ, Watkins WM. 1959. Some aspects of the biochemistry of the human blood-group substances. *Brit Med Bull*, 15:109-113.

Mukherjee A, Palcic MM, Hindsgaul O. 2000. Synthesis and enzymatic evaluation of modified acceptors of recombinant blood group A and B glycosyltransferases. *Carbohyd Res*, 326:1-21.

Murshudov GN, Vagin AA, Dodson EJ. 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D*, 53:240-255.

Nakahara T, Hindsgaul O, Palcic MM, Nishimura SI. 2006. Computational design and experimental evaluation of glycosyltransferase mutants: engineering of a blood type B galactosyltransferase with enhanced glucosyltransferase activity. *Protein Eng Des Sel*, 19:571-578.

Nguyen HP, Seto NOL, Cai Y, Leinala EK, Borisova SN, Palcic MM, Evans SV. 2003. The influence of an intramolecular hydrogen bond in differential recognition of inhibitory acceptor analogs by human ABO(H) blood group A and B glycosyltransferases. *J Biol Chem*, 278:49191-49195.

Nielsen MM, Suits MD, Yang M, Barry CS, Martinez-Fleites C, Tailford LE, Flint JE, Dumon C, Davis BG, Gilbert HJ, *et al.* 2011. Substrate and metal ion promiscuity in mannosylglycerate synthase. *J Biol Chem*, 286:15155-15164.

Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, Franken SM, Harel M, Remington SJ, Silman I, Schrag J, *et al.* 1992. The alpha/beta hydrolase fold. *Protein Eng*, 5:197-211.

Otter A, Lemieux RU, Ball RG, Venot AP, Hindsgaul O, Bundle DR. 1999. Crystal state and solution conformation of the B blood group trisaccharide alpha-L-Fucp-(1 -> 2)-[alpha-D-Galp]-(1 -> 3)]-beta-D-Galp-OCN3. *Eur J Biochem*, 259:295-303.

Palcic MM, Heerze LD, Pierce M, Hindsgaul O. 1988. The use of hydrophobic synthetic glycosides as acceptors in glycosyltransferase assays. *Glycoconjugate J*, 5:49-63.

Partha SK, Sadeghi-Khomami A, Slowski K, Kotake T, Thomas NR, Jakeman DL, Sanders DAR. 2010. Chemoenzymatic synthesis, inhibition studies, and X-ray crystallographic analysis of the phosphono analog of UDP-Galp as an inhibitor and mechanistic probe for UDP-galactopyranose mutase. *J Mol Biol*, 403:578-590.

Patenaude SI, Seto NOL, Borisova SN, Szpacenko A, Marcus SL, Palcic MM, Evans SV. 2002. The structural basis for specificity in human ABO(H) blood group biosynthesis. *Nat Struct Biol*, 9:685-690.

Pedersen LC, Dong J, Taniguchi F, Kitagawa H, Krahn JM, Pedersen LG, Sugahara K, Negishi M. 2003. Crystal structure of an alpha 1,4-N-acetylhexosaminyltransferase (EXTL2), a member of the exostosin gene family involved in heparan sulfate biosynthesis. *J Biol Chem*, 278:14420-14428.

Persson K, Ly HD, Dieckelmann M, Wakarchuk WW, Withers SG, Strynadka NCJ. 2001. Crystal structure of the retaining galactosyltransferase LgtC from Neisseria meningitidis in complex with donor and acceptor sugar analogs. *Nat Struct Biol*, 8:166-175.

Persson M, Letts JA, Hosseini-Maaf B, Borisova SN, Palcic MM, Evans SV, Olsson ML. 2007. Structural effects of naturally occurring human blood group B galactosyltransferase mutations adjacent to the DXD motif. *J Biol Chem*, 282:9564-9570.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF chimera - A visualization system for exploratory research and analysis. *J Comput Chem*, 25:1605-1612.

Pflugrath JW. 1999. The finer things in X-ray diffraction data collection. *Acta Crystallogr D*, 55:1718-1725.

Pham T, Loiselle D, Power A, Hickey AJR. 2014. Mitochondrial inefficiencies and anoxic ATP hydrolysis capacities in diabetic rat heart. *Am J Physiol-Cell Ph*, 307:C499-C507.

Pham TTK, Stinson B, Thiyagarajan N, Lizotte-Waniewski M, Brew K, Acharya KR. 2014. Structures of complexes of a metal-independent glycosyltransferase GT6 from Bacteroides ovatus with UDP-N-acetylgalactosamine (UDP-GalNAc) and its hydrolysis products. *J Biol Chem*, 289:8041-8050.

Qasba PK, Ramakrishnan B, Boeggeman E. 2005. Substrate-induced conformational changes in glycosyltransferases. *Trends Biochem Sci*, 30:53-62.

Rademacher C, Landstrom J, Sindhuwinata N, Palcic MM, Widmalm G, Peters T. 2010. NMR-based exploration of the acceptor binding site of human blood group B galactosyltransferase with molecular fragments. *Glycoconjugate J*, 27:349-358.

Ramakrishnan B, Boeggeman E, Qasba PK. 2012. Binding of N-acetylglucosamine (GlcNAc) beta1-6-branched oligosaccharide acceptors to beta4-galactosyltransferase I reveals a new ligand binding mode. *J Biol Chem*, 287:28666-28674.

Rini J, Esko J, Varki A. 2009. Glycosyltransferases and glycan-processing enzymes. In: Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME editors. *Essentials of Glycobiology*. Cold Spring Harbor (NY).

Rojas-Cervellera V, Ardevol A, Boero M, Planas A, Rovira C. 2013. Formation of a covalent glycosyl-enzyme species in a retaining glycosyltransferase. *Chem-Eur J*, 19:14018-14023.

Schuman B, Evans SV, Fyles TM. 2013. Geometric attributes of retaining glycosyltransferase enzymes favor an orthogonal mechanism. *Plos One*, 8.

Schuman B, Persson M, Landry RC, Polakowski R, Weadge JT, Seto NOL, Borisova SN, Palcic MM, Evans SV. 2010. Cysteine-to-serine mutants dramatically reorder the active site of human ABO(H) blood group B glycosyltransferase without affecting activity: structural insights into cooperative substrate binding. *J Mol Biol*, 402:399-411.

Seto NOL, Compston CA, Evans SV, Bundle DR, Narang SA, Palcic MM. 1999. Donor substrate specificity of recombinant human blood group A, B and hybrid A/B glycosyltransferases expressed in Escherichia coli. *Eur J Biochem*, 259:770-775.

Seto NOL, Compston CA, Szpacenko A, Palcic MM. 2000. Enzymatic synthesis of blood group A and B trisaccharide analogues. *Carbohyd Res*, 324:161-169.

Sindhuwinata N, Munoz E, Munoz FJ, Palcic MM, Peters H, Peters T. 2010. Binding of an acceptor substrate analog enhances the enzymatic activity of human blood group B galactosyltransferase. *Glycobiology*, 20:718-723.

Sinnott ML. 1990. Catalytic mechanisms of enzymatic glycosyl transfer. *Chem Rev*, 90:1171-1202.

Sobhanifar S, Worrall LJ, King DT, Wasney GA, Baumann L, Gale RT, Nosella M, Brown ED, Withers SG, Strynadka NC. 2016. Structure and Mechanism of Staphylococcus aureus TarS, the Wall Teichoic Acid beta-glycosyltransferase Involved in Methicillin Resistance. *PLoS pathogens*, 12:e1006067.

Sokolowski T, Haselhorst T, Scheffler K, Weisemann R, Kosma P, Brade H, Brade L, Peters T. 1998. Conformational analysis of a Chlamydia-specific disaccharide alpha-Kdo-(2 -> 8)-alpha-Kdo-(2 -> 0)-allyl in aqueous solution and bound to a monoclonal antibody: Observation of intermolecular transfer NOEs. *J Biomol Nmr*, 12:123-133.

Soya N, Fang Y, Palcic MM, Klassen JS. 2011. Trapping and characterization of covalent intermediates of mutant retaining glycosyltransferases. *Glycobiology*, 21:547-552.

Soya N, Shoemaker GK, Palcic MM, Klassen JS. 2009. Comparative study of substrate and product binding to the human ABO(H) blood group glycosyltransferases. *Glycobiology*, 19:1224-1234.

Sturgis CC. 1942. The history of blood transfusion. *Bulletin of the Medical Library Association*, 30:105-112.

Sujino K, Malet C, Hindsgaul O, Palcic MM. 1997. Acceptor hydroxyl group mapping for calf thymus α-(1 → 3) - galactosyltransferase and enzymatic synthesis of α-d-Galp-(1 → 3)-β-d-Gal p-(1 → 4)-βd-GlcpNAc analogs. *Carbohyd Res*, 305:483-489.

Szulman AE. 1962. The histological distribution of the blood group substances in man as Ddsclosed by immunofluorescence : Ii. the H antigen and its relation to A and B antigens. *J Exp Med*, 115:977-996.

Thiyagarajan N, Pham TT, Stinson B, Sundriyal A, Tumbale P, Lizotte-Waniewski M, Brew K, Acharya KR. 2012. Structure of a metal-independent bacterial glycosyltransferase that catalyzes the synthesis of histo-blood group A antigen. *Scientific reports*, 2:940.

Tumbale P, Brew K. 2009. Characterization of a metal-independent CAZy family 6 glycosyltransferase from Bacteroides ovatus. *J Biol Chem*, 284:25126-25134.

Tumbale P, Jamaluddin H, Thiyagarajan N, Brew K, Acharya KR. 2008. Structural basis of UDP-galactose binding by alpha-1,3-galactosyltransferase (alpha 3GT): Role of negative charge on aspartic acid 316 in structure and activity. *Biochemistry-Us*, 47:8711-8718.

Urresti S, Albesa-Jove D, Schaeffer F, Pham HT, Kaur D, Gest P, van der Woerd MJ, Carreras-Gonzalez A, Lopez-Fernandez S, Alzari PM, *et al.* 2012. Mechanistic insights into the retaining glucosyl-3-phosphoglycerate synthase from mycobacteria. *J Biol Chem*, 287:24649-24661.

Vagin A, Teplyakov A. 1997. MOLREP: an automated program for molecular replacement. *J Appl Crystallogr*, 30:1022-1025.

Walsh CT, Losey HC, Freel Meyers CL. 2003. Antibiotic glycosyltransferases. *Biochemical Society transactions*, 31:487-492.

Watkins WM. 1959. Enzymes of trichomonas-foetus - action of cell-free extracts on blood-group substances and low-Molecular-weight glycosides. *Biochem J*, 71:261-274.

Watkins WM. 1972. Structure, genetics and biosynthesis of blood-group-specific glycoproteins. *Biochem J*, 128:114P-116P.

Watkins WM. 1991a. Chemical-structure, biosynthesis and genetic-regulation of carbohydrate antigens - retrospect and prospect. *Pure Appl Chem*, 63:561-568.

Watkins WM. 1991b. Human fucosyltransferases Involved in the biosynthesis of the X-(3-fucosyl-lactosamine) antigenic determinant. *Abstr Pap Am Chem S*, 202:15-CARB.

Watkins WM. 2001. The ABO blood group system: historical background. *Transfus Med*, 11:243-265.

Weadge JT, Palcic MM, Begley TP. 2007. Chemistry of glycosyltransferases. *Wiley Encyclopedia of Chemical Biology*: John Wiley & Sons, Inc.

Winchell KR, Egeler PW, VanDuinen AJ, Jackson LB, Karpen ME, Cook PD. 2016. A structural, functional, and computational analysis of BshA, the first enzyme in the bacillithiol biosynthesis pathway. *Biochemistry-Us*, 55:4654-4665.

Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, *et al.* 2011. Overview of the CCP4 suite and current developments. *Acta Crystallogr D*, 67:235-242.

Yamamoto F, Cid E, Yamamoto M, Blancher A. 2012. ABO research in the modern era of genomics. *Transfus Med Rev*, 26:103-118.

Yamamoto F, Clausen H, White T, Marken J, Hakomori SI. 1990. Molecular Genetic-Basis of the Histo-Blood Group ABO System. *Nature*, 345:229-233.

Yamamoto F, Hakomori S. 1990. Sugar-nucleotide donor specificity of histo-blood group-A and group-B transferases Is based on amino-acid substitutions. *J Biol Chem*, 265:19257-19262.

Yazer MH, Palcic MM. 2005. The importance of disordered loops in ABO glycosyltransferases. *Transfus Med Rev*, 19:210-216.

Zhang H, Zhu F, Yang T, Ding L, Zhou M, Li J, Haslam SM, Dell A, Erlandsen H, Wu H. 2014. The highly conserved domain of unknown function 1792 has a distinct glycosyltransferase fold. *Nat Commun*, 5:4339.

Zhang Y, Swaminathan GJ, Deshpande A, Boix E, Natesh R, Xie Z, Acharya KR, Brew K. 2003. Roles of individual enzyme-substrate interactions by alpha-1,3-galactosyltransferase in catalysis and specificity. *Biochemistry*, 42:13512-13521.

Zhang YN, Swaminathan GJ, Deshpande A, Boix E, Natesh R, Xie ZH, Acharya KR, Brew K. 2003. Roles of individual enzyme-substrate interactions by alpha-1,3-galactosyltransferase in catalysis and specificity. *Biochemistry-Us*, 42:13512-13521.

Appendix I: Journal permissions for GTA/GTB 303 mutants in Chapter 3 and 4

OXFORD UNIVERSITY PRESS LICENSE TERMS AND CONDITIONS

Nov 22, 2018

This Agreement between Susannah Gagnon ("You") and Oxford University Press ("Oxford University Press") consists of your license details and the terms and conditions provided by Oxford University Press and Copyright Clearance Center.

License Number	4474341056800
License date	Nov 22, 2018
Licensed content publisher	Oxford University Press
Licensed content publication	Glycobiology
Licensed content title	Glycosyltransfer in mutants of putative catalytic residue Glu303 of the human ABO(H) A and B blood group glycosyltransferases GTA and GTB proceeds through a labile active site
Licensed content author	Blackler, Ryan J; Gagnon, Susannah M L
Licensed content date	Dec 15, 2016
Type of Use	Thesis/Dissertation
Institution name	
Title of your work	Advancing mechanistic understanding of glycosyltransferases
Publisher of your work	University of Victoria
Expected publication date	Apr 2019
Permissions cost	0.00 CAD
Value added tax	0.00 CAD
Total	0.00 CAD
Title	Advancing mechanistic understanding of glycosyltransferases
Institution name	University of Victoria
Expected presentation date	Apr 2019
Portions	Figures 1-4 Tables I-III Introduction Results Materials and methods Discussion Figure captions
Requestor Location	Susannah Gagnon 3800 Finnerty Road Victoria, BC V8P 5C2 Canada Attn:
Publisher Tax ID	GB125506730
Billing Type	Invoice
Billing Address	Susannah Gagnon 3800 Finnerty Road Victoria, BC V8P 5C2 Canada Attn: Susannah Gagnon

Appendix II: Journal permissions for GTA/GTB and chimera in Chapter 3 and 5

Confirmation Number: 11767528
Order Date: 11/22/2018

If you paid by credit card, your order will be finalized and your card will be charged within 24 hours. If you choose to be invoiced, you can change or cancel your order until the invoice is generated.

Payment Information

Susannah Gagnon
susannahjesse@hotmail.com
+1 (250) 721-8945
Payment Method: n/a

Order Details

Journal of biological chemistry

Order detail ID:	71673764
Order License Id:	4474321416507
ISSN:	1083-351X
Publication Type:	e-Journal
Volume:	
Issue:	
Start page:	
Publisher:	AMERICAN SOCIETY FOR BIOCHEMISTRY AND MOLECULAR BI
Author/Editor:	AMERICAN SOCIETY FOR BIOCHEMISTRY & MOLECULAR BIOL

Permission Status:	Granted
Permission type:	Republish or display content
Type of use:	Thesis/Dissertation
Requestor type	Author of requested content
Format	Electronic
Portion	chart/graph/table/figure
Number of charts/graphs/tables/figures	9
The requesting person/organization	Susannah Gagnon/University of Victoria
Title or numeric reference of the portion(s)	Figures 1-5, Tables 1-4, Glycosyltransferases (GTs) mediate glycoside synthesis, a ubiquitous enzymatic process...providing fresh insight into the glycosyl transfer reaction, Ligands —UDP-C-Gal, l-Fuc-(132)--d-Galp-O(CH2)7CH3 (HA), and -l-Fuc-(132)--d-(3-deoxy)-Galp-O(CH2)7CH3 (DI) were synthesized...all figures were produced with SetoRibbon, Loop Ordering —Data collection and refinement statistics...when in complex with GTA, AABB, ABBA, ABBB, and GTB (DI) and disordered when in complex with GTB (with HA), Four distinct donor conformations were observed... implements significant stepwise shifts to achieve the catalytically active conformation
Title of the article or chapter the portion is from	High Resolution Structures of the Human ABO(H) Blood Group Enzymes in Complex with Donor Analogs Reveal That the Enzymes Utilize Multiple Donor Conformations to Bind Substrates in a Stepwise Manner
Editor of portion(s)	N/A
Author of portion(s)	N/A
Volume of serial or monograph	N/A
Page range of portion	27040-27051
Publication date of portion	September 15 2015
Rights for	Main product
Duration of use	Life of current edition
Creation of copies for the disabled	no
With minor editing privileges	no
For distribution to	Worldwide
In the following language(s)	Original language of publication
With incidental promotional use	no
Lifetime unit quantity of new product	Up to 499

Institution name	University of Victoria
Expected presentation date	Apr 2019

Appendix III: Journal permissions for GTA/GTB 188 and 302 mutants in Chapter 3 and 6

OXFORD UNIVERSITY PRESS LICENSE
TERMS AND CONDITIONS

Nov 22, 2018

This Agreement between Susannah Gagnon ("You") and Oxford University Press ("Oxford University Press") consists of your license details and the terms and conditions provided by Oxford University Press and Copyright Clearance Center.

License Number	4474340709965
License date	Nov 22, 2018
Licensed content publisher	Oxford University Press
Licensed content publication	Glycobiology
Licensed content title	Conserved residues Arg188 and Asp302 are critical for active site organization and catalysis in human ABO(H) blood group A and B glycosyltransferases†
Licensed content author	Gagnon, Susannah M L; Legg, Max S G
Licensed content date	Jun 20, 2018
Type of Use	Thesis/Dissertation
Institution name	
Title of your work	Advancing mechanistic understanding of glycosyltransferases
Publisher of your work	University of Victoria
Expected publication date	Apr 2019
Permissions cost	0.00 CAD
Value added tax	0.00 CAD
Total	0.00 CAD
Title	Advancing mechanistic understanding of glycosyltransferases
Institution name	University of Victoria
Expected presentation date	Apr 2019
Portions	Figures 1-5 Tables I-VI Figure Captions Introduction Materials and methods Discussion Results
Requestor Location	Susannah Gagnon 3800 Finnerty Road Victoria, BC V8P 5C2 Canada Attn:
Publisher Tax ID	GB125506730
Billing Type	Invoice
Billing Address	Susannah Gagnon 3800 Finnerty Road Victoria, BC V8P 5C2 Canada Attn: Susannah Gagnon

Appendix IV: Journal permissions for GTA/GTB product complexes in Chapter 3 and 7

OXFORD UNIVERSITY PRESS LICENSE
TERMS AND CONDITIONS

Nov 22, 2018

This Agreement between Susannah Gagnon ("You") and Oxford University Press ("Oxford University Press") consists of your license details and the terms and conditions provided by Oxford University Press and Copyright Clearance Center.

License Number	4474340914829
License date	Nov 22, 2018
Licensed content publisher	Oxford University Press
Licensed content publication	Glycobiology
Licensed content title	High-resolution crystal structures and STD NMR mapping of human ABO(H) blood group glycosyltransferases in complex with trisaccharide reaction products suggest a molecular basis for product release
Licensed content author	Gagnon, Susannah M L; Legg, Max S G
Licensed content date	Jul 3, 2017
Type of Use	Thesis/Dissertation
Institution name	
Title of your work	Advancing mechanistic understanding of glycosyltransferases
Publisher of your work	University of Victoria
Expected publication date	Apr 2019
Permissions cost	0.00 CAD
Value added tax	0.00 CAD
Total	0.00 CAD
Title	Advancing mechanistic understanding of glycosyltransferases
Institution name	University of Victoria
Expected presentation date	Apr 2019
Portions	Figures 1-5 Tables I-III Introduction Materials and methods Results Discussion Figure captions
Requestor Location	Susannah Gagnon 3800 Finnerty Road Victoria, BC V8P 5C2 Canada Attn:
Publisher Tax ID	GB125506730
Billing Type	Invoice
Billing Address	Susannah Gagnon 3800 Finnerty Road Victoria, BC V8P 5C2 Canada Attn: Susannah Gagnon