Structural and Functional Analysis of Glycosyltransferase Mechanisms

by

Brock Schuman
BSc, University of Victoria, 2006

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

DOCTOR OF PHILOSOPHY

in the Department of Biochemistry & Microbiology

Supervisory Committee

Structural and Functional Analysis of Glycosyltransferase Mechanisms

by

Brock Schuman
BSc, University of Victoria, 2006

Supervisory Committee

Dr. Stephen Evans, Supervisor
(Department of Biochemistry and Microbiology)

Dr. Alisdair Boraston, Departmental member
(Department of Biochemistry and Microbiology)

Dr. Terry Pearson, Departmental member
(Department of Biochemistry and Microbiology)

Dr. Thomas Fyles, Outside member
(Department of Chemistry)

Abstract

Supervisory Committee

Dr. Stephen Evans, Supervisor
(Department of Biochemistry and Microbiology)

Dr. Alisdair Boraston, Departmental member
(Department of Biochemistry and Microbiology)

Dr. Terry Pearson, Departmental member
(Department of Biochemistry and Microbiology)

Dr. Thomas Fyles, Outside member
(Department of Chemistry)

Insight into the biochemical mechanisms utilized by retaining and inverting glycosyltransferase enzymes is an important stepping stone to the directed design of stereospecific inhibitor based drugs.

The suitability of proposed mechanisms was probed using site directed mutagenesis of catalytically relevant residues as well as the use of catalytically inactive substrate analogs UMP-PO_2-CH_2-D-Gal and α-L-Fuc-(1→2)-β-D-(3-deoxy)-Gal-O$(CH_2)_5CH_3$ with the retaining human enzyme a specific α-1,3-N-acetylglucosaminyltransferase (GTA) in conjunction with kinetic and structural approaches including two dozen high resolution X-ray structures and a 2.5 Å resolution neutron structure.

The neutron structure depicts a remarkably non-polar active site which lacks suitably positioned hydrogen atoms to support a dissociative mechanism. Site directed mutagenesis of residues which should be essential to initiate and stabilize a dissociative oxocarbenium ion do not abolish enzyme activity.

The catalytically inactive substrate analogs depict the acceptor nucleophile to lay very close to the anomeric carbon (2.5 Å), which is considerably closer than the closest observed enzymatic dipoles (4.8 Å). This is an indication that the active site architecture is more suited to facilitate a mechanism initiating with nucleophilic attack than dissociation.

To ensure that these observations are applicable to other glycosyltransferases, in depth geometric analysis of all published liganded structures of GT-A fold glycosyltransferase enzymes are reported that display conserved architectures in which the acceptor nucleophile approach is closer than enzymatic dipoles required for dissociation for both inverting and retaining enzymes. Inverting and retaining enzymes present the donor sugar through different conserved geometries about the divalent cation cofactor: all inverting enzymes position the donor for inline nucleophilic attack by the acceptor, the retaining enzymes position the sugar to be attacked from an orthogonal angle.

Such an orthogonal associative mechanism is the most direct proposed approach, and seems supported by all available evidence.

Table of Contents

List of Abbreviations

σ	cross section OR standard deviation
ADA	N-(2-acetamido)iminodiacetic acid
b	scattering length
BME	β -Mercaptoethanol
BSA	bovine serum albumin
CAZy	Carbohydrate Active enZyme databank
CCP4	Collaborative Crystallography Project 4
CD	cluster of differentiation
COOT	Crystallographic Object-Oriented Toolkit
CRL	Chalk River Laboratories
D	isotope ^{2}H
DA	α-L-Fuc-(1→2)-β-D-(3-deoxy)-Gal-O$(CH_2)_5CH_3$
EDTA	ethylenediaminetetraacetic acid
Extl2	α-1,4-N-acetylhexosaminyltransferase
H	isotope ^{1}H
Fm	femtometer
Fuc	L-fucose
Gal	galactose
GalNAc	N-acetylgalactosamine
GalT1	β-1,4-galactosyltransferase I
Glc	glucose
GlcAT-I	galactose 3-β-glucuronosyltransferase I
GlcNAc	N-acetylglucosamine
GNAc:PEP	polypeptide α-N-acetylgalactosaminyltransferase 2
GnT1	β-1,2-N-acetylglucosaminyltransferase I
GT	Glycosyltransferase family
GTA	A specific α-1,3-N-acetylglucosaminyltransferase
GT-A	Glycosyltranfserase fold A
GTB	B specific α-1,3-glucosyltransferase
GT-B	Glycosyltranfserase fold B
H antigen	Fuc-α(1→2)-Gal
HIFR	High Flux Isotope Reactor
HIV	human immunodeficiency virus
lac	lactose operon
ILL	Institut Laue-Langevin
ISIS	Institute for Science and International Security
J-PARC	Japan Proton Accelerator Research Complex
JSNS	Japan Spallation Neutron Source
k	wave number
KDO	Deoxymannooctulosonic acid
Kre2	α-1,2-mannosyltransferase
LANSCE	Los Alamos Neutron Science Center

LgtC	α-galactosyltransferase
ManT	mannosyl-3-phosphoglycerate synthase
Mgs	mannosylglycerate synthase
MpgS	mannosyl-3-phosphoglycerate synthase
MLF	Materials and Life Science Experimental Facility
MW	megawatt
NMR	nuclear magnetic resonance
Nu	nucleophile
PCR	Polymerase Chain Reaction
PCS	Protein Crystallography Station
PDB	Protein Data Bank
PEG	polyethylene glycol
PMSF	phenylmethanesulfonylfluoride
SeMet	selenomethionine
SINQ	Swiss Spallation Neutron Source
S_N1	substitution reaction unimolecular
S_N2	substitution reaction bimolecular
SNS	Spallation Neutron Source
SP	sulfopropyl
SpsA	β-1,3-glucan synthase
TOF	Time of Flight
Tris	tris(hydroxymethyl)aminomethane
U	uridine
UC-Gal	UMP-PO_2-CH_2-Gal
UDP	uridine diphosphate
UMP	uridine monophosphate
UV	ultraviolet
UVic	University of Victoria
vWF	von Willebrand factor

List of Tables

List of Figures

Acknowledgments

I am deeply indebted to my supervisor, Dr. Stephen Evans, without whom this leg of my life would not have been possible. His intimate understanding and zeal for crystallography are unparalleled, and are a gift he passes on to his students. He has been supportive and vigilant far beyond the capacity of most.

Thanks to all members of my committee, especially Dr. Thomas Fyles for helping me look at these enzymes from a new point of view and for helping to convince both others and myself the merit of my insights.

Thanks to all members of the Evans, Pearson and Boraston labs, past and present, with whom I have been very close over the years. Special thanks to Dr. Svetlana Borisova, Dr. Cory Brooks, Dr. Elizabeth Ficko-Blean and Dr. Alicia Lammerts van Bueren for their critical assistance and to Ryan Blackler for providing a safe pun-free buffer zone.

Extra special thanks and all of my love to Dr. Elizabeth Rose Woodward.

Chapter 1. Introduction to Glycosyltransferases

1.1 Overview of glycosyltransferase function — Biosynthesis of simple or complex oligosaccharides and glycoconjugates requires the existence of a repertoire of glycosyltransferase enzymes to catalyze the sequential transfer of sugars to form glycosidic linkages that are specific not only by the chemical moiety they connect to, but also by the stereochemistry of the bond formed. The stepwise transfer takes a single monosaccharide unit from a specific activated donor molecule, a nucleotide-diphospho-sugar in the case of Leloir-type glycosyltransferases, and adds it to a specific acceptor molecule.

Glycosyltransferases specific for acceptors of every macromolecular class have been described, including carbohydrates, polypeptides, nucleic acids, lipids, aromatic compounds and other molecules. Glycosyltransferases are a highly diverse group of enzymes with very few sequence similarities found even among members that share the same donor or acceptor substrates. The growing list of sequenced genomes shows that glycosyltransferases make up at least 1% of defined open reading frames. All forms of life have a bare minimum repertoire of three glycosyltransferases, and with the possible exception of kinases, there are more unique glycosyltransferases identified in most genomes than any other class of enzyme.

1.1.1 Glycosyltransferase classification — The International Union of Biochemistry and Molecular Biology enzyme nomenclature and classification is based on the reaction catalyzed, not the sequence or structure of the enzyme. While this is intuitive and useful

for many enzymes with specific molecular substrates, it is not sufficient to classify glycosyltransferases into a reasonable number of families.

As little as a single point mutation (*eg.* Seto *et al.*, 1997; Seto *et al.*, 1999; Marcus *et al.*, 2003) or conformational shift induced by a regulatory factor (*eg*. Patenaude *et al.*, 2002) can alter enzyme substrate specificity while leaving the enzyme's stereospecific mechanism and underlying structure unchanged.

Glycosyltransferases also display varying degrees of acceptor molecule specificity (Yoshida *et al.*, 2000; Li *et al.*, 2008). As such, the Carbohydrate Active enZyme databank (CAZy) has utilized general amino acid sequence homology to categorize glycosyltransferases into 94 families to date (Campbell *et al.*, 1997; Coutinho, 1999; Coutinho *et al.*, 2003; Rosen *et al.*, 2004). Family rosters are increasing in parallel with genomic revelation, and many glycosyltransferase enzymes have been discovered that are not yet formally classified. Families contain members that appear to share common evolutionary origins and presumably tertiary structure as well. At the time of writing, out of the 94 families, 35 currently have one or more members with determined structure, though only two structures have been determined containing both the acceptor and donor in a bisubstrate complex (Alfaro *et al.*, 2008; Persson *et al.*, 2001).

Functionally, glycosyltransferases have been segregated into "retaining" or "inverting" enzymes, according to whether the stereochemistry of the donor's anomeric bond is retained (*eg.* α→α) or inverted (*eg.* α→β) during the transfer. It has been suggested that retaining and inverting enzymes must require different mechanisms to explain the product stereochemistry, **Fig. 1.1**. The inverting reaction is mechanistically straightforward, requiring only nucleophilic attack on the non-hydroxylated face of the

donor anomeric carbon by the enzymatically deprotonated acceptor, which releases the diphosphonucleotide leaving group and inverts the anomeric center, **Fig. 1.1a**.

1.1.2 Proposed retaining mechanisms — In contrast to inverting enzymes, there has been significant debate concerning the mechanism of the retaining reaction (*eg*. Lairson *et al*., 2008). Suggested mechanisms can be broadly classified as proceeding through the replacement with primarily S_N1 or S_N2 character. Proposed mechanisms are outlined in **Fig. 1.1.**

The first postulated method for retention was the double displacement mechanism (Chelsky & Parsons, 1975), which requires two sequential S_N2 substitutions that invert the anomeric configuration of the donor sugar twice to yield a net retention of stereochemistry. This mechanism requires an initial nucleophilic attack from the enzyme to form a covalent enzyme-glycosyl intermediate that is subsequently attacked by the acceptor, **Fig. 1.1c**. Such a retaining mechanism has thorough enzymatic precedent for many transferase enzymes and carbohydrate acting enzymes such as glycoside hydrolases. This type of mechanism has been demonstrated in other enzymes using bisubstrate kinetics (Pabst *et al.*, 1974) where K_M of one substrate is largely unaffected by the concentration of the second, and by detection of the inverted covalent intermediate which has even been observed crystallographically (*eg.* (Howard *et al.*, 1998).

Because of the absence of corresponding evidence to support glycosyltransferases utilizing a double displacement mechanism, retaining substitution with dissociative (S_N1) character was proposed as an alternative (Martinez-Fleites *et al.*, 2006; Gibson *et al.*, 2002; Pedersen *et al.*, 2003; Lobsanov *et al.*, 2004; Sommer *et al.*, 2004; Reinert *et al.*, 2005; Flint *et al.*, 2005). UDP makes an excellent leaving group and donor sugar **C1** is

unsaturated which are both attributes favoring S_N1 substitution. However, the rate-determining unimolecular dissociation of RX → R^+ + X^- canonically loses product stereochemistry and leads to racemisation or partial inversion.

No racemic products have ever been described for a glycosyltransferase, and other enzymes which are understood to proceed by a dissociative mechanism are generally either hydrolytic or transferases that do not transfer stereocenters (reviewed in Nagano *et al.*, 2007). Nevertheless, advocates of an S_N1 pathway hypothesize that steric hindrance provided by the enzyme may allot room only for retained product generation (Persson *et al.*, 2001).

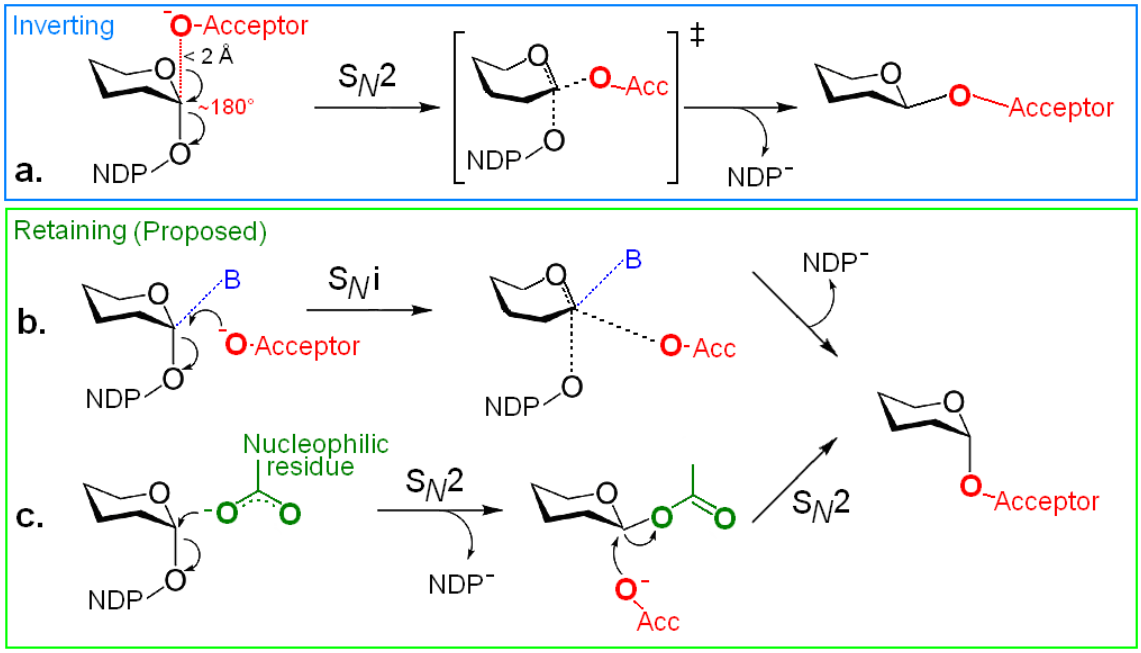


Fig. 1.1 Mechanisms proposed for glycosyltransferases. **(a)** Inverting enzymes promote a typical S_N2 nucleophilic attack on **C1** by the acceptor from an inline (~180°) position, with resulting inversion of the anomeric bond stereochemistry. Retaining enzymes have several postulated mechanisms, most popularly **(b)** S_Ni which uses the acceptor as an "internal return" intermediate and **(c)** a double displacement mechanism with two sequential S_N2 inversions leading to net retention of stereochemistry. Adapted from Schuman *et al.*, submitted to *PNAS*.

One widely accepted S_N1 variant, S_Ni (**Fig. 1.1b**), would require nucleophilic attack by the acceptor with concomitant release of NDP, with both bound to the anomeric carbon to form an "internal return" oxocarbenium ion intermediate (Sinnott & Jencks, 1980).

This is usually drawn as a standard S_N2 intermediate. Internal return was first proposed to describe the intramolecular return and anomeric stereochemistry retention of glucose derivatives during trifluoroethanol solvolysis in the gas phase (Sinnott & Jencks, 1980), and its acceptance as a suitable pathway for retaining glycosyltransferases has met with considerable resistance. Even Sinnot who first proposed the mechanism has warned "the evidence in favor of the internal return mechanism … is essentially negative" (Sinnott, 1990).

It has also been suggested that retaining transfer may contain both S_N1 and S_N2 elements (Lairson *et al.*, 2008), and to be sure the two are not mutually exclusive. Absolute distinction between associative and dissociative reaction pathways is not always possible: S_N2 pathways overlap into S_N1 pathways as the transition state develops a longer carbocation lifetime (Katritzky & Brycki, 1990).

There are several established techniques to differentiate some of these proposed mechanisms: Michaelis–Menten kinetics and kinetic isotope effect data can rule out a double displacement mechanism.

Other observations addressed later do not support a dissociative mechanism, which will prompt a novel mechanism proposal.

1.1.3 General glycosyltransferase structure — Most glycosyltransferases are peripherally or integrally membrane-associated, often assembled as functional dimers or multimers, which are challenges contributing to the paucity of glycosyltransferase structures deposited in the PDB. This obstacle can often be overcome by mutation or deletion of the membrane associated region (reviewed in Radominska-Pandya *et al*., 2005).

Mammalian glycosyltransferases pose a further problem in generally observed poor expression in bacteria, as well as abnormal glycosylation and folding in insect cell expression systems leading to inactive protein (Pak *et al.*, 2006). Only a handful of families with structural determination also have comprehensive studies of kinetics and specificity, in many cases largely due to the difficulty in synthesizing substrates.

The modest degree of sequence homology within and among the various glycosyltransferase families has made the prediction of tertiary structures difficult; however, structural determinations through single crystal x-ray diffraction have revealed that the catalytic domains of most glycosyltransferases which utilize nucleotide-sugars as the donor molecule display one of two fold types designated GT-A or GT-B (Bourne & Henrissat, 2001; Coutinho *et al.*, 2003). Examples of retaining and inverting enzymes have been observed in both the GT-A and GT-B type folds, so the fold is not determinative of stereospecificity.

Examples of the catalytic domains of the GT-A and GT-B fold types are presented in **Fig. 1.2**, where GT-A and GT-B fold types consist of two closely associated domains, at least one of which contains a Rossmann fold responsible for donor nucleotide recognition. The Rossmann fold is a common metal cation-dependent nucleotide binding domain characterized by β-sheets flanked by α-helices (Dodson *et al.*, 1966), and it is a

ubiquitous structural motif among glycosyltransferases and many other enzymes that utilize nucleotide substrates.

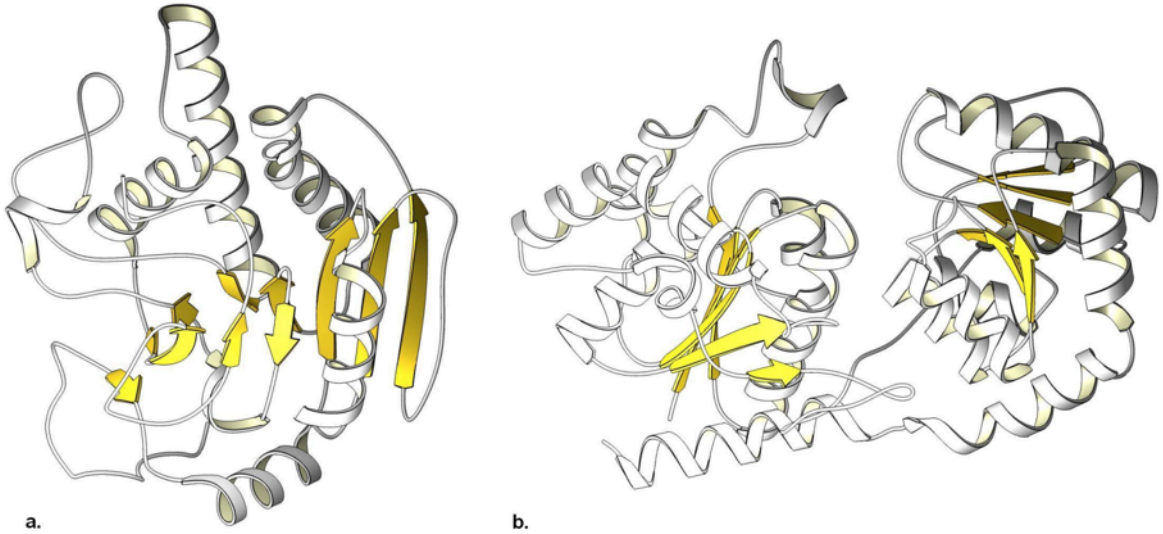


Fig. 1.2 Glycosyltransferase folds. SETOR (Evans, 1993) diagrams of the **(a)** GT-A fold type showing human ABO(H) blood group A enzyme with a single Rossmann fold, and **(b)** GT-B fold type with two Rossmann folds showing vancomycin acting enzyme GtfB. In both enzymes the active catalytic site is housed in a cleft between the two domains, albeit much more cavernous and apparent in the GT-B fold. Reproduced from (Schuman *et al.*, 2007) with permission (**Appendix I**).

The Rossmann fold constitutes a dominant portion of the catalytic centers in the cleft between the two domains and often contains much of the limited sequence homology that is observed across many glycosyltransferase families, likely due to a finite repertoire of donor nucleotides utilized (Campbell *et al.*, 1997). Though there are exceptions, most donor binding Rossmann folds contain a "DXD motif" that consists of an Asp-X-Asp amino acid triplet with which to coordinate the donor molecule's phosphates through a divalent cation. It is noteworthy that some inverting enzymes do not require a divalent metal cofactor, though to date there is only one retaining Leloir-type enzyme that has been characterized as metal independent (Tumbale and Brew 2009).

In the GT-A type fold, the donor binds the N-terminal domain's Rossmann fold and typically one or two of this domain's β-sheets extend into the C-terminal domain, forming a continuous β-sheet which can render unambiguous distinction of the domains difficult. The second domain is primarily responsible for acceptor recognition, though its C-terminal tail (which is often disordered) has been demonstrated for many enzymes to fold in and interact with the donor.

The C-terminal domain has greater sequence and structural variability among the different families than does the N-terminal nucleotide-binding domain, due in part to the increased structural variability of acceptor molecules. Different motifs have been observed for the C-terminal domain, including a second Rossmann fold and all α-helix. The GT-B type fold has two Rossmann or Rossmann-like folds in distinct domains which are separated by a deep wide crevice while the enzyme is in an unliganded "open" conformation.

Both fold families of enzymes have been suggested to have step-wise reaction mechanisms, in which substrate binding the open enzyme induces a conformational shift by main chain rotations to generate the closed conformation, aligning the substrates in their bioactive conformation for catalysis (Alfaro *et al.*, 2008; Unligil *et al.*, 2000; Coutinho *et al.*, 2003; Mulichak *et al.*, 2001).

1.2 Implications for disease management

Glycosyltransferases synthesize biological oligo- and polysaccharides, many of which have been associated with a myriad of disease processes. In addition to being the specific defect in a number of human genetic disorders (*eg.* (Pastores *et al.*, 2007; Yoshida *et al.*, 2001; Wennekes *et al.*, 2007), glycosyltransferases play critical roles in

almost all facets of infection (*eg.* Ma *et al.*, 2006; Umesiri *et al.*, 2010; Mas *et al.*, 1998; Mulichak *et al.*, 2001; Raetz & Whitfield, 2002; Harrington *et al.*, 2002; Wong *et al.*, 2011), immunity (*eg.* Freiberger *et al.*, 2007; Byrne *et al.*, 2011; Weil *et al.*, 1973) and cancer (*eg.* Hakomori, 1996; Gschaidmeier *et al.*, 1995; Yamamoto *et al.*, 1997; Werther, Rivera-MacMurray*, et al.*, 1994; Ravindranath *et al.*, 1991; Mathieu *et al.*, 2007; Marionneau *et al.*, 2002; Itzkowitz, Dahiya*, et al.*, 1990; Buzzi & Buzzi, 1974).

For each of these major domains of health research, one or many specific glycosyltransferase enzymes are implicated to cause or facilitate disease progression, and establishment of the underlying biochemical mechanism of these enzymes is an indispensable stepping stone towards directed drug design for a vast array of biomedical therapeutics. The role of glycosyltransferases in cancer, immunity and infection is not surprising as these processes all involve extracellular functions and nearly all extracellular proteins are glycosylated, often in a cell-type-dependent manner. All cluster of differentiation markers are either glycosylated proteins, or specific complex carbohydrates themselves (The latter category includes CD60, CD65, CD75, CD77 and CD175).

1.2.1 Glycosyltransferases in cancer progression — Specific glycosyltransferases and carbohydrates are known and in current clinical use as indicators of cancer progression. For example, the modulation of cell-surface glycoconjugates in neoplastic transformation has been attributed to the differential expression of various glycosyltransferases, and the relative concentrations of different gangliosides has long been used both in tissue histology and in serum testing to track the progress of cancers including breast, melanoma, gastric, colorectal, leukemia and liver metastases (Blixt *et al.*, 2011;

Itzkowitz, Dahiya*, et al.*, 1990; Itzkowitz, Bloom*, et al.*, 1990; Ravindranath *et al.*, 1991; Werther, Riveramacmurray*, et al.*, 1994; Cao *et al.*, 2008).

The change in relative abundance of many glycoconjugates on the cell surface during cancerous mutation arises due to differential expression of various glycosyltransferases, and treatment with unique oligosaccharides or *via* their destruction provides multiple possible therapeutic avenues (Albino *et al.*, 1986; Carubia *et al.*, 1984). Given the prominent role of glycoconjugates in cell adhesion and signalling, the differential expression and mutation of specific glycosyltransferases have been linked to tumor metastasis (Fidler *et al.*, 1978; Ciolczyk-Wierzbicka *et al.*, 2007; Kannagi, 1997), invasion (Yamamoto *et al.*, 1997) and angiogenesis (Gill *et al.*, 1987; Shima *et al.*, 1995; O'Donnell & Laffan, 2001).

1.2.2 Glycosyltransferases in diagnostics and drug design — Mammalian glycosyltransferases utilize only 9 monosaccharide donors in oligosaccharide biosynthesis (**Table 1.1**, left), making pathogen glycosyltransferases that utilize other donors (*e.g.* **Table 1.1**, right) excellent prospective targets for therapeutic inhibition. Interestingly, animal genomes contain glycosyltransferases able to synthesize both α and β-linked products of only 7 of the 9 utilized monosaccharide residues: glucuronic acid and N-acetylneuraminic acid are only found as the β stereoisomer.

Table 1.1 Biological monosaccharides. Mammals utilize exclusively the 9 monosaccharides listed, while many pathogenic bacteria utilize alternative monosaccharides. Bacterial enzymes using these sugars are potential targets for therapeutic inhibition.

The 9 Mammalian Monosaccharides		***Exclusively Bacterial Monosaccharides***
Glucose	N-acetylglucosamine	Deoxymannooctulosonic acid (KDO)
Galactose	N-acetylgalactosamine	Heptose
Mannose	N-acetylneuraminic acid	Rhamnose
Fucose	Glucuronic acid	Muramic acid
Xylose		Diacylglucosamine

In nature, N-acetylneuraminic acid is always in the β conformation even as a monosaccharide. This is imposed by the anomeric carbon's axially linked carboxylic acid. There is, however, no such stereochemical constraint for glucuronic acid which is also only β-linked by animals (as found in heparan sulphate). Glucuronic acid is found α-linked as glucuronoxylans in conifers, monocots and some phytopathogenic bacteria (Scheller & Ulvskov, 2010).

There are many human pathogenic bacteria species with related glycosyltransferase genes with as-of-yet-unidentified function. For example, many pathogens possess multiple uncharacterized glycogenin-like GT-8 enzymes; GT-8 enzymes have been described with glucosyl-, glucuronyl- and galactosyltransferase activity so the donor sugar identity cannot be assumed.

In addition to the use of novel donors, many bacterial glycosyltransferases recognize unique acceptor molecules which can be antibiotic targets *eg*. vancomycin which inhibits N-acetylglucosaminyltransferase by blocking the acceptor (Taku *et al.*, 1980).

The detailed understanding of glycosyltransferase substrate recognition at atomic resolution is a key element of directed drug design. The study of the structure and

function of glycosyltransferases involved in oligosaccharide biosynthesis is essential to understanding and exploiting them for diagnostic and therapeutic purposes.

1.3 The human ABO(H) blood group

1.3.1 Human ABO(H) blood group antigens — The human ABO(H) blood group antigens have a minimal carbohydrate epitope consisting of the H antigen disaccharide, Fuc-α(1→2)-Gal, which corresponds to the human ABO(H) blood type O and is generated by fucosylation of appropriate galactose or galactoside molecules. The H antigen is α1→3 N-acetylgalactosylated to generate the A antigen (**Fig. 1.3a**) or α1→3 galactosylated to generate the B antigen.

The O-blood type is the result of non-functional ABO(H) glycosyltransferases, most commonly a 261delG polymorphism which codes a truncated and unexpressed product, though other mutations exist (reviewed in Yazer & Palcic, 2005; Yazer *et al.*, 2008).

The A and B antigen trisaccharide termini are constituents of much larger common complex carbohydrates, where AB blood group antigen subtypes a, b, c and d have 6-14 monosaccharide constitutes with 1 or 3 branch points (**Fig 1.3b**). The proximity of the α1→2 and α1→3 glycosidic linkages to Gal makes a large contiguous surface area for adherence.

These two sugars differ only by the acetamido group of the A antigen, but this small difference is sufficient to cause a lethal agglutinating immune response in a mismatched blood transfusion, which was the reason for their early discovery (Landsteiner, 1901) and remains crucial to their biomedical relevance.

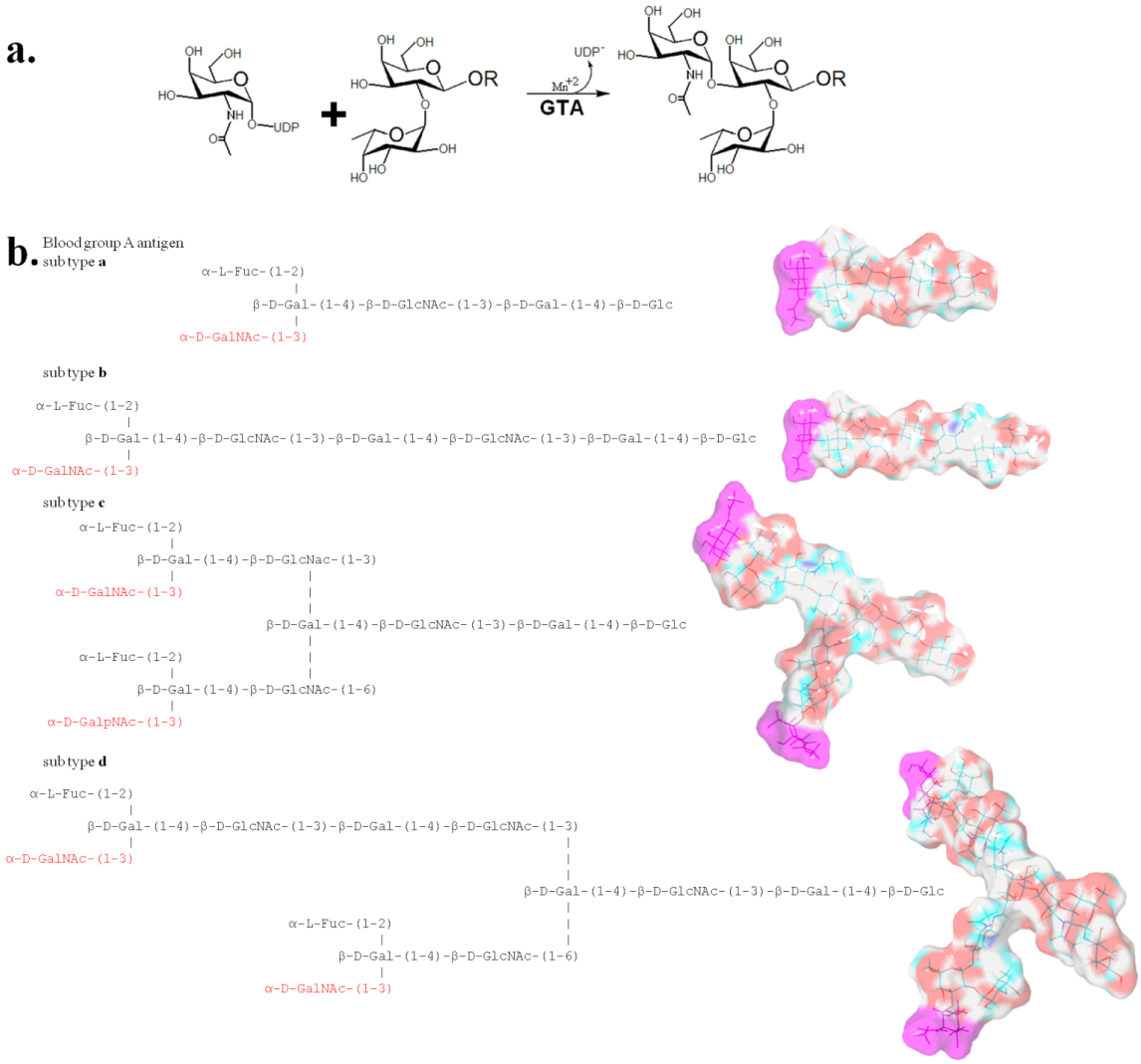


Fig. 1.3 The ABO(H) blood group A antigen. **a.** Biosynthesis of human blood group A antigen by GTA. N-acetylgalactosamine is transferred in the presence of a Mn^{+2} cofactor from UDP to Gal O3 of the H antigen (minimal epitope Fuc-α(1→2)Gal) with retention of the axial α anomeric stereochemistry. **b.** Structures of the 4 common human subtypes of A antigen carbohydrates, with the terminal A antigen GalNAc residues highlighted in red(text)/magenta(3D structure).

The blood group antigens are nearly ubiquitously expressed cell surface oligosaccharides found conjugated as glycoproteins and glycolipids (Watkins & Morgan, 1957; Finne *et al.*, 1980). They can also be found as soluble oligosaccharides in serum, saliva and the mucosal membranes of individuals with secretor status which is determined by a fucosyltransferase able to use soluble galactosides as acceptor substrates (Chester & Watkins, 1969).

Other soluble forms of the blood group antigens can be found in serum, including moieties of soluble glycoproteins and BSA-docked glycolipids, which may be incorporated into membranes of other cells and tissues which may not themselves produce the antigens. The most abundant glycoproteins decorated with these antigens are CD-31, PV-1, and von Willebrand factor (Tasaki *et al.*, 2009).

Von Willebrand factor (vWF) is a good representative of the protective function of these large branched complex carbohydrates. VWF participates in hemostasis and is a dominant constituent of endothelial Weibel-Palade bodies, megakaryocytes, and subendothelial connective tissue (Sadler, 1998; Sarode *et al.*, 2000). It is heavily glycosylated with ABO(H) blood group glycans along 12 N-glycosylation sites and 10 O-glycosylation sites, increasing the observed molecular weight of the ~300 KDa protein to over 500 KDa (Lenting *et al.*, 2010).

VWF oligomerizes through disulfide bonds into helical structures formed by hundreds of units, observed in formations up to 0.5 mm in length which bridge platelets to collagen at wound sites. Concatemers become susceptible to metalloproteolysis by ADAMTS13 only after tensile forces imparted by bridging platelets (via an A1 domain) and collagen (via an A3 domain) unravel the A2 domain tertiary structure to expose the buried cleavage site (O'Donnell *et al.*, 2005; Zhou *et al.*, 2011).

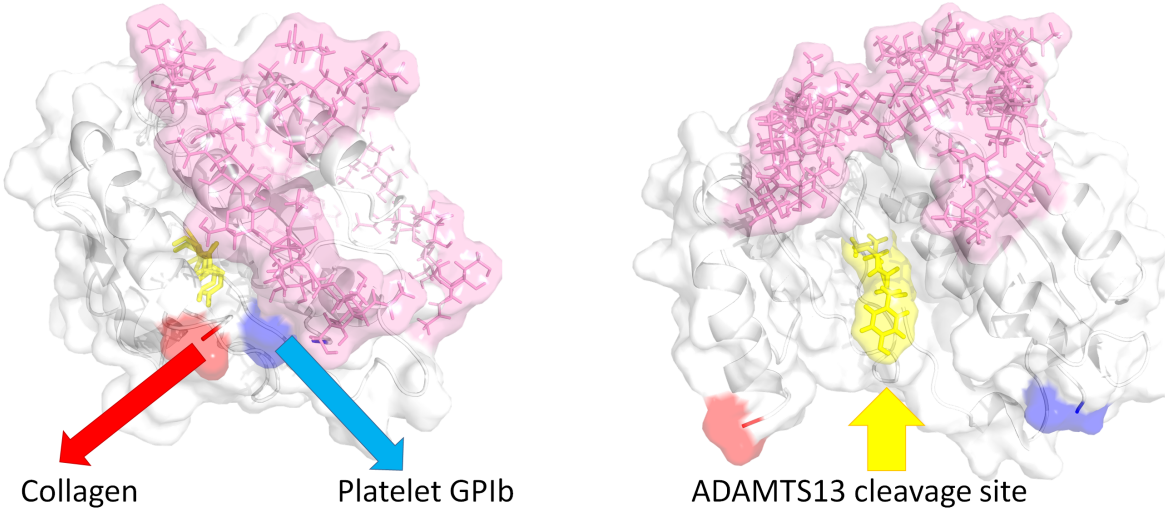


Fig 1.4 Protective function of blood group antigens. 2 N-linked A antigen moieties (pink) shielding the vWF A2 domain from mechanical forces applied by N-terminal A1 domain bound platelet GPIb anchors (Blue arrow) and C-terminal A3 domain binding collagen (Red Arrow). Protraction exposes the buried Tyr1605-Met1606 cleavage site (yellow) ADAMTS13 requires for wound clearance. Model built from PDB 3GXB.

As such, the role of the large branched ABO(H) antigens can be envisioned to prolong clot clearance by tethering and physiochemically protecting the A2 domain's tertiary structure from shear force (**Fig 1.4**). VWF molecules which lack the A and B antigens are cleared much more rapidly than are those with, suggesting that the additional surface area the A and B antigen monosaccharides impart contributes a great deal of stability (Gill *et al.*, 1987; Shima *et al.*, 1995; O'Donnell & Laffan, 2001).

Given its high serum abundance, the natural role of bridging serum components, and the high avidity provided by 22 glycosylation sites (each of which could potentially contain 2 blood group antigens if glycosylated with subtype c or d) vWF undoubtedly plays a central role in clinically catastrophic blood transfusion mismatch agglutination.

Mutations to the galactoside fucosyltransferase Fut2 can prevent formation of the H-antigen. This is known as the "Bombay blood group" phenotype, named after its discovery in populations of present day Mumbai (Hakim *et al.*, 1961). These mutations cause no major reported deleterious developmental or biomedical effects, but their effects do include increased susceptibility to eukaryotic pathogens, reduced plasma vWF and misregulated hemostasis (O'Donnell *et al.*, 2005).

This evidence, in conjunction with their ubiquitous distribution, indicates that ABO(H) blood group antigens can only have a limited impact on cellular communication or cell specific interactions. The blood group antigen's primary biological role is most likely protective.

A proposed indirect role of the antigens is to impart communal immunity between blood group mismatched individuals to enveloped viruses such as influenza and HIV. As these viruses bud in host cell membranes, their newly-formed envelopes will likely contain ABO(H) blood antigens which are particularly abundant in the mucosal membranes from where such viruses will shed. In an ABO-mismatched secondary host, the ABO(H) antigens present on the viral envelope may elicit a rapid host immune response which may work to minimize the spread of such viruses.

The humoral immune response to mismatched antigens is very strong as these antigens are utilized for adhesion and immunoevasion by common potential pathogens

(including *Campylobacter, Escherichia, Helicobacter, Salmonella, Shigella, Streptococcus, Vibrio* and *Yersinia)*. As such, exposure to the antigens is inevitable.

This may explain the high frequency (~44% globally according to bloodbook.com/) of the loss-of-function O-blood type mutation: O-type individuals would have some degree of established immunity during genetic bottleneck epidemics. If so, the ABO(H) antigens have played a role throughout human evolution.

1.3.2 ABO(H) blood group glycosyltransferases — Human retaining glycosyltransferases GTA and GTB (named long before the GT-A and GT-B fold types were described) are responsible for the retaining glycosyltransfer generation of the human ABO(H) blood group A and B antigens. The O blood group which corresponds to the H antigen is the result of an inactive, usually truncated, GTA or GTB enzyme (Lee *et al.*, 2005).

GTA and GTB are the two most homologous naturally occurring glycosyltransferases reported that utilize distinct naturally occurring donors, differing by only 4 critical amino acid substitutions: Arg/Gly-176, Gly/Ser-235, Leu/Met-266 and Gly/Ala-268 (Yamamoto *et al.*, 1990).

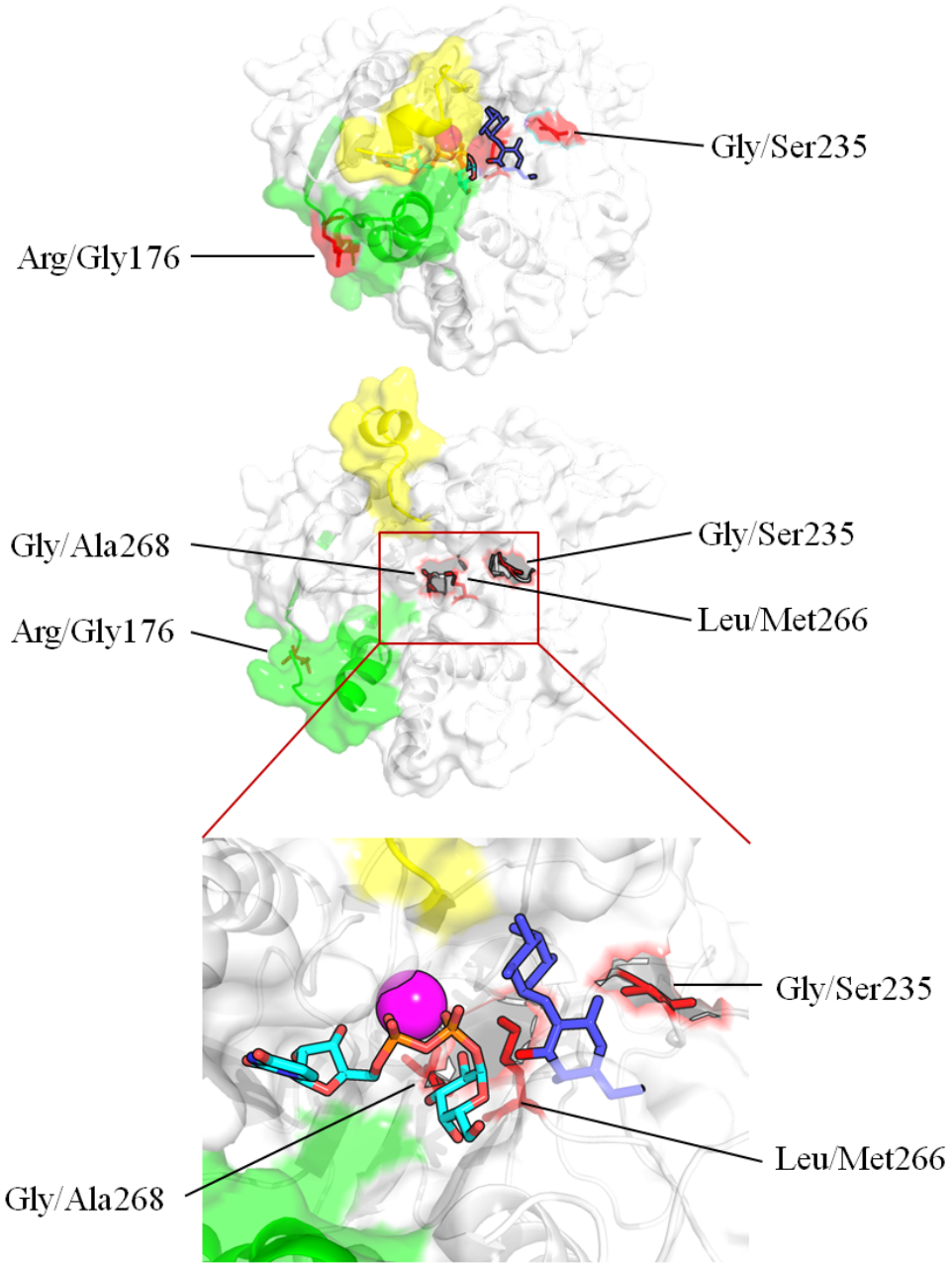


Fig. 1.5 Critical residues and mobile loops. Two large sections of the polypeptide chain form a closed conformation (top) to encase ligands Mn^{+2} (magenta), UDP-GalNAc (cyan) and the H-antigen (blue, O3 acceptor red): an internal loop comprised of residues 175-195 (green) and a C-terminal tail comprised of residues 346-354 (yellow). These loops must open (bottom) to allow substrate entry and product release. The four critical GTA/GTB distinguishing amino acids are indicated in red. Only Leu/Met266 and Gly/Ala268 are positioned to directly distinguish the C2 N-acetyl/hydroxyl of the GTA/GTB donor substrates. Modeled from PDB 2RJ7.

Arg/Gly-176 is found in the N-terminal UDP-donor binding Rossmann fold, while the other three are found on the C-terminal fold. Only residues 266 and 268 of the four critical amino acids are positioned to directly interact with the donor, and so directly confer specificity (Patenaude *et al.*, 2002). Arg/Gly-176 imparts some structural flexibility to an internal mobile loop it defines, and Gly/Ser-235 imparts less or more steric hindrance to the acceptor molecule. Like many transferases, GTA contains spans of mobile polypeptide which are not structurally observed in the absence of substrate, only becoming ordered in the presence of donor and acceptor substrates (**Fig. 1.5**).

With the two wild type enzymes differing by only four amino acids, it is possible to generate all 14 recombinant chimera permutations. It is convenient when discussing the GTA/GTB chimera to designate the four critical amino acid residues of GTA as AAAA and the four critical amino acid residues of GTB as BBBB.

Surprisingly, kinetic studies of all 14 chimera and the two wild type enzymes showed (in advance of structural determination) that of the four critical amino acids, only the final two (Leu/Met-266 and Gly/Ala-268) significantly affect donor specificity (Seto *et al.*, 1999). That is, any chimera XXAA (where X can be either A or B) would selectively transfer A donor (UDP-GalNAc), and any chimera XXBB would selectively transfer the B donor (UDP-Gal). Similarly, chimera that mixed the final two critical amino acid residues had mixed A and B activity. That is to say, any chimera XXAB or XXBA would transfer both the A and the B donor.

Although the first two critical amino acids (Arg/Gly-176 and Gly/Ser-235) do not affect donor specificity, they clearly impact enzyme turnover. It seems these residues impart stability to alter physiological turnover: right at the N-terminal end of the internal

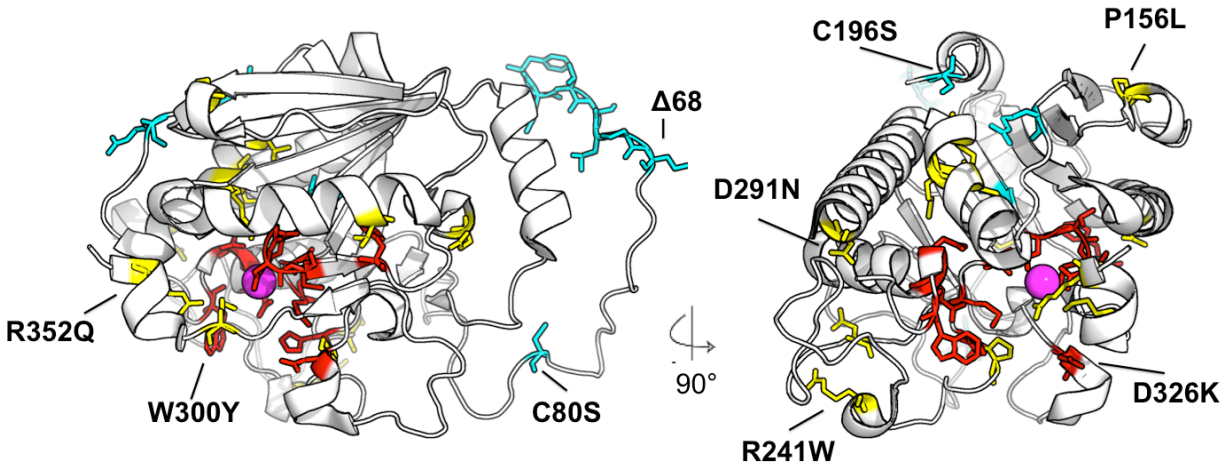
mobile loop for GTA to favor closure to overcome N-acetyl group steric hindrance and near the acceptor binding site for GTB to force the acceptor into the larger binding pocket. The chimeras BAAA, BBAA, ABAA all have GTA activity exceeding that of wild type GTA (AAAA) enzyme. The chimera BAAA has an 11-fold increase in k_{cat}, which remains the highest increase in k_{cat} reported for a single amino acid mutation.

There is a bank of over 300 expressing mutants of the human ABO(H) blood group glycosyltransferases, making them one of the most structurally and kinetically characterized glycosyltransferases in the literature (*eg.* Seto *et al*., 1999; Patenaude *et al*., 2002; Nguyen *et al*., 2003; Lee *et al*., 2005; Letts *et al*., 2006; Persson *et al*., 2007; Letts *et al*., 2007; Alfaro *et al*., 2008; Schuman *et al*., 2009; Schuman *et al*., 2011; Schuman *et al*., 2012; Johal *et al*., 2012). With the exception of glycogen phosphorylase there are more crystal structures of human ABO(H) blood group glycosyltransferases already deposited in the PDB than any other glycosyltransferase, making it a well-characterized model enzyme for the study of retaining glycosyltransfer.

Two rationales from determined structures are employed to commence structural and kinetic mutagenic analysis of residues identified: active site residues thought to be involved in catalysis and/or substrate recognition (*eg.* Glu303); and residues thought to be involved in crystallizability (*eg.* Cys209). A third rationale for mutations to analyse are clinically identified human mutations distal from the active site that produce weak phenotypes.

Global human populations with unusual phenotypes can be panned as every sample of blood destined for transfusion is tested for blood type. Those that show abnormal phenotypes such as AB_{weak}, where an individual with type A blood can display weak but

detectable levels of B antigen agglutination, are often sequestered and sequenced to reveal the location and nature of the polymorphism. Only active mutants with abnormal activity arise in the established screening process, as any mutation that renders the enzyme inactive yields blood type O. The location and rationale of some of these investigated mutations is shown in **Fig. 1.6.**



P74S, C80S, T136M, P156L, R168G, R176(A,H), R180H, S185(A,C,E,G,K,N,R), R188(E,F,H,K,S,Y), M189V, I192T, C196(A,S), C209(A,S), D211E, V212G, D213E, M214(G,R,S,T,V), F216I, E223D, H233(E,F,N,Q), P234(C,L,R,S,T), R241W, D262N, L266(G,R), G268(N,T), D291(E,N), W300(F,Y), D302(A,C,E,L), E303(A,C,D,L,Q), D326(F,K,R,W,Y), A343(F,K,R,W,Y),K346EM and R352RQ.

Fig. 1.6 Locations of mutations. Identified from panning human blood banks (yellow), analysed to improve crystallizability (cyan) or to probe the mechanism and substrate recognition (red, unmarked in sequence). The Mn^{+2} (magenta sphere) position indicates the substrate-binding pocket. Modeled from PDB 2RJ7.

1.4 Neutron diffraction for protein crystal structure analysis

Insight into enzymatic mechanisms, especially those that involve evolution, absorption or shuttling of protons, can be gained from analysis of active site hydrogen geometries including ordered solvation, hydrogen bonding patterns and imposed charges to substrates and amino acid residues. Neutron diffraction has been successfully and repeatedly implemented for this purpose (Kossiakoff & Spencer, 1981; Niimura & Bau,

2008; Bennett *et al.*, 2006; Blakeley *et al.*, 2008; Coates *et al.*, 2008; Blum *et al.*, 2009; Adachi *et al.*, 2009; Yagi *et al.*, 2009; Tomanicek *et al.*, 2010).

Water and enzyme hydrogen positions or even their presence cannot always be unambiguously determined by the chemical environment in which they reside by X-ray data alone, and are rarely directly observable at the resolution usually obtainable for protein crystals: hydrogen atoms, with between one and zero electrons in a biological setting, diffract X-rays poorly and are effectively invisible under normal experimental conditions. As such, the ionization states of residues possibly involved in hydrogen atom transfer and details concerning the hydrogen-bonding arrangement are often ambiguous.

1.4.1 Principles of neutron diffraction — Unlike x-rays, neutrons are not (directly) scattered by electrons. They are instead scattered by atomic nuclei. Hydrogen nuclei diffract strongly, and with negative scattering length, which is a measurement combining scattering strength and scattered phase. This is a useful quantification because for some atoms, such as H, the angular momentum vector imparted by nuclear spin shifts the observed phase of the diffracted neutron $^{1}/_{\lambda}$ compared to diffraction by other atoms. This is a ramification of polarizing the spin state of the diffracting neutron (Sears, 1992). Scattering length is defined as **b** in the relations:

$$\mathbf{b} = \sqrt{\frac{\sigma}{4\pi}} \qquad\qquad \boldsymbol{k}\cot\boldsymbol{\delta}(\boldsymbol{k}) = -\frac{\mathbf{1}}{\mathbf{b}}$$

where $\boldsymbol{k}$ is the wave number, $\boldsymbol{\delta(k)}$ is the s-wave phase shift and $\boldsymbol{\sigma}$ is the scattering cross section, the area surrounding the diffracting matter which effectively scatters. For the diffraction of X-rays, $\boldsymbol{\sigma}$ and **b** correlate with the number of electrons and are not appreciably affected by different isotopes.

This is not true for neutrons as the effective scattering area is orders of magnitude larger than the size of the nucleus: the precise diameters of atomic nuclei is a matter at the cutting edge of particle physics, but the current estimates of ~1.8 fm diameter for H (Pohl *et al.*, 2010) would produce a cross section of only 0.0023 barn, which is orders of magnitude smaller than the observed 1.75 barn cross section.

Neutron scattering cross sections have both an isotope sensitive nuclear component (mediated by the nuclear strong force), and a magnetic component (mediated by magnetic moments including spin states and atom oxidation). Neutron **σ** and **b** must be determined experimentally as current models of nuclear force are insufficient to calculate or predict them (Schober, 2009). A sample list of coherent X-ray and neutron scattering lengths and cross sections is provided as **Table 1.2.**

Table 1.2 Atomic scattering factors as reported in Sears, 1992

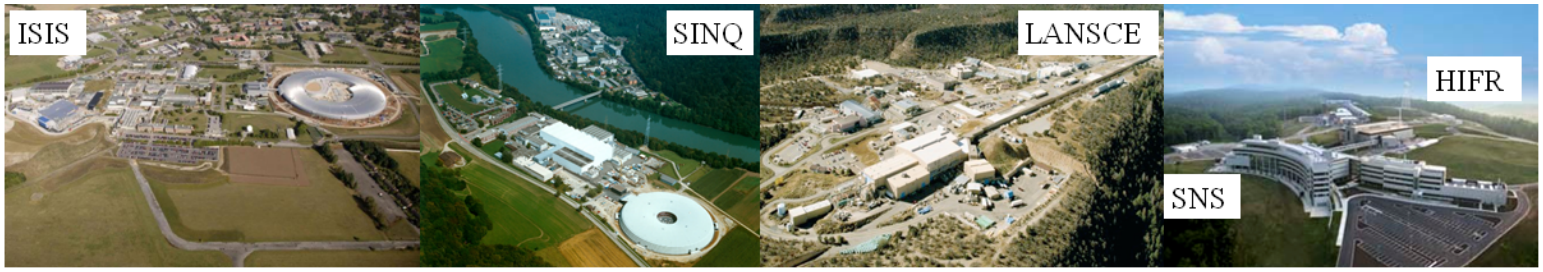
	Neutron		X-ray
	σ (barn)	**b** (fm)	**σ** (barn)
^{1}H	1.76	-3.74	1.67
^{2}H	5.59	6.67	1.67
^{12}C	5.56	6.65	19.94
^{13}C	4.81	6.19	19.94
^{14}N	11.03	9.37	23.26
^{15}N	5.21	6.44	23.26
O	4.23	5.8	26.57
^{32}S	0.99	2.8	53.25
^{34}S	4.21	3.48	53.24
^{46}Ti	3.05	4.93	79.53
^{48}Ti	4.65	-6.08	79.53
^{55}Mn	1.75	-3.73	91.22
^{74}Se	0.10	0.80	131.1
^{80}Se	7.03	7.48	131.1
^{197}Au	7.32	7.63	327.4

1.4.2 Isotopic labelling and use of X-ray data in joint refinement — The negative scattering of ^{1}H atoms can be used to model their positions; however, at non-atomic resolution (< 2.2 Å) the negative scattering peaks are shifted along the bond vectors from their correct positions and individual atom distinction is not always possible. As such, ^{1}H atoms (hereby designated H) are routinely replaced with ^{2}H (hereby designated D) which introduce a strong and positive scattering factor. Signal in a dataset where analyte atoms scatter in phase is much easier to interpret, as H-atom contribution to negative Fo-Fc maps can be, by and large, ignored until a good model is achieved.

Perdeuteration with deuterated protein expression media has been successfully employed to obtain strong signal from all protein hydrogen atoms (Shu *et al.*, 2000); however, since the labile solvent, hydroxyl, amine and sulfhydryl protons that readily exchange in solution are most likely to be of interest, detection of their exchange is usually sufficient for enzymological inquiry.

Proteins can be crystallized in D_2O liquor, but more often D_2O is introduced to the crystal by vapor diffusion: placing deuterated liquor drops adjacent to the crystal drop for an extended period of time, on average for nearly a year (Bennett *et al.*, 2008), though successful exchange has been reported after as little as a week (Blum *et al.*, 2009). The vapor diffusion method allows the slow introduction of D atoms without adversely affecting crystal stability.

Neutron data can be refined in tandem with X-ray data from the same or an isomorphous crystal. This has two distinct advantages. The first is to facilitate modeling. This is especially beneficial in non-perdeuterated samples of moderate resolution as the positive scattering signal can be quenched by riding hydrogens, particularly of methyl groups. Joint refinement allows one to build a very good starting model of the CNOS atoms with X-ray data, amiable to all modern crystallographic fitting software. Secondly, combining X-ray and neutron datasets of the same structure improves the ratio of observations to parameters, reducing systematic errors as they are unlikely to exist in both datasets (Wlodawer & Hendrickson, 1992).



Spallation sources	Centre	Location	Power (MW)	Established
	ISIS	Oxford, UK	0.16	1985
	SINQ	Nr Willigen, Switzerland	1	1996
	LANSCE	Los Alamos, USA	0.16	1972
	SNS	Oak Ridge, USA	5	2006
	JSNS/MLF	Tokaimura, Japan	5	2008
High-flux reactors	ILL	Grenoble, France	58	1971
	HIFR	Oakridge, USA	85	1965

Fig. 1.7. Linacs and nuclear reactors with stations dedicated to protein crystallography. Arial photographs are adapted from the organisation's websites, respectively: isis.rl.ac.uk/, sinq.web.psi.ch/, ess-scandinavia.eu/, lansce.lanl.gov/, sns.gov/, j-parc.jp/ and ill.eu/.

1.4.3 Neutron sources — The nuclear reactor at Chalk River Laboratories (CRL) in northern Ontario was the first neutron source to be adapted for neutron diffraction experiments (Brockhouse, 1953), and at 125 MW the CRL reactor built in 1962 remains the most powerful on the planet, seconded by the Oak Ridge National Laboratory's 85 MW reactor adapted for this purpose in 1990. High energy radiation can damage biological crystals bathed in X-rays; however, this is not a concern for neutrons even at prolonged exposures: inelastic scattering of neutrons, where kinetic energy of an incident particle is not conserved, is rare and does not excite electron energy levels as X-rays do.

Unfortunately the CRL reactor has never been equipped for protein crystal diffraction analysis. Furthermore, although power is important, it is not the rate limiting property to a neutron beam's diffracting capabilities. Despite the impressive power of the CRL reactor, it still has only moderate flux (particle density of ~ 10^{15} n/cm^2/s). High flux is important for crystal diffraction experiments to increase the total number of collision events.

Stronger neutron flux (> 10^{17} n/cm^2/s) of lower energy neutrons is obtainable at dedicated linear accelerator stations such as Los Alamos Nation Laboratory's Lujan Neutron Scattering Center (LANSCE). Facilities such as LANSCE produce higher flux through spallation: the collision of a high-velocity proton produced by a linear accelerator with a tungsten target shatters the tungsten nuclei to produce a spall of subatomic particles. A comprehensive list of functional linacs and nuclear reactors with stations dedicated to protein crystallography is provided as **Fig. 1.6.**

The precise timing of collision events allows calculation of the range of neutron velocities, and hence distinct neutron wavelengths through the de Broglie relationship.

The polychromatic neutron spall naturally separates along the path of the beam such that the particles with the shortest wavelengths arrive first.

This real-time wavelength resolution of distinct Bragg reflections circumvents the requirement for monochromation and allows for better utilization of the neutron source as multiwavelength time-of-flight Laue diffraction patterns, thus reducing the total data collection time in an experiment where the incoming flux is weaker, diffraction events are less often, and detection is more difficult than in equivalent X-ray experiments. Nonetheless, these limitations generally necessitate crystals several orders of magnitude larger than required for equivalent X-ray experiments.

1.4.4 Neutron detection — Gadolinium based neutron-sensitive image plates to detect neutrons have been developed (Cipriani *et al.*, 1994) and are implemented cylindrically (**Fig 1.8a**) or spherically (**Fig 1.8b**) to increase the effective detector surface area to volume ratio. In contrast, LANSCE's protein crystallography station utilizes a unique goniostat-mounted cylindrical 8-segment 2D multiwire area detector (**Fig. 1.8cd**). The multiwire enclosure is filled with the rare non-radioactive isotope ^{3}He, which evolves tritium (^{3}H) ions via the nuclear reaction:

$$\mathbf{n + {}^{3}He \rightarrow {}^{3}H + {}^{1}H + 0.764\ MeV}$$

These charged tritium ions are detected by the LANSCE PCS multiwire counter.

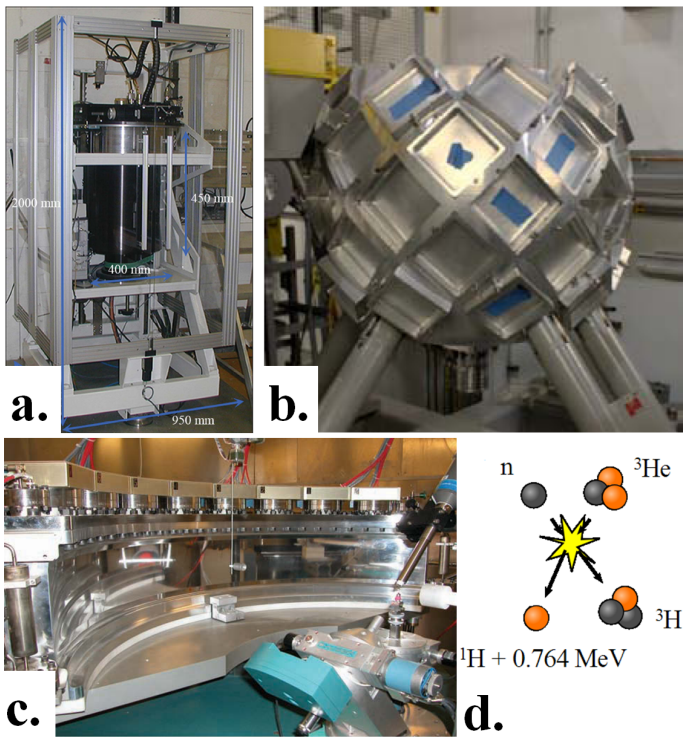


Fig. 1.8 Neutron detection. a. Cylindrical area detectors such as the quasi-Laue diffractometer LADI III in Grenoble (photo adapted from ill.eu/) are used in both Europe and Japan. **b.** The Oakridge Macromolecular Neutron Diffractometer MaNDi (photo adapted from ornl.gov/) and **(c)** LANSCE PCS multiwire area detector are both unique designs. **d.** The Nuclear reaction exploited at LANSCE PCS.

Chapter 2. Objectives

Understanding the fundamental structure-function relationships of glycosyltransferases is essential for exploiting their mechanisms of transfer and recognition, with the long-term goal of developing glycoconjugates and inhibitors of therapeutic significance. The absence of a clearly understood transfer mechanism for this important group of enzymes has been a major impediment to the design of such inhibitors.

The objectives of this work were: to continue in-depth structural (including neutron diffraction) and kinetic examination of glycosyltransferases, using human ABO(H) blood group glycosyltransferases and mutants as the primary model of retaining glycosyltransfer.

Hypothesis: Determination of GTA active site proton positions by neutronographic methods would give crucial insight into substrate binding and resolution of the enzyme reaction mechanism for proton transfer, with additional mechanistic insight to be gained by in depth kinetic and structural analysis of retaining and inverting glycosyltransferases in complex with donor and acceptor substrates (or analogs) produced by ourselves and other groups.

Chapter 3. Experimental approach

3.1 Human ABO(H) Glycosyltransferases

3.1.1 Mutagenesis — The original codon optimized genes for GTA (**Fig. 3.1**), GTB and chimera were produced by Dr. N.O.L. Seto at the National Research Council of Canada (Seto *et al.*, 1999) and inserted in a customized plasmid (pCWΔlac) for expression in *E. coli* BL21 (DE3). Subsequent site directed mutageneses where performed at UVic by myself or Joshua Moreau, or in the lab of Dr. Monica Palcic in Copenhagen. Site directed mutageneses used template isolated from the original plasmids was performed for these studies using either the QuikChange™ site directed mutagenesis kit (Stratagene) which allows one step PCR amplification of the entire plasmid with the introduced mutation(s), or more traditional overlap extensions using other high fidelity polymerases such as Phusion™ (New England Biolabs). Though not all are addressed in the discussion of results, the amino acid identities of mutations analysed independently or in tandem for these enzymes which produced soluble protein are shown in **Fig. 3.1.**

Fig. 3.1. DNA and protein sequences encoding the *E. coli* codon optimized GTA construct. Protein residues mutated for analysis are highlighted.

```
  1 aattcatgaa aaaaaccgct atcgcgatcg cagttgcact ggctggtttc gctaccgttg
 61 cgcaggccgc tgttcgtgaa ccggaccatc tgcaacgcgt ttccctgccg cgtatggttt
121 acccgcagcc gaaagttctg accccatgcc gtaaagacgt tctggttgtt accccgtggc
181 tggctccgat cgtttgggaa ggcaccttca acatcgatat cctgaacgaa cagttccgtc
241 tgcaaaacac caccatcggt ctgaccgttt tcgctatcaa aaaatacgtt gctttcctga
301 aactgttcct ggaaactgct gaaaaacact tcatggttgg tcaccgtgtt cactactacg
361 ttttcaccga ccagccggcc gcggttccgc gtgttaccct gggcaccggt cgtcaactgt
421 ccgttctgga agtgcgcgcc tacaaacgtt ggcaggacgt ttccatgcgt cgtatggaaa
481 tgatcagcga cttctgcgaa cgtcgtttcc tgtccgaagt tgactacctg gtttgcgttg
541 acgttgacat ggagttccgt gaccacgttg gtgttgaaat cctgaccccg ctgttcggta
601 ccctgcaccc gggcttctac ggttcctccc gtgaagcatt cacctacgaa cgtcgtccgc
661 agtcccaggc ctacatcccg aaagacgaag gtgacttcta ctacctgggt ggtttcttcg
721 gtggttccgt tcaggaagtt cagcgtctga cccgtgcatg ccaccaggct atgatggttg
781 accaggcgaa cggtatcgaa gctgtttggc acgacgaatc ccacctgaac aaatacctgc
841 tgcgtcacaa accgaccaaa gttctgtccc cggaatacct gtgggaccag caactgctgg
901 gttggccggc tgttctgcgt aaactgcgtt tcactgcagt tccgaaaaac caccaggctg
961 ttcgtaaccc gg

 62  fmvslprmvy pqpkvltpcr kdvlvvtpwl apivwegtfn idilneqfrl qnttigltvf
122  aikkyvaflk lfletaekhf mvghrvhyyv ftdqpaavpr vtlgtgrqls vlevraykrw
182  qdvsmrrmem isdfcerrfl sevdylvcvd vdmefrdhvg veiltplfgt lhpgfygssr
242  eaftyerrpq sqayipkdeg dfyylggffg gsvqevqrlt rachqammvd qangieavwh
302  deshlnkyll rhkptkvlsp eylwdqqllg wpavlrklrf tavpknhqav rnp
```

3.1.2 Protein purification — The initial purification protocol that produced crystallizable protein (Marcus *et al.*, 2003) was cation exchange (Sulfopropyl Sepharose Fast Flow™, GE Healthcare) followed by addition of Mn^{+2} and affinity chromatography using immobilized UDP-Hexanolamine-agarose as the stationary phase. The protein was eluted with UDP/Mn^{+2} which can yield up to ~50 mg protein per litre of culture. All GTA/GTB and mutant proteins reported in this work were produced by myself.

Optimizations to the original protocol have led to larger, better diffracting crystals without Hg containing compounds such as 3-chloro-Hg-2-methoxy-propylurea in the crystallization conditions. The Hg atoms destabilized large regions in the polypeptide chain (Letts *et al.*, 2007). These optimizations include the following: replacement of EDTA with PMSF in the cell lysis buffer and rapid resuspension at 4°C; addition of reducing agent BME throughout the purification including cell lysis buffer; a second round of cation exchange following dialysis, both performed on a column with a vast excess of resin at controlled speeds of 1 ml/min; cation exchange with a 0-1 M NaCl gradient at room temperature; discarding the delayed affinity shoulder peak; centrifuging samples before and after the affinity purification step; washing or dialyzing eluted UDP/Mn^{+2} and concentration to ~40 mg/ml (higher concentration causes precipitation); and centrifuging the concentrated protein sample with supernatant transfer to a fresh vessel and addition of BME ~daily until no new precipitate was formed, which often led to improved crystallizability compared to freshly produced protein.

Protein concentrations of semi-purified samples were estimated by UV absorption, and by the methods of Bradford following the addition of UDP/Mn^{+2}. A typical elution profile for such a purification scheme is given as **Fig. 3.2**.

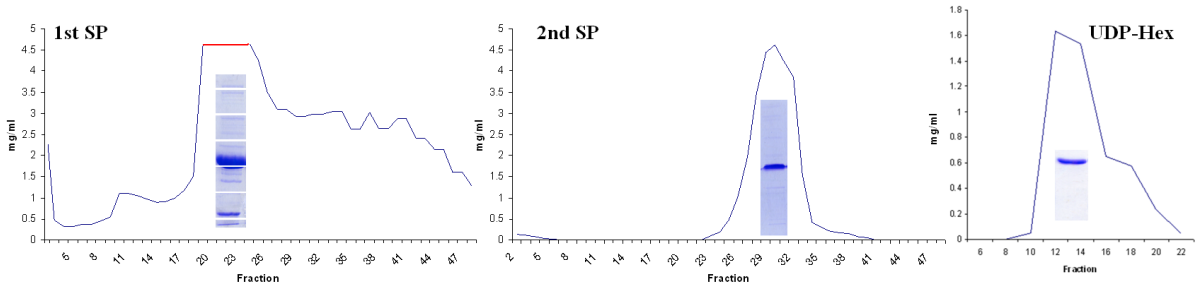


Fig 3.2. Typical GTB purification elution profiles and associated SDS-PAGE gels.

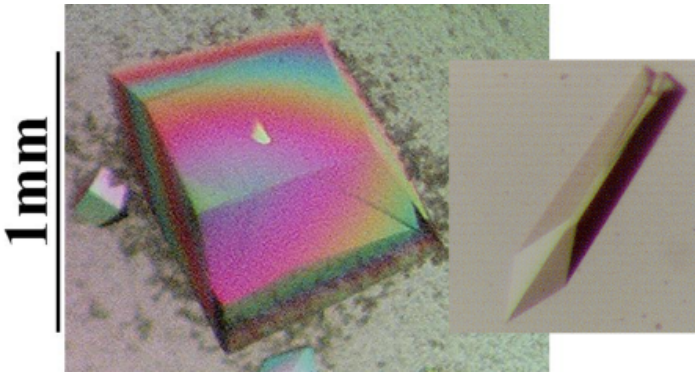
3.1.3 Kinetics — Bisubstrate kinetic assays were not performed on site, but by our collaborators in the lab of Dr. Monica Palcic at Carlsberg Laboratory. The method is described elsewhere (Seto *et al.*, 1997). Briefly, the radiochemical assay uses an immobilized acceptor analog (α-Fuc-(1→2)-β-D-Gal-O$(CH_2)_7CH_3$) and radiolabeled UDP-Gal[6^{3}H] or UDP-GalNAc[6^{3}H] as donor (Palcic *et al.*, 1988) with several concentrations of acceptor or donor used at saturating concentrations of the alternate substrate.

At the end of assays, the mixture was applied to a Sep-Pak reverse phase C18 cartridge and washed with water to remove unreacted radiolabeled donor. Unreacted acceptor and radiolabeled product were eluted from the cartridge with methanol and counted in a liquid scintillation counter after the addition of ECOLITE(+)™ liquid scintillation cocktail (MP Biomedicals LLC). V_{max} and K_A (acceptor K_M) and K_B (donor K_M) are derived using nonlinear regression analysis with GraphPad Prism 3.0 (GraphPad Software, San Diego, CA).

3.1.4 Crystallization — The initial hanging drop vapor diffusion crystallization conditions (native and SeMet), as well as all crystal structures reported in this work were carried out by Dr. Svetlana Borisova in Parafilm (Pechiney Plastic Packaging Company) wrapped

35x10 mm tissue culture dishes with 5-30 µl drops reported as containing 6−8 mg/ml enzyme, 70 mM ADA pH 7.5, 50 mM sodium acetate pH 4.6, 40 mM NaCl, 8 mM $MnCl_2$, 2.5% MPD, 5% glycerol, 2% PEG 4000 and 0.5 mM 3-chloro-Hg-2-methoxy-propylurea incubated at 16°C with 30% glycerol added to the mother liquor as cryoprotectant (Patenaude *et al.*, 2002). This tissue culture dish method (affectionately known as a "Svetlana-drop" in homage to Dr. S.N. Borisova) was employed to obtain all crystals 1 mm in length or larger using large sitting drops (*eg.* **Fig. 3.3**).

Fig. 3.3 Examples of large GTA crystals. Left crystal produced by Dr. S.N. Borisova, right crystal by myself.



3-chloro-Hg-2-methoxy-propylurea was omitted from all crystals that produced the closed conformation. A typical GTA crystallization screen for closed conformation crystals consists of reservoirs containing 25mM NaCl, 7% MPD, 3% PEG 3400, 150mM ADA pH 7.5 and 300mM ammonium sulfate with a range of 0-100mM sodium acetate pH 4.6, 5mM $CoCl_2$ or 10mM $MnCl_2$, and the 6µl drops containing ratios of reservoir-solution: enzyme from 1:2 – 1:5 in Hampton 24 well plates incubated for 2 days at 4°C then transferred to 16°C. Closed conformation crystals of the ABBB chimera could be grown overnight without substrate in just PEG 3400 and tris pH 8.5 at least in part due to

the rigidity imparted by Arg176 and Ser235 which bookend the internal mobile loop. GTB P234S spontaneously crystallized after concentration to ~100 mg/ml.

Co-crystals with a number of substrate analogs were also investigated, the two reported here (**Fig. 3.4**) are α-L-Fuc-(1→2)-β-D-(3-deoxy)-Gal-O$(CH_2)_5CH_3$, a 3-deoxy-acceptor (DA) provided by Dr. O. Hindsgaul (Kamath *et al*., 1999) and UMP-PO_2-CH_2-Gal (UC-Gal), a C-glycosidic phosphonate donor analog provided by Dr. T. Lowary (Laferte *et al*., 2000).

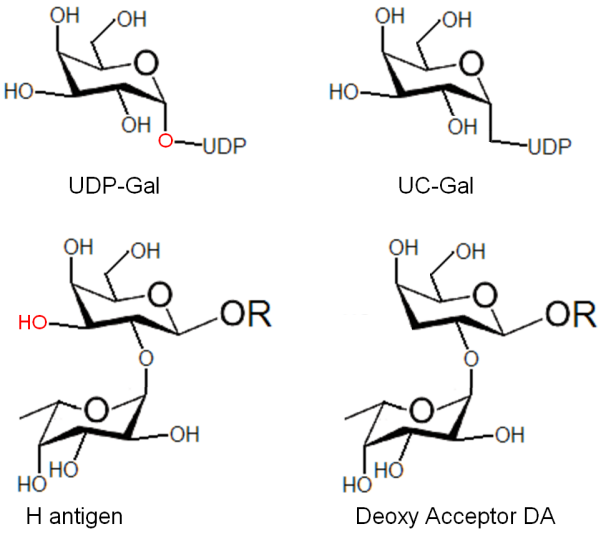


Fig. 3.4 Donor and acceptor analogs Natural donor and acceptor molecules are compared to catalytically inactive analogs. Substituted groups are highlighted in red.

3.1.5 X-ray data collection and refinement — X-ray collection and refinement statistics can be found as **Appendix II**, where the bulk of refinement was carried out by the lead author cited (namely myself or Asha Johal), and the unpublished structures GTA/C209A/E303C, GTB/M214G, GTA/P234S and GTA/L266V were solved myself.

MicroMax-002 (Rigaku/MSC) generated X-rays coupled to Osmic 'Blue' X-ray mirrors produced diffraction data collected on a Rigaku RAXIS IV++ area detector by 0.5° oscillations on a single φ vertical goniometer with 2-8 minute exposure to single crystals at a distance of 72mm. When cryogenic conditions were employed crystals were

frozen and maintained at a temperature of −160°C using a CryoStream700 crystal cooler (OxfordCryosystems).

These data were collected, integrated, scaled and merged using CrystalClear™ (Rigaku/MSC) and either d*TREK (Pflugrath, 1999) or HKL2000 (Otwinowski & Minor, 1997). Refinement strategies varied from dataset to dataset, but in general phases were solved by molecular replacement using Phaser (McCoy *et al.*, 2007) with previously determined structures as starting models. Solved structures were subsequently refined using the CCP4 module REFMAC5 (Murshudov *et al.*, 1997; The CCP4 suite: programs for protein crystallography, 1994) and/or Phenix.refine (Adams *et al.*, 2002) with intermittent model fitting using SetoRibbon (Evans, unpublished) and/or the Crystallographic Object-Oriented Toolkit (Emsley *et al.*, 2010). Ligand libraries were generated with either ReadySet (Adams *et al.*, 2002) or the PRODRG server (Schuttelkopf & van Aalten, 2004). All GTA/GTB crystals described are of orthorhombic space group $C222_1$ with unit cells dimensions of 52.4-53.5 Å, 149.0-151.9 Å, 78.2-80.2 Å.

3.1.6 Deuteration, neutron data collection and refinement — Neutron diffraction data were collected, integrated, scaled and averaged by Dr. S.Z. Fisher at Los Alamos National Laboratories, New Mexico, using d*TREK (Pflugrath, 1999) with custom modifications for use with the wavelength-resolved Laue neutron data (Langan & Greene, 2004). Refinement was carried out by myself.

Undeuterated crystals were flanked in a quartz capillary by drops of deuterated mother liquor for vapor diffusion over the span of ~1 year. 18 Laue images of LANSCE spallation-generated TOF neutron diffraction data were collected on the PCS multiwire

detector by 30˚ oscillations about φ under 5 different κ and ω orientations on a Huber κ-circle goniometer (**Fig. 1.8c**) with 18 h exposure. Examples of neutron and X-ray diffraction patterns utilized for joint refinement are displayed in **Fig. 3.5**. The integrated neutronographic Laue dataset was wavelength normalized using LAUENORM and merged using SCALA of the CCP4 suite of crystallographic programs (Diederichs & Karplus, 1997; Helliwell *et al.*, 1989; Evans, 2006), and solved assuming isomorphism with a molecular replacement solved x-ray data collected at room temperature with a compatible unit cell for joint x-ray/neutron refinement.

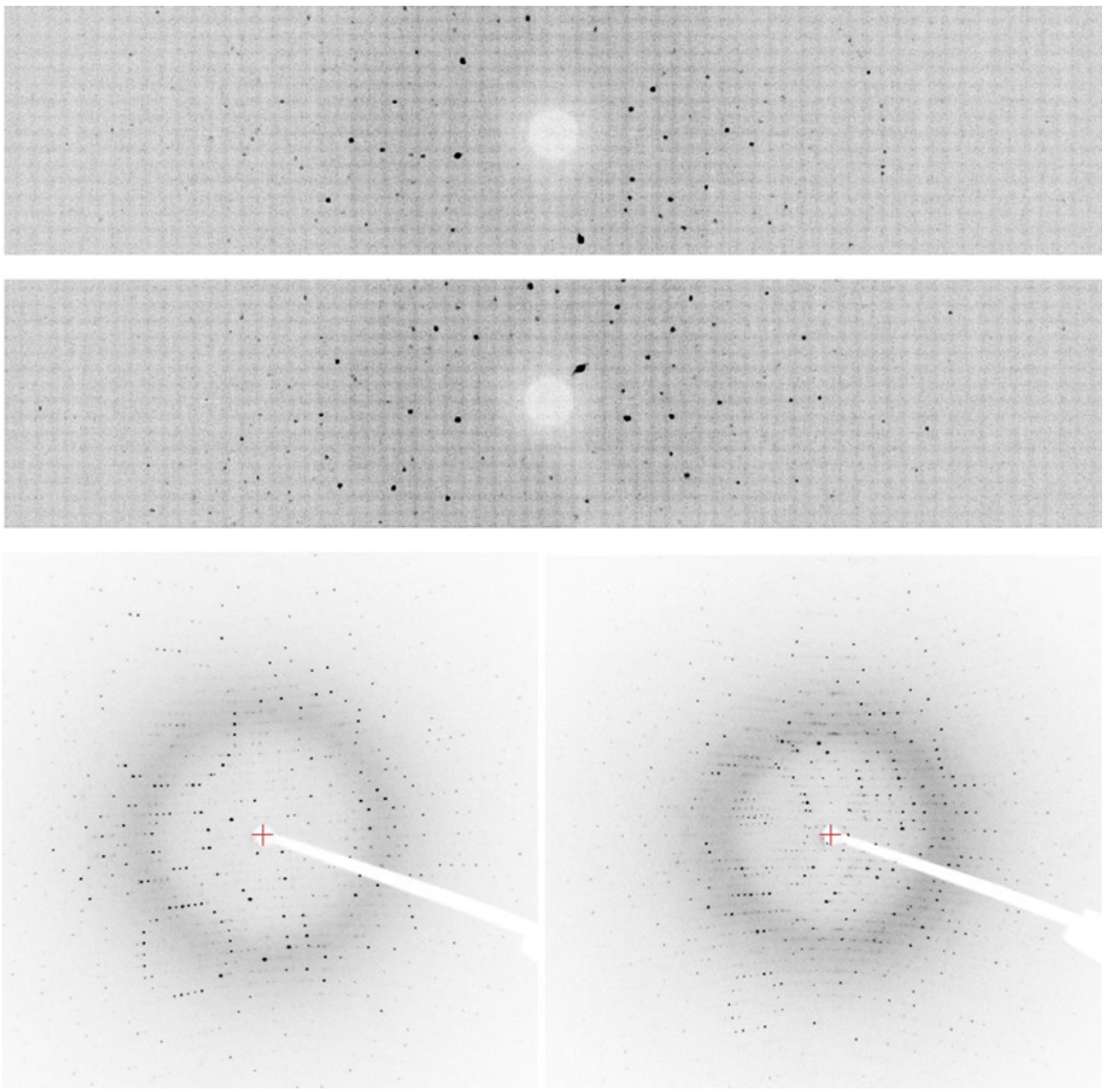


Fig. 3.5 Diffraction images. (top) Neutron Laue time-of-flight diffraction images of GTA collected at the PCS at two different ϕ settings. For this representation, the time-of-flight data were overlaid to produce a conventional multiwavelength Laue pattern. (bottom) X-ray diffraction images of GTA collected at room temperature with relative angles of 0 and 90° with 0.5° oscillations and 2 minute exposure. Adapted from (Schuman *et al.*, 2011) with permission (**Appendix I**).

Once this compatible X-ray dataset was obtained, it was used for a standard CNOS atom refinement to act as the starting model for joint refinement with the neutron dataset. Readily exchangeable hydrogen atoms (backbone amides as well as the functional groups of CHKNQRSTW and Y residues) were modeled as occupational conformers of

deuterium and hydrogen using ReadySet (Adams *et al.*, 2002) with riding hydrogen atoms added after multiple rounds of real-space, atomic displacement and occupancy refinement against joint X-ray and neutron data using nCNS (Adams, 2009) and phenix.refine (Adams *et al.*, 2002). The models were fitted between refinement rounds with SetoRibbon (Evans, unpublished) as the advanced rigid body and auto fitting tools of COOT (Emsley *et al.*, 2010) are not compatible with models which include hydrogen and/or deuterium atoms.

The model (PDB ID **4DHH**) presented in this work represents a final phenix.refine round with joint x-ray/neutron reflections using n Gaussian neutron/X-ray scattering distribution with a minimal refinement strategy lacking individual xyz parameters. The D atom omit map used to display H atom negative scattering without D atom model bias, and was generated using phenix.autobuild (Adams *et al.*, 2002).

Molecular depth indices from the solvent accessible surface were calculated with DPX (Pintar *et al.*, 2003) using 1.4 Å water probe spheres, then binned in ranges of 0-1, >1-2, >2-3, >3-4, and >4. Deuterium atoms are considered exchanged when their occupancies are refined to ≥ 20%.

3.2 Attempts to capture the GTB E303C covalent intermediate

Personal attempts to replicate the covalent intermediate identified in (Soya, 2010) crystallographically were made using both wild type GTA and postulate nucleophile mutant enzyme GTA-E303C. These attempts were made with both formed crystals and with protein/substrate mixtures prior to crystallization.

Briefly, samples were incubated with 0-20 mM substrates for 2-24 H at 37°C followed by exchange with 50 mM CH_3COONH_4 5 mM EDTA then addition of 10% CH_3COOH to pH ~3.3 (estimated with the Henderson–Hasselbalch equation).

Attempts were also made to observe the intermediate by ^{1}H and ^{31}P NMR analysis. Experiments were conducted with 50-500 mM substrate analytes in D_2O solvent using a BRUKER 500 magnet with 10-15 min acquisition time under the supervision of Christine Greenwood at UVic's NMR facilities. Mg^{+2} was used in lieu of natural metal cofactor Mn^{+2} to avoid paramagnetic decrease of relaxation time and consequent peak broadening, and to shift equilibrium in favour of the steady state intermediates of interest. In retrospect, cation choice may not appreciably lengthen the lifetime of the intermediate state, but only prolong the establishment of the catalytically competent conformer. An example of an ^{1}H NMR spectrum of the B-antigen product with the expected chemical shifts for α and β anomers is given in **Fig. 3.5**, with the chemical shifts reported in parts per million. Tetramethylsilane was used as an internal standard.

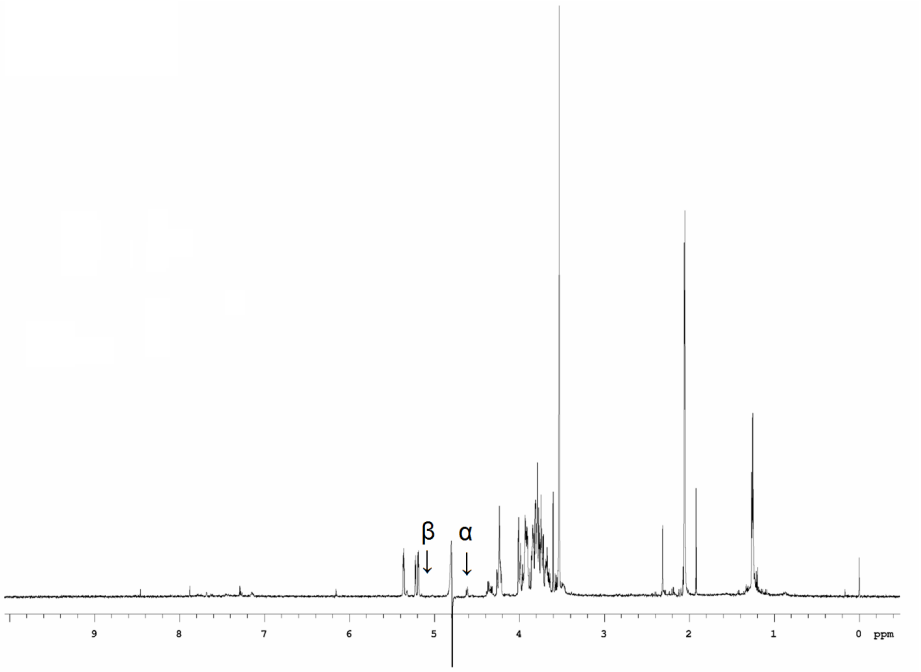


Fig. 3.6. An example ^{1}H NMR spectra of the B-antigen product including the expected chemical shifts for the α and β Gal C1 anomers.

3.3 Reanalysis of glycosyltransferase structural data

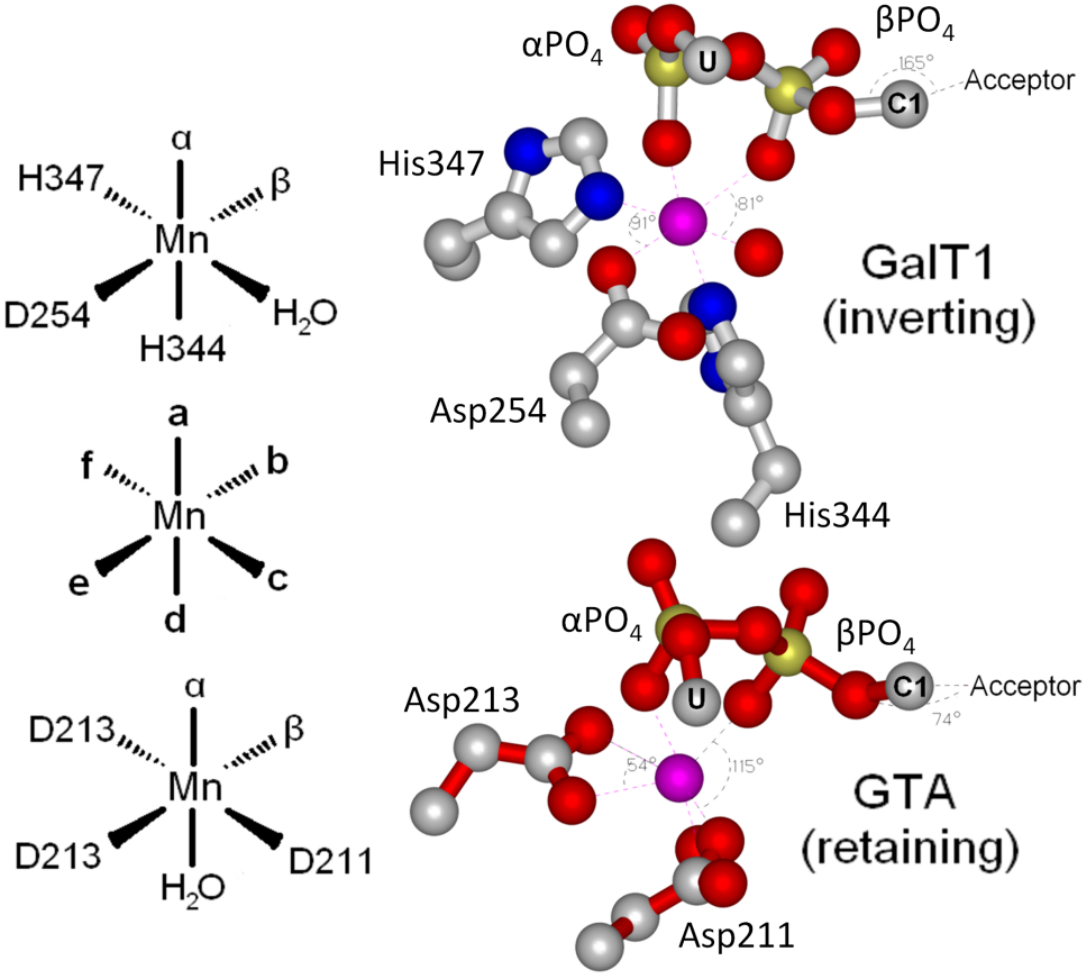
3.3.1 Geometric analysis — Deposited GT-A fold PDBs identified by CAZY were personally analyzed for geometric parameters using SetoRibbon (Evans, unpublished). Of the 11 GT-A fold families with deposited structures (listed in **Table 3.2**), 2 were found which had unambiguous complete (nucleotide and monosaccharide) donor density as well as acceptor molecule analogs in a single structure: GT-6 retaining enzyme GTA (Alfaro *et al.*, 2008) and GT-8 retaining enzyme LgtC (Persson *et al.*, 2001). Donor and acceptor molecules were unambiguously observed in separate structures for GT-7 inverting enzyme GalT1 (Ramakrishnan *et al.*, 2004); and GT-43 inverting enzyme GlcAT-1 (Kakuda *et al.*, 2004; Pedersen *et al.*, 2002). Inverting enzyme GnT1 (Gordon *et al.*, 2006) as well as retaining enzymes Extl2 (Pedersen *et al.*, 2003) and ManT (Goncalves *et al.*, 2010) have deposited structures with entire donor density but without available acceptor bound structures.

Enzymes other than GTA and LgtC have the acceptor absent, modeled as glycerol or required combining separate donor-bound and acceptor-bound PDBs for analysis. It should be noted that the accuracy of their analyzed geometry is somewhat mitigated as the same steric constrains as the bisubstrate liganded active site are not necessarily applicable. GTA and LgtC both have deposited structures with both donor and deoxy acceptor analogs (as seen in Fig. 3.4) bound simultaneously, and the deviation of these analogs from observed natural acceptors is only ~0.1 Å. Some of the parameters of interest are absent entirely for enzymes GnT1, Extl2 and MpgS.

Table 3.1 GT-A fold glycosyltransferases with deposited structures. Blue families were assessed to have structures with unambiguous whole acceptor and donor molecule electron density for analysis (underlined in the same structure); those in green have only structures with donor density.

Family	Example enzyme	Stereo-specificity	Example Complex(es)
GT-2	SpsA	Inverting	UDP
GT-6	**GTA**	**Retaining**	**UDP-Gal + Gal-Fuc**
GT-7	**GalT1**	**Inverting**	**UDP-Gal, GlcNAc-GlcNAc**
GT-8	**LgtC**	**Retaining**	**UDP-Gal + Gal-Glc**
GT-13	**GnT1**	**Inverting**	**UDP-GalNAc, UDP-Glc**
GT-15	Kre2	Retaining	GDP + Man + GlcNAc
GT-27	GNAc:Pep	Retaining	UDP + GlcNAc
GT-43	**GlcAT-I**	**Inverting**	**UDP-GlcUA, UDP + Gal-Gal-Xyl**
GT-55	**MpgS**	**Retaining**	**GDP-Man**
GT-64	**Extl2**	**Retaining**	**UDP-GalNAc**
GT-78	MgS	Retaining	GDP
GT-81	ManT	Retaining	GDP

Fig. 3.7. The **a-f** nomenclature used to describe metal binding partners with example inverting enzyme GalT1 (top) and retaining enzyme GTA (bottom). **U** is uridine, **C1** is donor galactose **C1**.



One set of geometric parameters of particular interest are the angles between the phosphates and acidic residues that coordinate the divalent metal cofactor (**M**). The molecular geometry surrounding the metal is roughly octahedral, though when acute bidentate Asp coordination is employed this can be skewed to nearly trigonal prismatic dimensions. For consistency the α-phosphate **O2** is labeled as **a** at the apex of the coordination octahedron, and place the β-phosphate **O1** as **b** and the remaining coordinating atoms **c-f** in a clockwise layout as illustrated in **Fig. 3.7**.

The metal cation is the focal point of the alignments analyzed: instead of optimizing RMS for the entire polypeptide chain, alignments were made using metal cation **M**, α-phosphate **O1** and β-phosphate **O2** as fixed positions from which to compute relative distances and angles.

Other geometric parameters of interest include: the nucleophilic distance from the acceptor nucleophilic oxygen atom **Nu** to the donor monosaccharide electrophilic center **C1;** the angle between the incoming nucleophile and leaving group β-phosphate oxygen **O3**; the distance between **Nu** and **O3**; the distance to the closest observed **O3** and **C1** enzyme dipoles; the angle between **O3**, **C1** and their nearest dipole vectors; the nearest dipole vectors to ring oxygen **O5**; and the angles between donor β-phosphate **O1** and the adjacent coordinating acidic ligands of the metal ion. Protein macrodipoles were estimated with the Protein Dipole Moments Server (Felder *et al.*, 2007).

3.4 Figure generation

All figures (with the exeption of photographs) were produced by myself. **Fig. 1.2** was generated with SETOR (Evans, 1993), **Figures 1.3-1.6** and **4.3a** with Pymol (DeLano, 2002). All other structures were generated with SetoRibbon (Evans, unpublished). Carbohydrate PDBs for **Figures 1.3** and **1.4** were generated with SWEET-DB (Loß *et al.*, 2002). The C-terminal position of the open conformation of ABBB (PDB 2R7J) in **Fig. 1.5** and protraction of vWF A2 tertiary structure in **Fig. 1.4** (PDB 3GXB) were modeled with Foldit Standalone (Cooper *et al.*, 2010). Neutron and X-ray diffraction in **Fig. 3.5** were displayed with PCS (Langan & Greene, 2004) and CrystalClear (Rigaku), respectively.

Chapter 4. Results and discussion

4.1 ***Human ABO glycosyltransferase X-ray structure-function studies***

4.1.1 Ligand conformations — Recently (Alfaro *et al.*, 2008) crystallization conditions have been optimized to reveal unambiguous density for the donor sugar and two previously disordered sections of polypeptide which both directly interact with the substrates (**Fig. 1.5**): an internal loop spanning residues 176-196 and a C-terminal residues 346-354. With the exceptions of chimeric mutant ABAA and triple mutant GTB/C80S/C196S/C209S this closed conformation is stable (at neutral pH) only in the presence of both donor and acceptor substrates or substrate analogs.

Four distinct conformations of the UDP-Gal galactose moiety have been observed with unambiguous electron density (**Fig. 4.1**, **Table 4.1**). Observing the entire donor molecule at all, let alone in concert with an acceptor, is a rare occurrence with only 4 GT-A fold enzyme examples in the literature with unambiguous density for both donor and acceptor molecules. Of these, only GTA and GalT1 have densities for both donor and acceptor molecules observed in the same structure (**Table 4.2**).

Conformer 1 (**Fig. 4.2a**), first identified in (Alfaro *et al.*, 2008), is compatible with formation of the closed conformation of the enzyme and thus is the only catalytically competent conformation.

Conformer 2 (**Fig. 4b**) was first identified in mutant enzymes GTB/C80S/C196S and GTB/C80S/C196S/C209S (Schuman *et al.*, 2010), though the same conformation has since been identified in wild type enzymes and other mutants. The change in conformation 2 begins at the UDP β-phosphate, which rotates about the Mn^{2+}

coordination site such that it forms a direct hydrogen bond to Asp211 and a water molecule.

This invokes different hydrogen-bonding partners for each and every sugar oxygen atom (**Table 1.1**) and is a conformation more similar to the dominant solution rotamer identified by both transferred nuclear Overhauser enhancement (Angulo *et al.*, 2006) and molecular dynamic analysis with MarvinView (ChemAxon, data not shown). This suggests that the enzyme is able to bind the more abundant solution rotamers and facilitate gyration to active donor **conformer 1**.

Residues in the internal loop and the C-terminus of GTB are essential for catalysis (Yazer & Palcic, 2005; Persson *et al.*, 2001; Boix *et al.*, 2001), and **Conformer 2** is incompatible with closure of the enzyme loops for three reasons. Firstly, the new position of Gal O6 (**Fig. 4.2c**) sterically hinders approach of internal loop Trp181. Secondly, new position of UDP-Gal β-phosphate is inaccessible to Arg352 of the C-terminus. Thirdly, a water molecule participates in stabilisation of this rotomer only 3.2 Å from O3 (**Fig. 4b**), which could allow destructive donor hydrolysis through nucleophilic attack by water.

UDP-Gal's C4 epimer UDP-Glc is a poor donor for GTA and GTB, with enzymatic activities 4 orders of magnitude lower than the native donor substrates (Seto *et al.*, 2000). ES-MS studies determined UDP-Glc to have binding interactions similar to UDP-Gal (Soya *et al.*, 2009). Like Gal, the donor glucosyl moiety directly interacts with Arg188, Asp211 and Asp302, however it does so with different hydrogen bonds at each carbohydrate hydroxyl (**Table 4.1**). This shift does not sterically hinder mobile loop closure, which accounts for the residual activity of UDP-Glc.

Fig. 4.1 The four observed donor rotamer conformations and matching 2Fo-Fc density maps contoured to 1 σ. **Conformer 1** is catalytically active, **conformer 2** represents a binding intermediate that prevents closure of the mobile loops, and in doing so may facilitate cooperative binding. The other 2 conformers have only been observed using non-native donors UDP-Glc and UC-Gal.

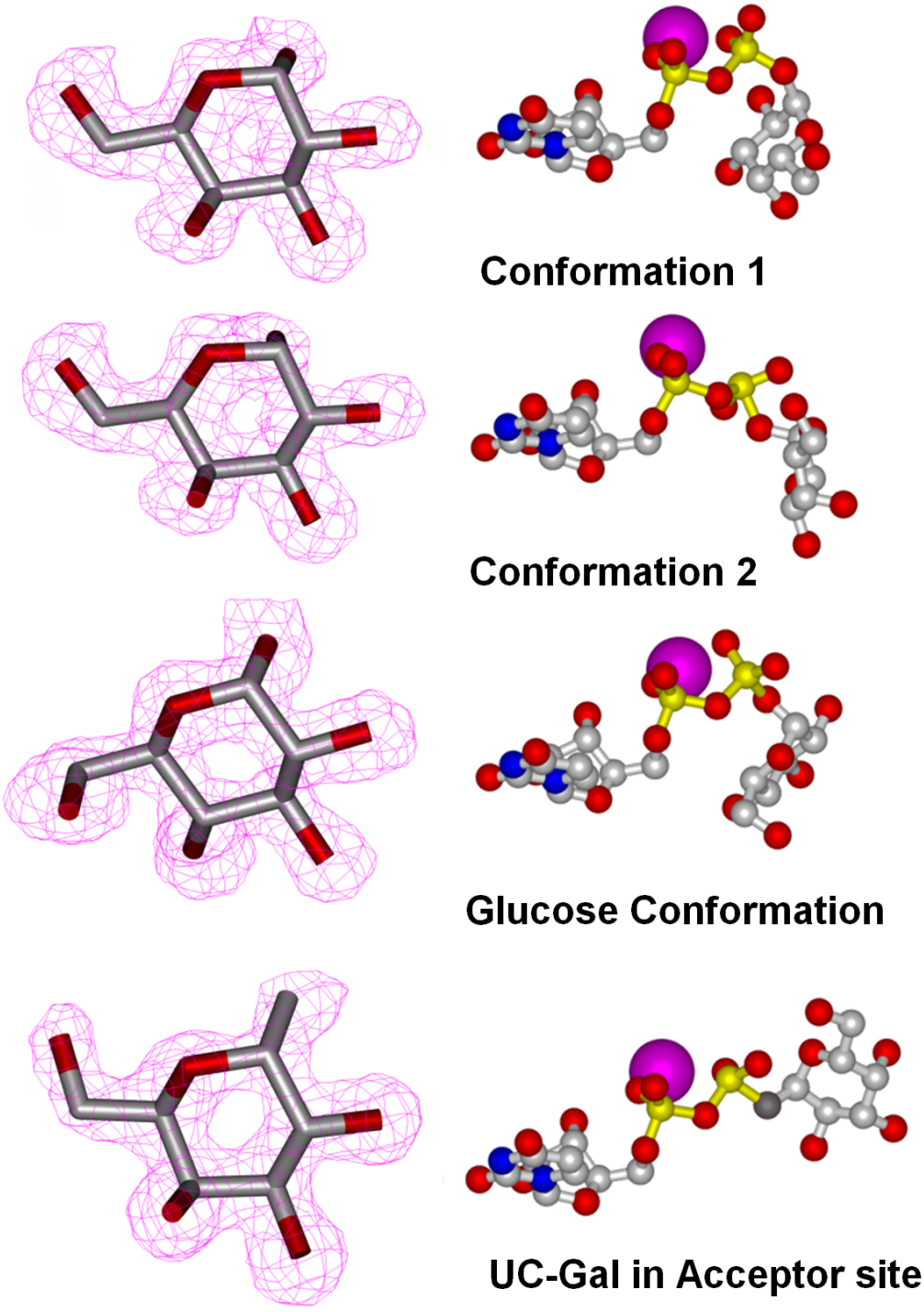


Table 4.1 Observed donor intramolecular and carbohydrate-enzyme hydrogen bonds

	Conformation 1	Conformation 2	Glucose Conformation	UC-Gal in Acceptor site
Gal O2	Glu303-Oε	Asp211-Oδ, βPO_4	—	Glu303-Oε1
Gal O3	Asp302-Oδ, His301-O	Arg188-NH_2, Asp211-Oδ2	Gly267-N, G/A268-N, Glu303-Oε	Glu303-Oε2
Gal O4	His301-O, His301Nδ	Asp302-Oδ	Arg188-NH_2, Asp211-Oδ	His233-Nδ
Gal O5	αPO_4, βPO_4	—	αPO_4	βPO_4
Gal O6	—	His301-Nδ	Asp302-Oδ	βPO_4

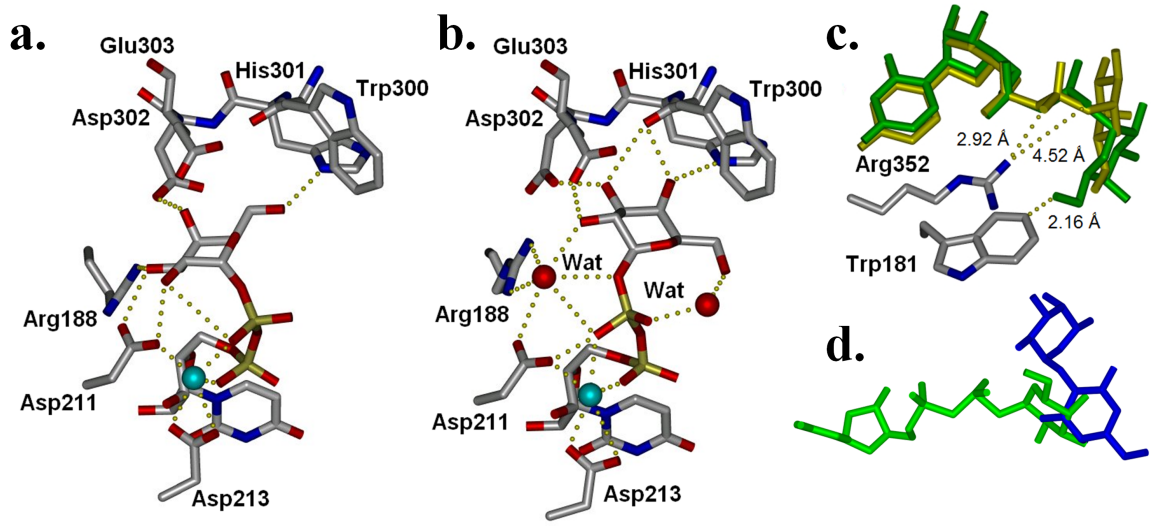


Fig. 4.2 Bioactive and inactive conformations. Observed binding partners for **(a) conformers 1** and **(b) 2** (Cyan spheres Mn^{+2}, red spheres H_2O). **(c)** Compared to **conformer 1** (yellow), **conformer 2** (green) inhibits enzyme closure by steric hindrance of Trp181 and rotation of β-phosphate out of Arg352 access. **(d)** A conformation observed for donor analog UC-gal positions the Gal moiety in the same position as acceptor Gal, and so is also incompatible with enzymatic turnover. **a-c** Adapted from (Schuman *et al.*, 2010) with permission (**Appendix I**).

UDP-Gal hydrolysis is thought to be one contributing factor to the paucity of glycosyltransferase structures with ordered donor carbohydrates. To circumvent hydrolysis, a C-glycosidic phosphonate donor analog UC-Gal (**Fig. 3.3b**) was produced. Most UC-Gal structures adopt **conformation 1**; however, complexes GTB, ABBA and ABBB were found to adopt a novel inactive donor conformation in which the galactosyl moiety is extended into the acceptor site (**Fig. 4.2d**).

While this conformation has not been observed for the native UDP-Gal donor, it is not dependent on the phosphonate tetrahedral geometry or the methylene hydrogens that are in no position to make any interactions, so it is a plausible intermediate state following initial UDP binding of a dominant solution UDP-Gal conformer while the donor gyrates into the bioactive conformation. Glycerol has been observed to bind in the donor binding site along with UDP, poorly mimicking UDP-Gal binding and distorting UDP > 0.5 Å away from the acceptor (Johal *et al.*, 2012). The 3-deoxy-H antigen (**Fig. 3.4**) is observed to bind within 0.1 Å of where the native H antigen binds (Alfaro *et al.*, 2008), and is an excellent model for the natural binding site

4.1.2 Mutations to increase crystallizability — Primary mutagenic analysis to improve structural determination of the mobile loops included serine and alanine mutagenesis of the enzyme's four cysteine residues, individually and in tandem (Letts *et al.*, 2007; Schuman *et al.*, 2010). These were constructed to avoid cysteine coordination of Hg atoms introduced with 3-chloro-Hg-2-methoxy-propylurea in the initial crystallization conditions. These Hg atoms caused significant disorder the mobile loops precluding enzyme closure (**Fig. 4.3a**). The mutations did not appreciably alter enzyme kinetics, and C209A markedly increased crystallizability. As such, C209A has been successfully used in conjunction with other mutations which were otherwise resistant to crystallization such as E303C.

These mutations were successful in showing the first glimpses of the previously disordered internal loop (residues 175-195), albeit not in the closed conformation as identified in (Alfaro *et al.*, 2008) (**Fig. 4.3b**). C80S also introduced a second UDP binding site that bridged the dimer interface (**Fig. 4.3cd**) with no perceived physiological role.

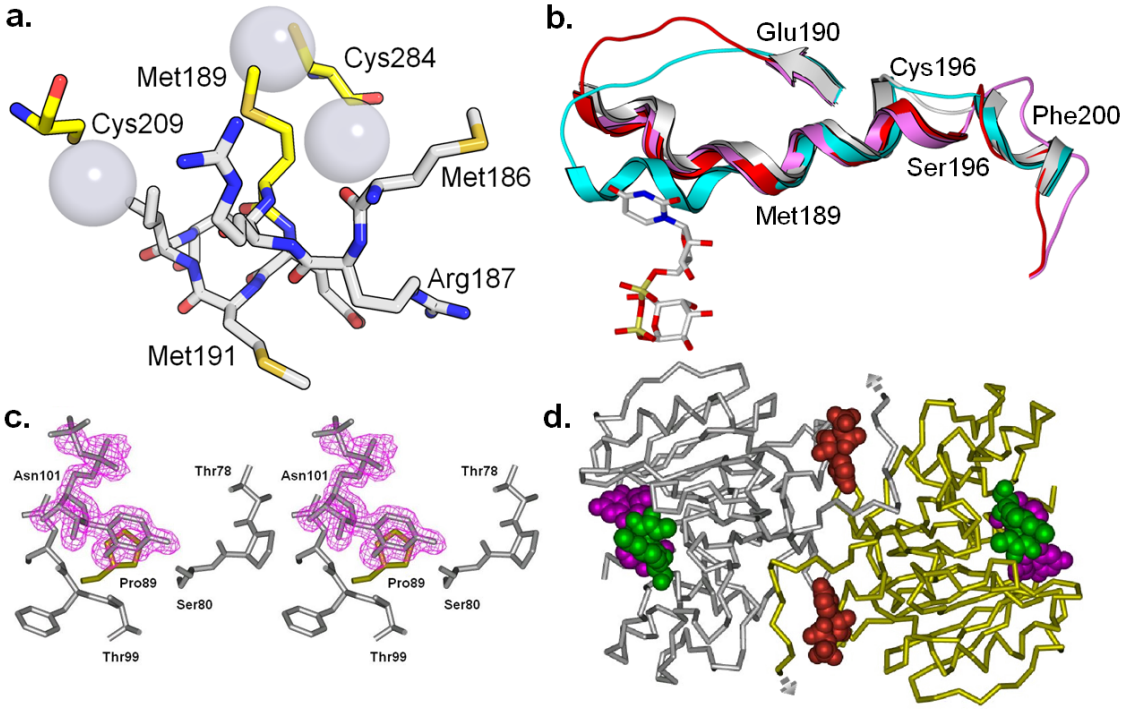


Fig. 4.3a. Three cysteine coordinated Hg atoms superimposed on the closed position of the internal loop they displace. **b.** Secondary structure of this loop displaying open (white) and closed (cyan) conformations overlapped with GTB/C80S/C196S/C209S (red) and GTB/C80S/C196S (magenta), which display an ordered open conformation **c.** Stereoview of 2Fo-Fc density surrounding the observed second UDP molecule, and **d.** its location on the biological dimer assembly (red spheres) compared to UDP-Gal (green) and the H antigen (magenta). **b&c** Adapted from (Schuman *et al.*, 2010) with permission (**Appendix I**).

4.1.3 Mutants identified from human blood banks — Particularly interesting mutations identified from blood bank data include those distal from the active site. Some rationale for the loss of activity observed for many of these mutations are given below.

GTB P156L, a mutation identified in human blood banks which coincides with the O^2 phenotype caused by mutation G268R (**Fig. 4.4a**) which sterically hinders donor

binding (Lee *et al.*, 2005). P156L by itself had the bizarre effect of spontaneously crystallizing into the $C222_1$ space group after concentration. These crystals (and crystals obtained from their melts) diffracted poorly, so a publishable dataset was never obtained. It can only be loosely inferred that this freakish and impractical crystallizability was the result of altered crystal contacts which served to extend the a axis by ~10 Å, thus altering all contacts on the long b axis (**Fig. 4.4b**).

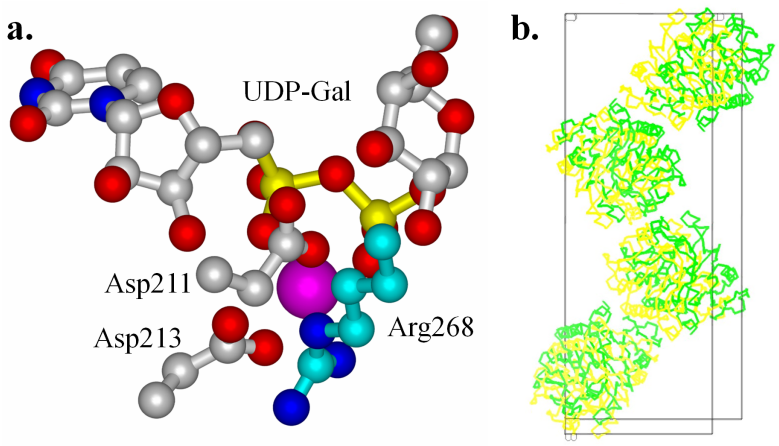


Fig. 4.4a. The G268R (cyan) point mutation causes the O^2 phenotype. R268 cannot be modeled without sterically clashing with donor Gal, Mn^{+2} and/or the DXD motif. **b.** P156L (green) unit cell molecules superimposed over those of wt GTB (yellow) demonstrating the changes between a and b axes.

GTB P234S transfers GalNAac at ~82% the rate of GTA, and Gal at < 4% the rate of GTB. The mechanism by which substrate specificity is changed remained enigmatic even in light of initial structure determination: the previous proposition that Ser234 repositions the GTB critical residue Met266 to allow GalNAc binding (Marcus *et al.*, 2003) was based on an incorrect substrate binding model.

Met266 adopts the aforementioned conformation upon enzyme closure, and the proposition also did not explain the reduction in UDP-Gal turnover. It is far more likely that P234S imparts substrate specificity in a similar fashion to its critical residue neighbour Gly/Ser235 (BBAB has 10x more GTA activity than GTB activity (Yamamoto & Hakomori, 1990)): by relaxing the acceptor binding pocket φ/ψ restraints to apportion room for the larger donor substrate. GTB natural mutations K346E (Olsson *et al.*, 2001),

K346M (Seltsam *et al.*, 2003), R352W (Yamamoto *et al.*, 1993) and R352Q (Olsson *et al.*, 2001) are located on the mobile C-terminal tail, and all have severely retarded activity. They abolish the positive dipoles necessary to stabilize the UDP nucleofuge, a concept descussed in more detail later.

*4.2 **GTA joint X-ray/neutron structural analysis***

4.2.1 Challenging conventions about size limitation — Neutron diffraction data were collected from a ~.03 mm^3 crystal grown in the absence of substrates and cofactors, which exhibited disorder in the mobile loops (residues 176-183 and 346-354). This is a sample size over an order of magnitude smaller than is usually required for neutron diffraction experiments, which have been reported in range from 0.4 to 50 mm^3 (Bennett *et al.*, 2008). Surprisingly, the largest GTA crystals tested (*eg*. **Fig. 3.3**) did not offer the best neutron diffraction during 24H screening.

Table 4.2 Data collection and joint refinement statistics for neutron and X-ray diffraction of $C222_1$ GTA crystals collected at room temperature.

Source **Detector**	Rigaku RAXIS IV++	LANSCE PCS ^{3}He-detector
Wavelength(s) (Å)	1.5418	0.7 - 6.4
Unit cell volume ($Å^3$)	653283	654195
Resolution (Å)	19.47 - 1.85 (1.92 - 1.85)	27.0 - 2.5 (2.64 - 2.50)
No. reflections Measured (unique)	120226 (26867)	27675 (2219) 9759 (1163)
Redundancy	4.47 (4.19)	2.8 (1.9)
Mean I/σ(I)	14.1 (4.1)	3.6 (1.6)
R_{merge} (%)	4.9 (29.5)	25.1 (32.4)
Completeness (%)	94.7 (98.2)	84.1 (70.3)
Refinement	X-ray	Neutron
R_{work} (%)	18.3	31.0
R_{free} (%)	20.7	34.8
RMS Bond length (Å)	0.013	
RMS Bond angles (°)	1.306	
PDB ID	4DHH	

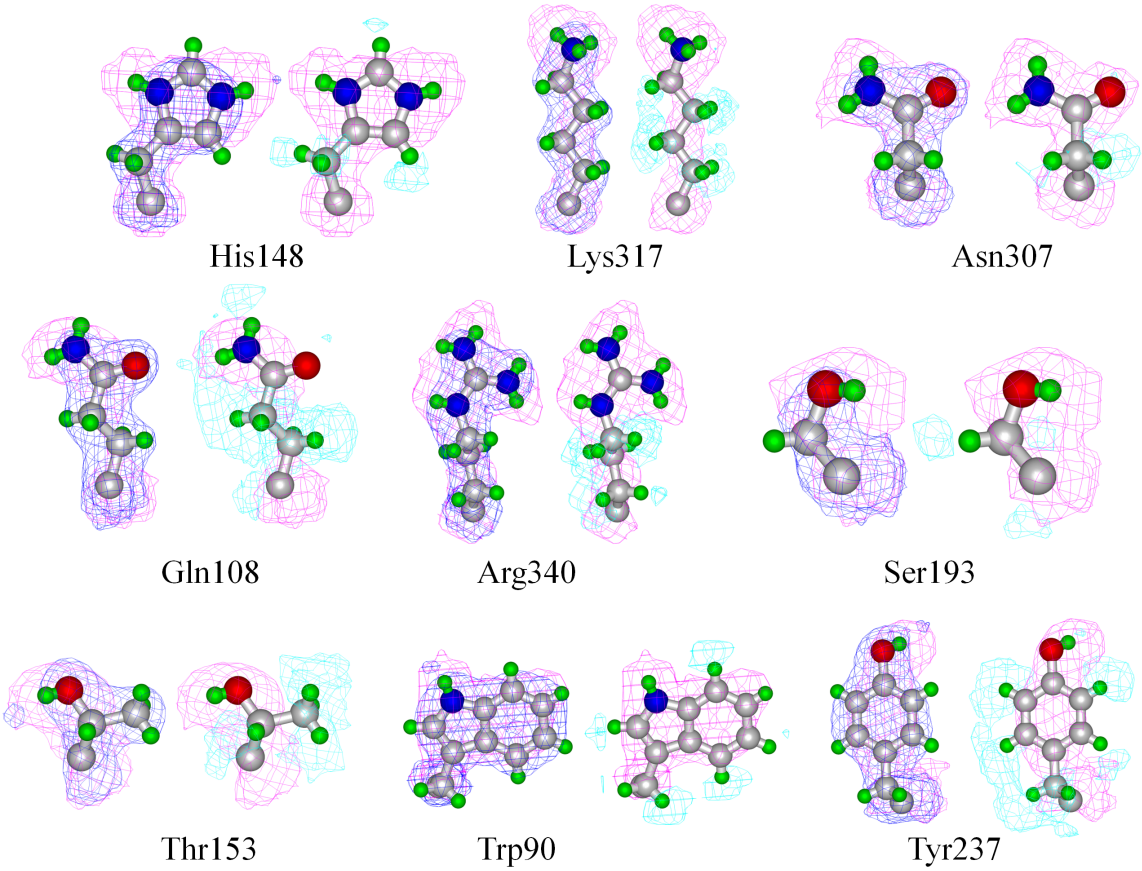


Fig. 4.5 Examples of amino acid species with exchangeable protons. Neutron 2Fo-Fc density maps (magenta) demonstrate exchanged deuterium atoms compared to 2Fo-Fc electron density maps (blue). Though the negative scattering peaks of the -1σ contoured 2Fo-Fc neutron omit maps (cyan) are shifted along the bond vectors and individual atom distinction is not always possible, they give good approximation of H locations, especially for methyl groups which can ablate positive signal (*eg.* Thr153).

The quality of these data is evident in the refinement statistics (**Table 4.2**), in examples of diffraction density from exchanged D atoms, and in negative H atom diffraction (**Fig. 4.5**).

One contributing factor that allowed rapid data collection from a small crystal was the relatively high symmetry of the space group (orthorhombic $C222_1$), which allowed data to be collected in fewer settings. 75% of neutron structures reported to date have lower symmetry space groups (Bennett *et al.*, 2008), and GTA has the largest reported unit cell. The other major contributing factor was the higher than average degree of isotopic labeling.

That such high quality data were collected for GTA from such relatively small crystals is an indication that neutron crystallography may still be an accessible approach for proteins in which massive crystals cannot be obtained.

4.2.2 H/D exchange penetrates the entire enzyme — A clear contribution to the high quality of the collected data is the level of H/D exchange, which extends throughout the enzyme. All non-proline residues contain an exchangeable amide (**Fig. 4.6a**), so this atom is a useful indicator of overall exchange. GTA exhibits H/D exchange throughout every secondary structure motif. Of the 482 ordered exchangeable H atoms in the enzyme, 400 (or 83%) were observed to be exchanged with occupancies of at least 0.2 after refinement and 98% of backbone amide H atoms were observed as exchanged (**Fig. 4.6bc**). This higher than average backbone exchange observation is also applicable at more stringent occupancies (**Fig. 4.6d**).

One might think a D occupancy of 35% should correspond to equal positive scattering from D (35% of the 6.7 fm scattering length) and negative scattering from H atoms (65% of the -3.8 fm scattering length) (**Fig. 4.6e**). However, this is not what is observed (*eg*. **Fig. 4.6f**) or expected. Peaks from atoms negative to the phasing model are statistically expected at one-half their true weight (Luzzati, 1953; Henderson & Moffat, 1971). As such, D can be confidently assigned as totally exchanged at ≥ 60% refined occupancy, H must be ≥ 90% occupancy for confident assignment of no exchange, and the area between represents an area of partial exchange (Kossiakoff & Spencer, 1981). The 20% occupancy used here for exchange comparison between neutron structures represents double the confidence interval required to assume there has been significant exchange.

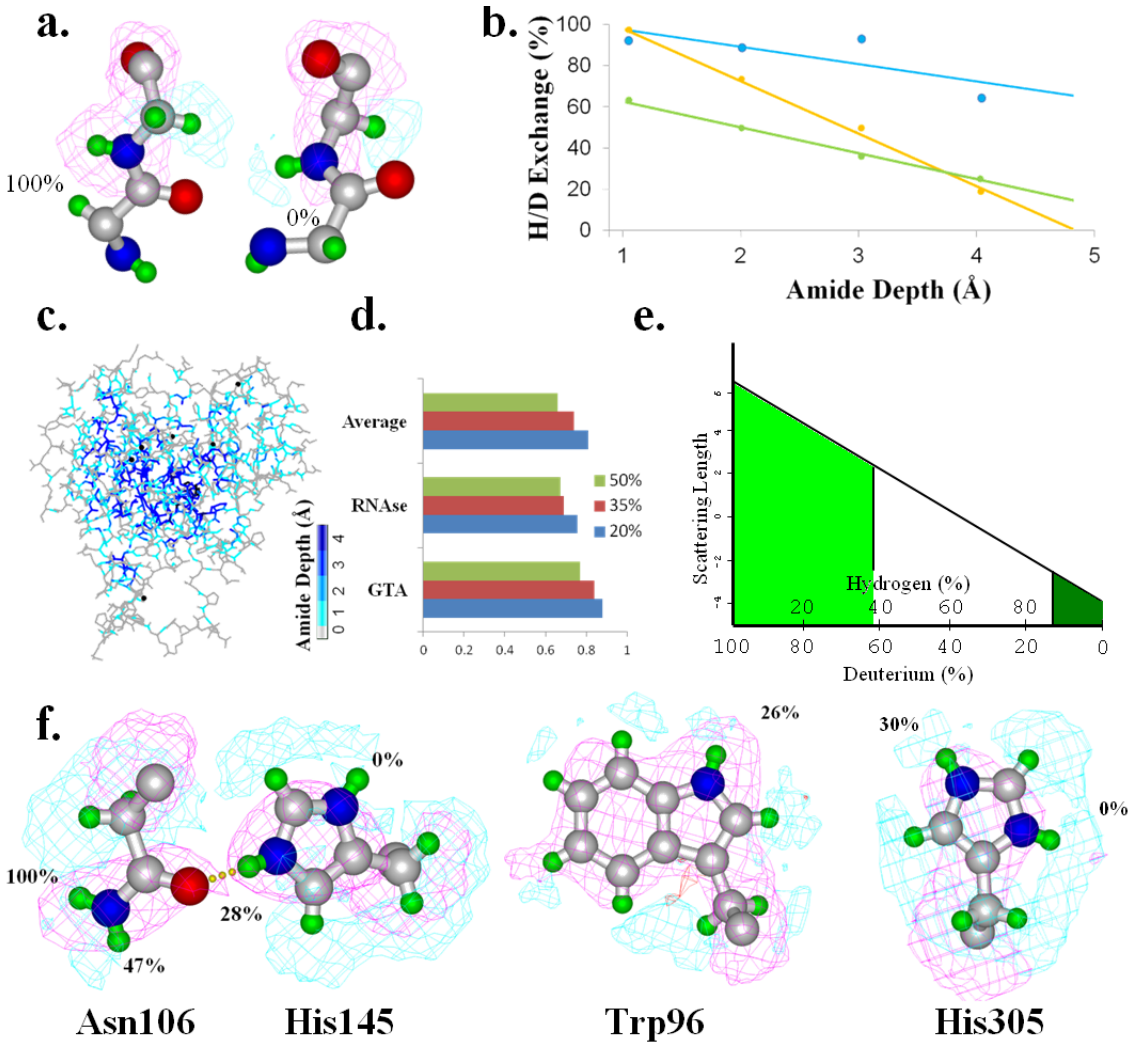


Fig. 4.6a. Examples of peptide bonds displaying 100% and 0% D atom occupancy. **b.** Backbone H/D exchange as a function of depth as observed for GTA (blue, R^2 61%) compared to RNase A (yellow, R^2 97%) and the average of all other neutron structures (green, R^2 92%). **c.** GTA depth from the solvent accessible surface represented as a color gradient, with unexchanged backbone D atoms represented as black spheres. **d.** Overall backbone exchanges of $\geq$ 20, 35 and 50% for GTA, RNAse A and the average of all other neutron structures. **e.** Scattering length as a function of the H/D exchange ratio. Shaded areas of the plot indicate the H/D ratios where unambiguous assignment can be made. **f.** Neutron $2F_o$-F_c (magenta) and -1.0 σ contoured $2F_o$-F_c D omit maps (cyan) displaying positive D signal at low occupancies (labeled).

All previous neutron structures display peptide deuterium occupancy to be inversely and linearly proportional to depth from the solvent exposed surface (Bennett *et al.*, 2008). GTA, however, is a departure from this pattern, and displays a non-uniform distribution (**Fig. 4.6b).**

That H/D exchange in GTA is observed to be unrelated to depth indicates that the native folded state of the enzyme may be more solvent accessible than structures determined from other proteins. The exchange is likely facilitated though a dynamic 'breathing' process that exposes labile H/D atoms to solvent and deuteron channels even in the crystalline state.

Given that successful neutron structure determination typically requires proteins with highly stable structures to form crystals with orders of magnitude more volume than those needed for X-ray diffraction experiments (Blum *et al.*, 2007), it is perhaps to be expected that most proteins for which neutron structures have been reported have had relatively protected cores less permeable to solvent exchange.

Glycosyltransferases, as do most transferase enzymes, possess inherently flexible portions of polypeptide which can contribute to difficulty in crystallization. Paradoxically, this same flexibility that is an impediment to the growth of large GTA crystals may have facilitated the high level of H/D exchange which allowed data collection from a crystal with relatively small volume.

4.2.3 H/D atoms absence supports an S_N2 mechanism — It is well established that in solution associative mechanisms such as S_N2 favor aprotic solvents. In contrast, dissociative mechanisms such as S_Ni require polar protic solvent for initiation and stabilization the intermediate ions (Grote & Hynes, 1980; Gertner *et al.*, 1988; Kim & Hynes, 1990). The active site of GTA is void of proximal solvent molecules, as might be expected to prevent destructive donor hydrolysis. In the absence of solvent, it falls upon the enzyme itself to provide an appropriate environment to facilitate the mechanism of transfer.

The focal point of the GTA active site is fundamentally aprotic. No H/D atoms are located proximal to the reaction center, the anomeric carbon of the donor sugar **C1** or nucleofuge UDP β-phosphate **O3**. In the open conformation of the neutron structure, the closest observed enzymatic H or D atom to the binding position of **C1** is the amide of Gly267 at 4.9 Å. The closest observed enzymatic H or D atom to **O3** is His233 Dε at 5.4 Å, and even this atom would be physically separated from the donor by the acceptor. Modeling H atoms to X-ray structures of the closed conformation brings Arg352 within 3.9 Å (**Fig. 4.7a**), but even this is too distal to initiate UDP dissociation.

As the catalytic site does not contain H atoms to support the generation and stabilization of an oxocarbenium ion, S_Ni would appear to be disfavoured. The active site is more consistent with substitution initiated by acceptor nucleophilic attack with predominantly S_N2 character, such as the recently proposed orthogonal associative mechanism (Schuman *et al*., submitted to *Glycobiology*). This mechanism has the advantage of a lower energy of activation and is consistent with the observed 100% stereoselectivity of glycosyltransfer observed of this class of enzymes.

To the best of our knowledge, our evaluation of proton and dipole distances as a method of distinction between associative and dissociative mechanisms is a novel approach, and one which appears applicable to other enzymes.

Enzymes proposed to proceed through S_N1 mechanisms which also have solved liganded structures are almost invariably hydrolases with solvent molecules in the active site. Correctly oriented enzyme proton dipoles can be modeled very close to the cleavage sites, *eg*. β-amylase (**Fig. 4.7b** modeled from PDB 1BYC,(Mikami *et al.*, 1994)) where the hydroxyl proton of Thr342 may be as close as 2.2 Å to the nucleofuge, and Glu186 presents an observed negative dipole at 2.5 Å.

Correspondingly, the active sites of liganded nonhydrolytic enzymes proposed to utilize S_N2 mechanisms are comparatively void of solvent and enzymatic dipoles do not approach the electronegative nucleofuge at the < 3 Å distance necessary to initiate dissociation. Examples include DNA polymerases (**Fig. 4.7c** modeled from PDB 1ZYQ, (Brieba *et al.*, 2005)) and inverting glycosyltransferases (*eg*. **Fig. 4.7d** modeled from PDB 1TVY/1TW5 (Ramakrishnan *et al.*, 2004)), which gives energetic and evolutionary consistency to the mechanisms utilized by inverting and retaining glycosyltransferases.

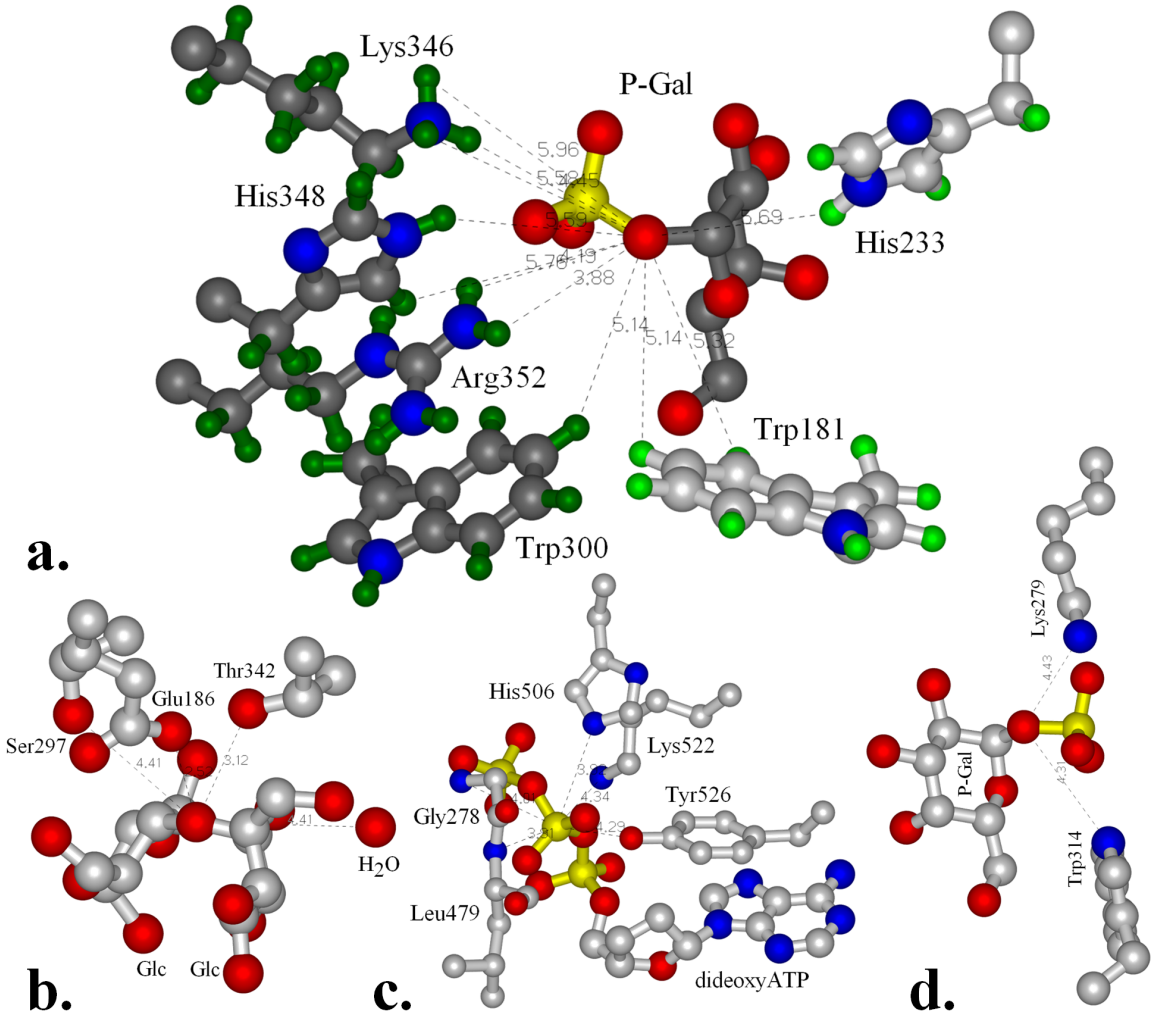


Fig. 4.7. Dipole distances as a mechanism differentiation tool. The aprotic active site of GTA displaying the absence of H atoms within 3.9 Å of nucleofuge phosphate O3. All residues within a 6 Å shell are displayed, lighter residues are directly observed in the neutron structure (PDB 4DHH), darker residues are modeled from X-ray data (PDB 2RJ7). To initiate a dissociative carbocation intermediate dipoles such as H atoms should approach at < 3 Å. **b.** β-amylase, which is thought to proceed through a dissociative mechanism, demonstrates this close dipole approach as Glu186 is observed at 2.5 Å and the Thr342 hydroxyl could present an H atom as close as 2.1 Å. Other associative enzymes such as T7 DNA polymerase **(c)** or inverting glycosyltransferase β4Gal-T1 **(d)** display a void of potential H atoms surrounding their nucleophiles. His and Leu H atoms in **(c)** can be modeled with a high degree of confidence, and both would approach the nucleofuge at > 3 Å. All residues within a 5 Å shell are displayed in **b-d**.

4.3 Reanalysis of previously released Glycosyltransferases data

As the active site of GTA is void of enzymatic dipoles surrounding the nucleophilic acceptor atom (**Nu**) and **C1**, and such an environment is not obviously compatible with dissociation to an ionic intermediate, investigation was undertaken to determine if these observations were applicable to other retaining and inverting glycosyltransferases.

4.3.1 Analysis of available structural data — The closest observed enzymatic dipoles (to donor sugar **C1**, β-phosphate nucleofuge atom **O3** and ring **O5**, which could share the charge of an oxocarbenium ion) for a number of retaining and inverting enzymes are outlined in **Table 4.3**. All lie too far away (~4.5 Å) to initiate dissociation, but may extend the lifetime of an associative transition state.

For retaining enzymes GTA and ManT, the closest **C1** dipoles (Glu303 and Asp167, respectively) could be considered candidate nucleophiles for a double displacement reaction; however, structurally conserved nucleophiles are absent in many reported retaining enzymes including LgtC and Extl2 which have respective Gln and Arg residues in this position. The closest dipoles to donor sugar **C1**, **O5** and β Phosphate **O3** vary considerably, can carry either positive or negative charges, and often their mutation does not inhibit catalysis (*eg.* (Soya, 2010). For example GTB D302C retains 7% activity, E303C retains 21% activity and D303Q retains 1% activity (unpublished). Furthermore, glycosyltransfer still proceeds when **O5** is substituted with sulfur (Adlercreutz *et al*., to be submitted), and as such an intermediate thiocarbenium ion is unlikely to be stabilized to the same extent by donation from sulfur as in a regular oxocarbenium intermediate.

Table 4.3 Active site residue identities and geometric values. With the exception of the GTA neutron diffraction studies (Schuman *et al.*, 2011) hydrogen atoms are not directly observed, so distances are given between centers of non-hydrogen atoms. Italic enzyme names indicate the model did not contain an acceptor molecule.

Stereospecificity	**Inverting**			**Retaining**			
Example enzyme	GlcAT-I	GalT1	*GnT1*	LgtC	GTA	*Extl2*	*ManT*
PDB(s)	1V84 1KWS	1TVY 1TW5	2AM3	1GA8	2RJ7	1OMZ	2WVL
Resolution	1.82 Å 2.10 Å	2.30 Å 2.30 Å	1.80 Å	2.00 Å	1.70 Å	2.10 Å	2.81 Å
Nu – C1	4.4 Å	4.2 Å	4.0 Å[a]	2.2 Å	2.5 Å		
∠Nu-C1-O3	160°	165°	151°[a]	90°	74°	NA	NA
Nu – O3	5.8 Å	5.6 Å	5.4 Å[a]	2.8 Å	2.2 Å		
O3 dipole	H_2O[b]	K279	Y184	H78	K346	H_2O[b]	Y268
O3-dipole	4.4 Å	4.4 Å	5.4 Å	4.7 Å	5.6 Å	3.8 Å	4.4 Å
∠dipole-O3-C1	91°	80°	87°	171°	149°	131°	59°
C1 dipole	H308	W314	D211	Q189	E303	R293	D167
C1-dipole	3.6 Å	4.5 Å	5.2 Å	3.5 Å	4.8 Å	3.7 Å	3.5Å
∠dipole-C1-O3	67°	75°	71°	162°	155°	167°	142°
O5 dipole	R156	W314	D291	Q189	R352	R293	D168
O5-dipole	5.9 Å	3.4 Å	3.9 Å	4.2 Å	5.8 Å	3.2 Å	3.8 Å
∠dipole-O5-C1	113°	123°	96°	82°	83°	97°	68°
c [c]	H2O[b]	H2O[b]	H2O[b]	D103	D211	H2O[b]	NA
∠**b**M**c** [c]	89°	82°	87°	105°	116°	101°	
f [c]	D196	H347	H2O[b]	D105	D213	H2O[b]	N313
∠**b**M**f** [c]	114°	104°	95°	92°	90°	88°	82°

[a] PDB 2AM3 has a glycerol molecule modeled as an acceptor
[b] It is likely that the active species are not actually water molecules, but residues in disordered regions of the polypeptide
[c] **b**, **c** & **f** are octahedral binding partners to the coordinated metal atom **M** as described in **Fig. 3.7**

It is noteworthy that the active site architecture for these dipoles is conserved and distinct for retaining and inverting enzymes, but is consistent with an associative mechanism for both. For the inverting enzymes, the **C1** – **O3** bond lie at angles of 90° or less to the active site dipoles, while for retaining enzymes the dipoles form strongly obtuse angles (**Table 4.3**, **Fig. 4.8**). For the identities of the substrates investigated, please refer to **Table 3.1**.

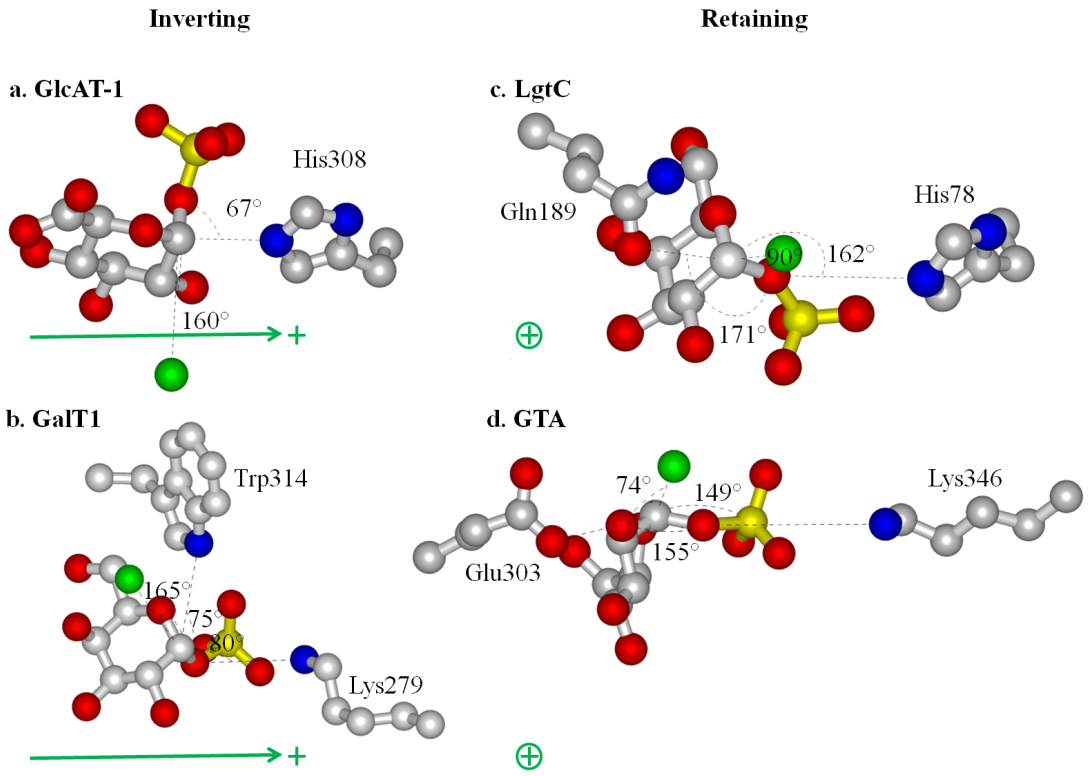


Fig. 4.8 Reaction center dipoles. Opposed to the placement of the acceptor nucleophile (Green spheres), the closest polar residues to leaving group β-phosphate **O3** and **C1** lay acutely for inverting enzymes **(a,b)** and lie nearly in-line for retaining enzymes GTA **(c,d)**. This may help to stabilize the associative intermediates without hindering the opposite angle of attack from the acceptor molecule nucleophile. Also, the **O3** – **C1** vectors lie loosely perpendicular to the enzyme macrodipole vectors to stabilize the inverting transition states (green arrows) (**a,b**), and loosely parallel to stabilize the retaining transition states (**c,d**) (green ⊕, dipole oriented with the cationic end above the page and the anionic end in the page).

Significantly, this positions the dipoles and enzyme macrodipoles, which stabilize the retaining and inverting transition states, orthogonal to each other (**Fig. 4.8**). The orientations of the protein macrodipoles are conserved among retaining enzymes where they lie roughly perpendicular to the nucleophile approach. This would be as expected for stabilization of the developing partial cationic charge while minimally influencing leaving group departure or nucleophilic attack. The macrodipoles of inverting enzymes

are similarly conserved, but are oriented parallel to the line of nucleophile approach, in position to assist such an attack.

Further, the proximity of nucleophile (**Nu**) and electrophile (**C1**) in the retaining enzymes calls into question a dissociative mechanism. The acceptor nucleophiles are observed at distances much less than 3 Å (2.5 Å for GTA and 2.2 Å for LgtC) from donor C1, whereas they should reside greater than 3 Å away to allow UDP dissociation prior to nucleophilic attack (Schramm & Shi, 2001). The computed transition states for glucopyranosyl fluoride solvolyses place the acceptor oxygen 3.02 Å from **C1** in the S_Ni transition state and at 2.25 Å in the associative "front side" transition state (Chan *et al.*, 2011). **Nu** and **C1** can be much greater than 3 Å in a precatalytic conformation as is observed for both analyzed models of inverting enzymes (4.4 Å for GlcAT-I and 4.2 Å for GalT1).

Comparing the biologically active GTA structures to inverting enzymes such as galactosyltrasferases β4GalT1 reveals that the enzymes bind distinct metal-nucleotide-sugar conformers (**Figs. 3.7** & **4.8**), where the metal coordinating angle **∠b-M-c** is less than **∠b-M-f** for inverting enzymes and the inverse is true for retaining enzymes (**Table 4.3**). Inverting enzymes position **C1** for inline nucleophilic attack from the acceptor at an angle nearly 180° to the leaving group, while retaining enzymes position these groups at roughly 90° with respect to the **C1**-leaving group axis.

This is accomplished by inverting and retaining enzymes orienting their metal-nucleotide-sugar binding Rossmann folds approximately perpendicular to one another (**Fig. 4.9**). Although geometrically distinct, in-line (inverting) and orthogonal (retaining) associative transitions states are not dissimilar; retaining enzymes apparently orient their

acceptors to an apical position of a trigonal bipyramidal transition state with the leaving group occupying one of the equatorial positions.

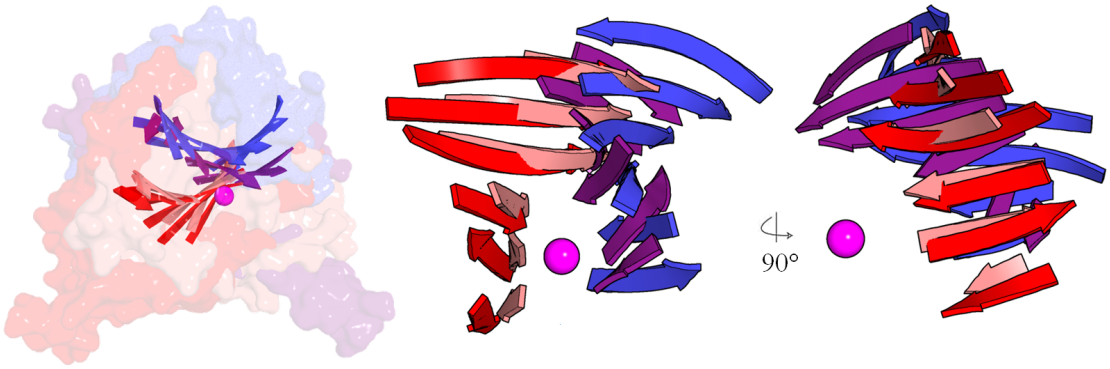


Fig. 4.9 Retaining and inverting enzymes are entirely orthogonal. Overlapping the analyzed GT-A fold glycosyltransferases structures (purple, inverting GlcAT-I; blue, inverting GalT1; red, retaining GTA; pink, retaining LgtC) by the metal ion (magenta sphere) and phosphates reveal that the general architecture of entire inverting or retaining enzymes are skewed by ~90°, which is especially evident when looking at the β-sheets of the metal-nucleotide-sugar binding Rossmann folds.

This orientation is formally accessible as a pseudo-rotation of the trigonal bipyramidal geometry of the S_N2 transition state, which is facilitated by structurally conserved obtusely oriented enzymatic dipoles for retaining enzymes, and is complemented by conserved acute dipoles for retaining enzymes (**Fig. 4.8**). Concurrent closing of the Nu-C1 distance with leaving group loss and concurrent opening of the H-C1-Nu angle would result in associative retention of the donor's anomeric stereochemistry.

A similar reaction pathway has been suggested based on structural studies involving glycomimetic inhibitors (Errey *et al.*, 2010; Lee *et al.*, 2011) and quantum chemical calculations (André *et al.*, 2003), however the proposed mechanisms were still referred to as "S_Ni-like" implying the mechanism proceeds with a rate-limiting dissociative transition state and an intermediate of some finite lifetime.

NMR analysis of donor hydrolysis facilitated by retaining glycosyltransferase enzymes in the absence of acceptor indicates that the cleaved monosaccharides are attacked by water from a retained position (André *et al.*, 2003; Sindhuwinata *et al.*, 2010). This is inconsistent with a dissociative mechanism, as the steric constraints imparted by the enzyme's fully liganded closed position would be mitigated, with solvent molecules occupying both equatorial and axial positions. This assertion can be applied to some liganded closed conformations of retaining enzymes as well: there is no room for dissociation without opening the mobile loops of GTA or LgtC, and an open enzyme conformation during transfer does not support the dissociative rationale that the enzyme architecture might prohibit attack from the inverted position.

4.3.2 Kinetic and energetic analysis — The distinction between an S_Ni and an orthogonal associative mechanism is found in the reaction profile and the timing of bond formation and bond breakage: if nucleophilic attack precedes or is concurrent with leaving group dissociation (André *et al.*, 2003) with no enzymatic cage required to stabilize an oxocarbenium intermediate, then there can be little dissociative character to the mechanism.

The physical organic literature describes the associative mechanism as A_ND_N indicating a single transition state; the S_Ni pathway involving an intermediate is described as $D_N{*}A_{NSS}$ or $D_N^{\ddagger}{*}A_{NSS}$ (Guthrie & Jencks, 1989), differing in the depth of the energetic well corresponding to the intermediate of the two-step process. The available structural data presented above are most consistent with an orthogonal associative mechanism of the A_ND_N type.

The assumed geometric and energetic consequences of the various relevant transitions states and intermediates are sketched in **Fig. 4.10** to visually highlight the distinctions. The proposed orthogonal associative transition state likely lies close in energy to the transition state leading to a S_Ni intermediate; the pathways differ solely in the number of barriers/intermediates invoked.

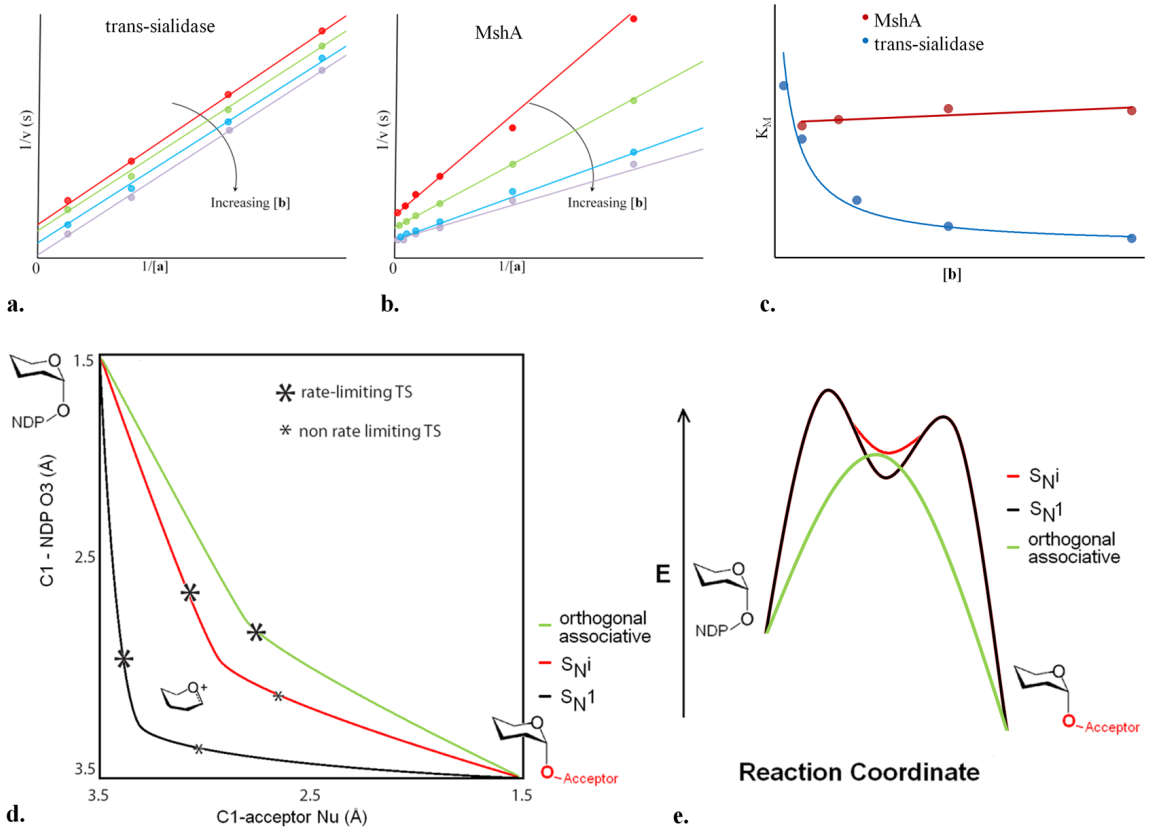


Fig. 4.10 Kinetics and energetics of proposed mechanisms. Two-step mechanisms like double displacement or that of trans-sialidases (**(a)** adapted from (Cheng *et al.*, 2010) should display Lineweaver-Burk plots with parallel slopes for any given concentration of donor ([**donor**]). Retaining glycosyltransferases such as MshA ((**b**) adapted from (Vetting *et al.*, 2008) do not, and have a single K_M regardless of acceptor concentration, which is an indication that retaining glycosyltransferases do not use a double displacement reaction. (**c**) This can also be seen by plotting [**donor**] against acceptor K_M which is ideally constant for single step reactions (red horizontal line, MshA), and parabolic for multistep reactions (blue parabola, trans-sialidase). **(d)** A More-O'Ferrall-Jencks plot illustrating the reaction coordinate potential energy surfaces of the proposed mechanisms. This displays that an orthogonal assosiative mechanism would take the shortest path for retaining glycosyltransfer. **(e)** Comparison of free energy and reaction coordinates for dissociative (S_N1, S_Ni) and orthogonal associative pathways. Both S_N1 and S_Ni displacement would involve an intermediate.

Glycosyltransferases are bi-substrate enzymes and some mechanistic features can be inferred from the overall kinetic schemes observed. Double displacement should follow ping-pong kinetics as it develops a covalent intermediate, which can be identified on a Lineweaver-Burke plot as parallel lines at varied donor substrate concentrations; such a pattern is seen for trans-sialidase, for which such a mechanism has good precedent (**Fig. 4.10a**, adapted from (Cheng *et al.*, 2010). This is not observed for retaining glycosyltransferases such as MshA (**Fig. 4.10b**), in which single values for the acceptor K_M have been reported even when detailed bisubstrate Michaelis–Menten kinetic data has been collected (**Fig. 4.10c**). Kinetic evidence clearly does not support a 2-step mechanism with a covalent intermediate.

The double displacement mechanism has strong precedent for enzymes that do not use metallic co-factors (reviewed in (Nagano *et al.*, 2007) such as glycoside hydrolases, in which covalent glycosyl-enzyme intermediates have been trapped in crystal structures by using fluoridated substrates (*eg*. (Howard *et al.*, 1998; Numao *et al.*, 2003). Such a species should be easier to trap for a glycosyltransferase, as the strong donor leaving groups would leave a covalent sugar-enzyme intermediate in an energy well with the second attack being the rate limiting step (Lairson *et al.*, 2008), and there have been intensive attempts to trap such an intermediate.

The only reports of enzyme-glycosyl intermediates have come from two independent ESI-MS studies, which identified apparent covalent intermediates using postulated nucleophile mutants (Soya, 2010; Lairson *et al.*, 2004). One case showed the covalent species substituted remotely (~10 Å from the acceptor and produced at a rate

much slower than enzymatic turnover, making the catalytic relevance suspect. The other case showed the enzyme-glycosyl intermediate bound to the mutated cysteine; however, such an intermediate has not been observed by means other than MS. It has been suggested that these species could be the results of charged carbocation monosaccharides introduced in the gas phase by the electrospray conditions, which undergo reaction with enzyme nucleophiles to produce such glycosylated species (reviewed in (Di Marco & Bombi, 2006). Kinetic isotope effect data (Lee *et al.*, 2011) are also strong evidence against a stable covalent intermediate, and support a short lived intermediate such as the S_Ni oxocarbenium ion, or even a simple orthogonal associative transition state.

While double-displacement should follow ping-pong bi-substrate kinetics, S_Ni and orthogonal associative mechanisms should follow either random associative or Theorell-Chance mechanisms. The latter is followed by a retaining galactosyltransferase (Guthrie & Jencks, 1989). The distinction between these mechanisms is whether or not the ternary species builds to an extent that is kinetically significant. A developed S_Ni intermediate must avoid water attack, so a well-structured ternary complex formed in a random associative scheme is a reasonable possibility. There is no need for a long-lived ternary complex in an orthogonal associative mechanism. Thus the observation of Theorell-Chance kinetics is consistent with, but does not compel, an orthogonal associative mechanism for the group transfer transition state.

Chapter 5. Conclusions and future work.

The foregoing structural and kinetic analyses are most consistent with an orthogonal associative pathway (**Fig. 5**) for glycosyltransferases that retain anomeric stereochemistry. This has the advantage of unifying associative mechanisms for both retaining and inverting glycosyltransferases without the inconsistencies found in previously proposed mechanisms. From the structural perspective, retaining and inverting enzymes are observed to bind and to act upon distinct conformers of the metal nucleotide sugar complex. The donor substrate trajectory architecture observed for retaining enzymes is conserved so as to present the transferring monosaccharide anomeric electrophile from an orthogonal orientation. The distances observed between the approaching nucleophile and C1 are too close to support dissociation in the structures of both LgtC and GTA. The chemical environment of the active site is aprotic, which disfavors dissociative initiation to glycosyltransfer.

A double displacement mechanism requires an appropriately positioned and structurally conserved nucleophile in the active site, and correspondingly S_Ni requires a structurally conserved base. The active sites of many retaining enzymes do not contain well-positioned candidate nucleophiles, and those that have been proposed are often not sequentially or spatially conserved. Finally, alanine mutagenesis of the proposed nucleophiles does not always abolish enzyme activity (Lairson *et al.*, 2004; Lee *et al.*, 2011).

Structural and kinetic evidence lies in favor of a single step orthogonal associative displacement. The substitution is positioned to initiate with nucleophilic attack and proceeds through a trigonal bipyramidal transition state with incoming acceptor Nu axial

and concurrent transfer of a proton to an equatorial leaving β phosphate **O3**. This makes **C1** the focal point for a pseudorotation that pivots **O3** towards axial and **Nu** towards equatorial for retention of stereochemistry (**Fig. 1.1e**).

This mechanism provides the shortest physical route to glycosyltransfer (**Fig. 5de**), avoiding energy wells and intermediates that have been elusive to previous attempts at detection. An orthogonal associative process avoids generation of even a short-lived oxocarbenium intermediate; it is therefore the simplest of the alternatives and the only one consistent with all available evidence.

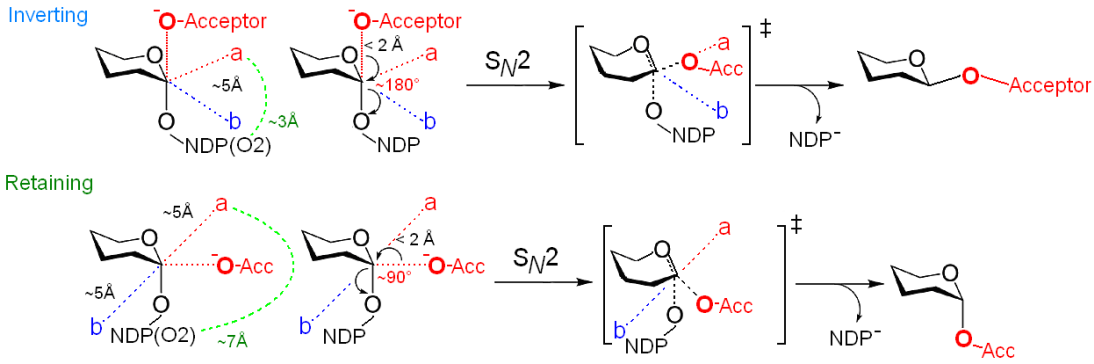
At least three experiments might further support the orthogonal associative mechanism. First, if any enzymatic activity can be detected in a mild aprotic solvent such as acetonitrile, it would be a clear indication that the reaction is proceeding through an associative mechanism. Such solvents may very well abolish enzymatic activity for other reasons, so lack of activity is not necessarily an indication of a dissociative mechanism.

Second, an additional neutron diffraction experiment including the natural H-antigen in the closed conformation should give more support, in part by indicating the location of the transferred proton. Our collaborators Drs. L. Coates and P. Langan have expressed interest in a second neutron structure to be collected at the new protein crystal facility at Oak Ridge. Previously, liganded GTA crystals were not brought for screening as larger crystals cracked even when soaked with ligand very slowly; however, our observations indicate that even moderate sized crystals may be amiable to neutron diffraction experiments. Perdeuteration would be beneficial but not absolutely necessary,

and care must be taken to use a suboptimal metal cofactor such as Mg or Co as Mn scatters negatively (**Table 1.2**) and would distort the active site signal.

Thirdly, the feasibility of the transition state can be computed in a Gausian 09 environment with B3LYP or B2LYP function, experiments which are already underway with Dr. I. Paci at UVic.

Fig. 5. Proposed mechanisms



References

Adachi, M., Ohhara, T., Kurihara, K., Tamada, T., Honjo, E., Okazaki, N., Arai, S., Shoyama, Y., Kimura, K., Matsumura, H., Sugiyama, S., Adachi, H., Takano, K., Mori, Y., Hidaka, K., Kimura, T., Hayashi, Y., Kiso, Y. & Kuroki, R. (2009). Structure of HIV-1 protease in complex with potent inhibitor KNI-272 determined by high-resolution X-ray and neutron crystallography. *Proc Natl Acad Sci U S A* **106**, 4641-4646.

Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C. (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr D Biol Crystallogr* **58**, 1948-1954.

Adams, P. D., Mustyakimov, M., Afonine, P.V., Langan, P. (2009). Generalized X-ray and neutron crystallographic

analysis: more accurate and complete structures for

biological macromolecules. *Acta Crystallogr.* **D65**, 567-573.

Albino, A. P., Houghton, A. N., Eisinger, M., Lee, J. S., Kantor, R. R., Oliff, A. I. & Old, L. J. (1986). Class II histocompatibility antigen expression in human melanocytes transformed by Harvey murine sarcoma virus (Ha-MSV) and Kirsten MSV retroviruses. *J Exp Med* **164**, 1710-1722.

Alfaro, J. A., Zheng, R. B., Persson, M., Letts, J. A., Polakowski, R., Bai, Y., Borisova, S. N., Seto, N. O., Lowary, T. L., Palcic, M. M. & Evans, S. V. (2008). ABO(H) blood group A and B glycosyltransferases recognize substrate via specific conformational changes. *J Biol Chem* **283**, 10097-10108.

André, I., Tvaroska, I. & Carver, J. P. (2003). On the reaction pathways and determination of transition-state structures for retaining [alpha]-galactosyltransferases. *Carbohyd Res* **338**, 865-877.

Angulo, J., Langpap, B., Blume, A., Biet, T., Meyer, B., Krishna, N. R., Peters, H., Palcic, M. M. & Peters, T. (2006). Blood group B galactosyltransferase: insights into substrate binding from NMR experiments. *J Am Chem Soc* **128**, 13529-13538.

Bennett, B., Langan, P., Coates, L., Mustyakimov, M., Schoenborn, B., Howell, E. E. & Dealwis, C. (2006). Neutron diffraction studies of Escherichia coli dihydrofolate reductase complexed with methotrexate. *Proc Natl Acad Sci U S A* **103**, 18493-18498.

Bennett, B. C., Gardberg, A. S., Blair, M. D. & Dealwis, C. G. (2008). On the determinants of amide backbone exchange in proteins: a neutron crystallographic comparative study. *Acta Crystallographica Section D* **64**, 764-783.

Blakeley, M. P., Ruiz, F., Cachau, R., Hazemann, I., Meilleur, F., Mitschler, A., Ginell, S., Afonine, P., Ventura, O. N., Cousido-Siah, A., Haertlein, M., Joachimiak, A., Myles, D. & Podjarny, A. (2008). Quantum model of catalysis based on a mobile proton revealed by subatomic x-ray and neutron diffraction studies of h-aldose reductase. *Proc Natl Acad Sci U S A* **105**, 1844-1848.

Blixt, O., Lavrova, O. I., Mazurov, D. V., Cló, E., Kračun, S. K., Bovin, N. V. & Filatov, A. V. (2011). Analysis of Tn-antigenicity with a panel of new IgM and IgG1 monoclonal antibodies raised against leukemic cells. *Glycobiology*.

Blum, M. M., Koglin, A., Ruterjans, H., Schoenborn, B., Langan, P. & Chen, J. C. (2007). Preliminary time-of-flight neutron diffraction study on diisopropyl fluorophosphatase (DFPase) from Loligo vulgaris. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **63**, 42-45.

Blum, M. M., Mustyakimov, M., Ruterjans, H., Kehe, K., Schoenborn, B. P., Langan, P. & Chen, J. C. (2009). Rapid determination of hydrogen positions and protonation states of diisopropyl fluorophosphatase by joint neutron and X-ray diffraction refinement. *Proc Natl Acad Sci U S A* **106**, 713-718.

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

Bourne, Y. & Henrissat, B. (2001). Glycoside hydrolases and glycosyltransferases: families and functional modules. *Curr Opin Struct Biol* **11**, 593-600.

Brieba, L. G., Kokoska, R. J., Bebenek, K., Kunkel, T. A. & Ellenberger, T. (2005). A lysine residue in the fingers subdomain of T7 DNA polymerase modulates the miscoding potential of 8-oxo-7,8-dihydroguanosine. *Structure* **13**, 1653-1659.

Brockhouse, B. N. (1953). RESONANT SCATTERING OF SLOW NEUTRONS. *Canadian Journal of Physics* **31**, 432-452.

Buzzi, S. & Buzzi, L. (1974). Cancer immunity after treatment of Ehrlich tumor with diphtheria toxin. *Cancer Res* **34**, 3481-3486.

Byrne, G. W., Stalboerger, P. G., Du, Z., Davis, T. R. & McGregor, C. G. (2011). Identification of new carbohydrate and membrane protein antigens in cardiac xenotransplantation. *Transplantation* **91**, 287-292.

Campbell, J. A., Davies, G. J., Bulone, V. & Henrissat, B. (1997). A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. *Biochem J* **326 (Pt 3)**, 929-939.

Cao, Y., Merling, A., Karsten, U., Goletz, S., Punzel, M., Kraft, R., Butschak, G. & Schwartz-Albiez, R. (2008). Expression of CD175 (Tn), CD175s (sialosyl-Tn) and CD176 (Thomsen-Friedenreich antigen) on malignant human hematopoietic cells. *International Journal of Cancer* **123**, 89-99.

Carubia, J. M., Yu, R. K., Macala, L. J., Kirkwood, J. M. & Varga, J. M. (1984). Gangliosides of normal and neoplastic human melanocytes. *Biochem Biophys Res Commun* **120**, 500-504.

The CCP4 suite: programs for protein crystallography. (1994). *Acta Crystallogr D Biol Crystallogr* **50**, 760-763.

Chan, J., Tang, A. & Bennet, A. J. (2011). A Stepwise Solvent-Promoted SNi Reaction of α-d-Glucopyranosyl Fluoride: Mechanistic Implications for Retaining Glycosyltransferases. *Journal of the American Chemical Society* **134**, 1212-1220.

Chelsky, D. & Parsons, S. M. (1975). Stereochemical course of the adenosine triphosphate phosphoribosyltransferase reaction in histidine biosynthesis. *J Biol Chem* **250**, 5669-5673.

Cheng, J., Huang, S., Yu, H., Li, Y., Lau, K. & Chen, X. (2010). Trans-sialidase activity of Photobacterium damsela alpha2,6-sialyltransferase and its application in the synthesis of sialosides. *Glycobiology* **20**, 260-268.

Chester, M. A. & Watkins, W. M. (1969). α-L-Fucosyltransferases in human submaxillary gland and stomach tissues associated with the H, Lea and Leb blood-group characters and ABH secretor status. *Biochem Bioph Res Co* **34**, 835-842.

Ciolczyk-Wierzbicka, D., Bodzioch, M., Gil, D., Zmudzinska, D., Dembinska-Kiec, A. & Laidler, P. (2007). Expression of fucosyltransferases contributes to melanoma invasive phenotype. *Med Chem* **3**, 418-424.

Cipriani, F., Dauvergne, F., Gabriel, A., Wilkinson, C. & Lehmann, M. S. (1994). Image plate detectors for macromolecular neutron diffractometry. *Biophysical Chemistry* **53**, 5-13.

Coates, L., Tuan, H. F., Tomanicek, S., Kovalevsky, A., Mustyakimov, M., Erskine, P. & Cooper, J. (2008). The catalytic mechanism of an aspartic proteinase explored with neutron and X-ray diffraction. *J Am Chem Soc* **130**, 7235-7237.

Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., Baker, D., Popovic, Z. & players, F. (2010). Predicting protein structures with a multiplayer online game. *Nature* **466**, 756-760.

Coutinho, P. M., Deleury, E., Davies, G. J. & Henrissat, B. (2003). An evolving hierarchical family classification for glycosyltransferases. *J Mol Biol* **328**, 307-317.

Coutinho, P. M. a. H., B. (1999). Carbohydrate-Active Enzymes server at URL: http://afmb.cnrs-mrs.fr/~cazy/CAZY/index.html.

DeLano, W. L. (2002). The PyMOL Molecular Graphics System.

Di Marco, V. B. & Bombi, G. G. (2006). Electrospray mass spectrometry (ESI-MS) in the study of metal-ligand solution equilibria. *Mass Spectrom Rev* **25**, 347-379.

Diederichs, K. & Karplus, P. A. (1997). Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat Struct Biol* **4**, 269-275.

Dodson, E., Harding, M. M., Hodgkin, D. C. & Rossmann, M. G. (1966). The crystal structure of insulin. 3. Evidence for a 2-fold axis in rhombohedral zinc insulin. *J Mol Biol* **16**, 227-241.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica Section D* **66**, 486-501.

Errey, J. C., Lee, S. S., Gibson, R. P., Martinez Fleites, C., Barry, C. S., Jung, P. M., O'Sullivan, A. C., Davis, B. G. & Davies, G. J. (2010). Mechanistic insight into enzymatic glycosyl transfer with retention of configuration through analysis of glycomimetic inhibitors. *Angew Chem Int Ed Engl* **49**, 1234-1237.

Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallographica Section D* **62**, 72-82.

Evans, S. V. (1993). SETOR: hardware-lighted three-dimensional solid model representations of macromolecules. *J Mol Graph* **11**, 134-138, 127-138.

Felder, C. E., Prilusky, J., Silman, I. & Sussman, J. L. (2007). A server and database for dipole moments of proteins. *Nucleic Acids Research* **35**, W512-W521.

Fidler, I. J., Gersten, D. M. & Hart, I. R. (1978). The biology of cancer invasion and metastasis. *Adv Cancer Res* **28**, 149-250.

Finne, J., Krusius, T., Rauvala, H. & Jarnefelt, J. (1980). Molecular nature of the blood-group ABH antigens of the human erythrocyte membrane. *Rev Fr Transfus Immunohematol* **23**, 545-552.

Flint, J., Taylor, E., Yang, M., Bolam, D. N., Tailford, L. E., Martinez-Fleites, C., Dodson, E. J., Davis, B. G., Gilbert, H. J. & Davies, G. J. (2005). Structural dissection and high-throughput screening of mannosylglycerate synthase. *Nat Struct Mol Biol* **12**, 608-614.

Freiberger, F., Claus, H., Günzel, A., Oltmann-Norden, I., Vionnet, J., Mühlenhoff, M., Vogel, U., Vann, W. F., Gerardy-Schahn, R. & Stummeyer, K. (2007). Biochemical characterization of a Neisseria meningitidis polysialyltransferase reveals novel functional motifs in bacterial sialyltransferases. *Molecular Microbiology* **65**, 1258-1275.

Gertner, B. J., Wilson, K. R., Zichi, D. A., Lee, S. & Hynes, J. T. (1988). Non-equilibrium solvation in SN1 and SN2 reactions in polar solvents. *Faraday Discussions of the Chemical Society* **85**, 297-308.

Gibson, R. P., Turkenburg, J. P., Charnock, S. J., Lloyd, R. & Davies, G. J. (2002). Insights into trehalose synthesis provided by the structure of the retaining glucosyltransferase OtsA. *Chem Biol* **9**, 1337-1346.

Gill, J. C., Endres-Brooks, J., Bauer, P. J., Marks, W. J., Jr. & Montgomery, R. R. (1987). The effect of ABO blood group on the diagnosis of von Willebrand disease. *Blood* **69**, 1691-1695.

Goncalves, S., Borges, N., Esteves, A. M., Victor, B. L., Soares, C. M., Santos, H. & Matias, P. M. (2010). Structural analysis of Thermus thermophilus HB27 mannosyl-3-phosphoglycerate synthase provides evidence for a second catalytic metal ion and new insight into the retaining mechanism of glycosyltransferases. *J Biol Chem* **285**, 17857-17868.

Gordon, R. D., Sivarajah, P., Satkunarajah, M., Ma, D., Tarling, C. A., Vizitiu, D., Withers, S. G. & Rini, J. M. (2006). X-ray Crystal Structures of Rabbit N-acetylglucosaminyltransferase I (GnT I) in Complex with Donor Substrate Analogues. *Journal of Molecular Biology* **360**, 67-79.

Grote, R. F. & Hynes, J. T. (1980). The stable states picture of chemical reactions. II. Rate constants for condensed and gas phase reaction models. *The Journal of Chemical Physics* **73**, 2715-2732.

Gschaidmeier, H., Seidel, A., Burchell, B. & Bock, K. W. (1995). Formation of mono- and diglucuronides and other glycosides of benzo(a)pyrene-3,6-quinol by V79 cell-expressed human phenol UDP-glucuronosyltransferases of the UGT1 gene complex. *Biochem Pharmacol* **49**, 1601-1606.

Guthrie, R. D. & Jencks, W. P. (1989). IUPAC recommendations for the representation of reaction mechanisms. *Accounts of Chemical Research* **22**, 343-349.

Hakim, S. A., Vyas, G. N., Sanghvi, L. D. & Bhatia, H. M. (1961). Eleven cases of "Bombay" phenotype in six families: suppression of ABO antigen demonstrated in two families. *Transfusion* **1**, 218-222.

Hakomori, S. (1996). Tumor malignancy defined by aberrant glycosylation and sphingo(glyco)lipid metabolism. *Cancer Res* **56**, 5309-5318.

Harrington, P. R., Lindesmith, L., Yount, B., Moe, C. L. & Baric, R. S. (2002). Binding of Norwalk virus-like particles to ABH histo-blood group antigens is blocked by antisera from infected human volunteers or experimentally vaccinated mice. *J Virol* **76**, 12335-12343.

Helliwell, J. R., Habash, J., Cruickshank, D. W. J., Harding, M. M., Greenhough, T. J., Campbell, J. W., Clifton, I. J., Elder, M., Machin, P. A., Papiz, M. Z. & Zurek, S. (1989). The recording and analysis of synchrotron X-radiation Laue diffraction photographs. *J Appl Crystallogr* **22**, 483-497.

Henderson, R. & Moffat, J. K. (1971). The difference Fourier technique in protein crystallography: errors and their treatment. *Acta Crystallographica Section B* **27**, 1414-1420.

Hosseini-Maaf, B., Letts, J. A., Persson, M., Smart, E., LePennec, P. Y., Hustinx, H., Zhao, Z., Palcic, M. M., Evans, S. V., Chester, M. A. & Olsson, M. L. (2007). Structural basis for red cell phenotypic changes in newly identified, naturally occurring subgroup mutants of the human blood group B glycosyltransferase. *Transfusion* **47**, 864-875.

Howard, S., He, S. & Withers, S. G. (1998). Identification of the active site nucleophile in jack bean alpha-mannosidase using 5-fluoro-beta-L-gulosyl fluoride. *J Biol Chem* **273**, 2067-2072.

Itzkowitz, S. H., Bloom, E. J., Kokal, W. A., Modin, G., Hakomori, S. & Kim, Y. S. (1990). Sialosyl-Tn. A novel mucin antigen associated with prognosis in colorectal cancer patients. *Cancer* **66**, 1960-1966.

Itzkowitz, S. H., Dahiya, R., Byrd, J. C. & Kim, Y. S. (1990). Blood group antigen synthesis and degradation in normal and cancerous colonic tissues. *Gastroenterology* **99**, 431-442.

Johal, A. R., Schuman, B., Alfaro, J. A., Borisova, S., Seto, N. O. L. & Evans, S. V. (2012). Sequence-dependent effects of cryoprotectants on the active sites of the human ABO(H) blood group A and B glycosyltransferases. *Acta Crystallographica Section D* **68**, 268-276.

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.

Kannagi, R. (1997). Carbohydrate-mediated cell adhesion involved in hematogenous metastasis of cancer. *Glycoconj J* **14**, 577-584.

Katritzky, A. R. & Brycki, B. E. (1990). The mechanisms of nucleophilic substitution in aliphatic compounds. *Chemical Society Reviews* **19**, 83-105.

Kim, H. J. & Hynes, J. T. (1990). Role of solvent electronic polarization in electron-transfer processes. *The Journal of Physical Chemistry* **94**, 2736-2740.

Kossiakoff, A. A. & Spencer, S. A. (1981). Direct determination of the protonation states of aspartic acid-102 and histidine-57 in the tetrahedral intermediate of the serine proteases: neutron structure of trypsin. *Biochemistry* **20**, 6462-6474.

Lairson, L. L., Chiu, C. P. C., Ly, H. D., He, S. M., Wakarchuk, W. W., Strynadka, N. C. J. & Withers, S. G. (2004). Intermediate trapping on a mutant retaining alpha-galactosyltransferase identifies an unexpected aspartate residue. *Journal of Biological Chemistry* **279**, 28339-28344.

Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. (2008). Glycosyltransferases: Structures, functions, and mechanisms. *Annu Rev Biochem* **77**, 521-555.

Landsteiner, K. (1901). Über Agglutinationserscheinungen normalen menschlichen Blutes. *Wien Klin Wochenschr.* 1132-1134.

Langan, P. & Greene, G. (2004). Protein crystallography with spallation neutrons: collecting and processing wavelength-resolved Laue protein data. *J Appl Crystallogr* **37**, 253-257.

Lee, H. J., Barry, C. H., Borisova, S. N., Seto, N. O. L., Zheng, R. X. B., Blancher, A., Evans, S. V. & Palcic, M. M. (2005). Structural basis for the inactivity of human blood group O-2 glycosyltransferase. *Journal of Biological Chemistry* **280**, 525-529.

Lee, S. S., Hong, S. Y., Errey, J. C., Izumi, A., Davies, G. J. & Davis, B. G. (2011). Mechanistic evidence for a front-side, S(N)i-type reaction in a retaining glycosyltransferase. *Nat Chem Biol* **7**, 631-638.

Lenting, P. J., Pegon, J. N., Christophe, O. D. & Denis, C. V. (2010). Factor VIII and von Willebrand factor – too sweet for their own good. *Haemophilia* **16**, 194-199.

Letts, J. A., Persson, M., Schuman, B., Borisova, S. N., Palcic, M. M. & Evans, S. V. (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 Biol Crystallogr* **63**, 860-865.

Li, M., Liu, X. W., Shao, J., Shen, J., Jia, Q., Yi, W., Song, J. K., Woodward, R., Chow, C. S. & Wang, P. G. (2008). Characterization of a novel (alpha 1,2-fucosyltransferase of Escherichia coli O128 : B12 and functional investigation of its common motif. *Biochemistry* **47**, 378-387.

Lobsanov, Y. D., Romero, P. A., Sleno, B., Yu, B. M., Yip, P., Herscovics, A. & Howell, P. L. (2004). Structure of Kre2p/Mnt1p - A yeast alpha 1,2-mannosyltransferase involved in mannoprotein biosynthesis. *Journal of Biological Chemistry* **279**, 17921-17931.

Loß, A., Bunsmann, P., Bohne, A., Loß, A., Schwarzer, E., Lang, E. & von der Lieth, C.-W. (2002). SWEET-DB: an attempt to create annotated data collections for carbohydrates. *Nucleic Acids Research* **30**, 405-408.

Luzzati, V. (1953). Resolution d'un structure cristalline lorsque les positions d'une partie des atoms sont connues: traitement statistique. *Acta Crystallographica* **6**, 142-152.

Ma, B., Simala-Grant, J. L. & Taylor, D. E. (2006). Fucosylation in prokaryotes and eukaryotes. *Glycobiology* **16**, 158R-184R.

Marcus, S. L., Polakowski, R., Seto, N. O., Leinala, E., Borisova, S., Blancher, A., Roubinet, F., Evans, S. V. & Palcic, M. M. (2003). A single point mutation reverses the donor specificity of human blood group B-synthesizing galactosyltransferase. *J Biol Chem* **278**, 12403-12405.

Marionneau, S., Le Moullac-Vaidye, B. & Le Pendu, J. (2002). Expression of histo-blood group A antigen increases resistance to apoptosis and facilitates escape from immune control of rat colon carcinoma cells. *Glycobiology* **12**, 851-856.

Martinez-Fleites, C., Proctor, M., Roberts, S., Bolam, D. N., Gilbert, H. J. & Davies, G. J. (2006). Insights into the synthesis of lipopolysaccharide and antibiotics through the structures of two retaining glycosyltransferases from family GT4. *Chemistry & Biology* **13**, 1143-1152.

Mas, E., Pasqualini, E., Caillol, N., El Battari, A., Crotte, C., Lombardo, D. & Sadoulet, M. O. (1998). Fucosyltransferase activities in human pancreatic tissue:

comparative study between cancer tissues and established tumoral cell lines. *Glycobiology* **8**, 605-613.

Mathieu, S., Gerolami, R., Luis, J., Carmona, S., Kol, O., Crescence, L., Garcia, S., Borentain, P. & El-Battari, A. (2007). Introducing alpha(1,2)-linked fucose into hepatocarcinoma cells inhibits vasculogenesis and tumor growth. *International Journal of Cancer* **121**, 1680-1689.

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). Phaser crystallographic software. *J Appl Crystallogr* **40**, 658-674.

Mikami, B., Degano, M., Hehre, E. J. & Sacchettini, J. C. (1994). Crystal structures of soybean beta-amylase reacted with beta-maltose and maltal: active site components and their apparent roles in catalysis. *Biochemistry* **33**, 7779-7787.

Mulichak, A. M., Losey, H. C., Walsh, C. T. & Garavito, R. M. (2001). Structure of the UDP-Glucosyltransferase GtfB that modifies the heptapeptide aglycone in the biosynthesis of vancomycin group antibiotics. *Structure* **9**, 547-557.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-255.

Nagano, N., Noguchi, T. & Akiyama, Y. (2007). Systematic comparison of catalytic mechanisms of hydrolysis and transfer reactions classified in the EzCatDB database. *Proteins: Structure, Function, and Bioinformatics* **66**, 147-159.

Nakahara, T., Hindsgaul, O., Palcic, M. M. & Nishimura, S. I. (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.

Niimura, N. & Bau, R. (2008). Neutron protein crystallography: beyond the folding structure of biological macromolecules. *Acta Crystallogr A* **64**, 12-22.

Numao, S., Kuntz, D. A., Withers, S. G. & Rose, D. R. (2003). Insights into the mechanism of Drosophila melanogaster Golgi alpha-mannosidase II through the structural analysis of covalent reaction intermediates. *Journal of Biological Chemistry* **278**, 48074-48083.

O'Donnell, J. & Laffan, M. A. (2001). The relationship between ABO histo-blood group, factor VIII and von Willebrand factor. *Transfusion Med* **11**, 343-351.

O'Donnell, J. S., McKinnon, T. A. J., Crawley, J. T. B., Lane, D. A. & Laffan, M. A. (2005). Bombay phenotype is associated with reduced plasma-VWF levels and an increased susceptibility to ADAMTS13 proteolysis. *Blood* **106**, 1988-1991.

Olsson, M. L., Irshaid, N. M., Hosseini-Maaf, B., Hellberg, Å., Moulds, M. K., Sareneva, H. & Chester, M. A. (2001). Genomic analysis of clinical samples with serologic ABO blood grouping discrepancies: identification of 15 novel A and B subgroup alleles. *Blood* **98**, 1585-1593.

Otwinowski, Z. & Minor, W. (1997). Vol. Volume 276*, Methods in Enzymology*, edited by Charles W. Carter, Jr., pp. 307-326. Academic Press.

Pabst, M. J., Habig, W. H. & Jakoby, W. B. (1974). Glutathione S-Transferase A. *Journal of Biological Chemistry* **249**, 7140-7148.

Pak, J. E., Arnoux, P., Zhou, S., Sivarajah, P., Satkunarajah, M., Xing, X. & Rini, J. M. (2006). X-ray crystal structure of leukocyte type core 2 beta1,6-N-

acetylglucosaminyltransferase. Evidence for a convergence of metal ion-independent glycosyltransferase mechanism. *J Biol Chem* **281**, 26693-26701.

Pastores, G. M., Elstein, D., Hrebicek, M. & Zimran, A. (2007). Effect of miglustat on bone disease in adults with type 1 Gaucher disease: a pooled analysis of three multinational, open-label studies. *Clin Ther* **29**, 1645-1654.

Patenaude, S. I., Seto, N. O., Borisova, S. N., Szpacenko, A., Marcus, S. L., Palcic, M. M. & Evans, S. V. (2002). The structural basis for specificity in human ABO(H) blood group biosynthesis. *Nat Struct Biol* **9**, 685-690.

Pedersen, L. C., Darden, T. A. & Negishi, M. (2002). Crystal structure of beta 1,3-glucuronyltransferase I in complex with active donor substrate UDP-GlcUA. *J Biol Chem* **277**, 21869-21873.

Pedersen, L. C., Dong, J., Taniguchi, F., Kitagawa, H., Krahn, J. M., Pedersen, L. G., 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. *Journal of Biological Chemistry* **278**, 14420-14428.

Persson, K., Ly, H. D., Dieckelmann, M., Wakarchuk, W. W., Withers, S. G. & Strynadka, N. C. (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.

Pflugrath, J. W. (1999). The finer things in X-ray diffraction data collection. *Acta Crystallogr D Biol Crystallogr* **55**, 1718-1725.

Pintar, A., Carugo, O. & Pongor, S. (2003). DPX: for the analysis of the protein core. *Bioinformatics* **19**, 313-314.

Pohl, R., Antognini, A., Nez, F., Amaro, F. D., Biraben, F., Cardoso, J. M. R., Covita, D. S., Dax, A., Dhawan, S., Fernandes, L. M. P., Giesen, A., Graf, T., Hänsch, T. W., Indelicato, P., Julien, L., Kao, C.-Y., Knowles, P., Le Bigot, E.-O., Liu, Y.-W., Lopes, J. A. M., Ludhova, L., Monteiro, C. M. B., Mulhauser, F., Nebel, T., Rabinowitz, P., dos Santos, J. M. F., Schaller, L. A., Schuhmann, K., Schwob, C., Taqqu, D., Veloso, J. F. C. A. & Kottmann, F. (2010). The size of the proton. *Nature* **466**, 213-216.

Raetz, C. R. H. & Whitfield, C. (2002). Lipopolysaccharide endotoxins. *Annu Rev Biochem* **71**, 635-700.

Ramakrishnan, B., Boeggeman, E. & Qasba, P. K. (2004). Effect of the Met344His mutation on the conformational dynamics of bovine beta-1,4-galactosyltransferase: crystal structure of the Met344His mutant in complex with chitobiose. *Biochemistry* **43**, 12513-12522.

Ravindranath, M. H., Tsuchida, T., Morton, D. L. & Irie, R. F. (1991). Ganglioside Gm3-Gd3 Ratio as an Index for the Management of Melanoma. *Cancer* **67**, 3029-3035.

Reinert, D. J., Jank, T., Aktories, K. & Schulz, G. E. (2005). Structural basis for the function of Clostridium difficile toxin B. *Journal of Molecular Biology* **351**, 973-981.

Rosen, M. L., Edman, M., Sjostrom, M. & Wieslander, A. (2004). Recognition of fold and sugar linkage for glycosyltransferases by multivariate sequence analysis. *Journal of Biological Chemistry* **279**, 38683-38692.

Sadler, J. E. (1998). Biochemistry and genetics of von Willebrand factor. *Annu Rev Biochem* **67**, 395-424.

Sarode, R., Goldstein, J., Sussman, I. I., Nagel, R. L. & Tsai, H. M. (2000). Role of A and B blood group antigens in the expression of adhesive activity of von Willebrand factor. *Brit J Haematol* **109**, 857-864.

Scheller, H. V. & Ulvskov, P. (2010). Hemicelluloses. *Annual Review of Plant Biology* **61**, 263-289.

Schober, H. (2009). *Neutron Scattering Applications and Techniques*, edited by L. Liang, R. Rinaldi & H. Schober, pp. 37-104. Springer US.

Schramm, V. L. & Shi, W. (2001). Atomic motion in enzymatic reaction coordinates. *Current Opinion in Structural Biology* **11**, 657-665.

Schuman, B., Alfaro, J. A. & Evans, S. V. (2007). Glycosyltransferase structure and function. *Bioactive Conformation I* **272**, 217-257.

Schuman, B., Fisher, S. Z., Kovalevsky, A., Borisova, S. N., Palcic, M. M., Coates, L., Langan, P. & Evans, S. V. (2011). Preliminary joint neutron time-of-flight and X-ray crystallographic study of human ABO(H) blood group A glycosyltransferase. *Acta Crystallographica Section F* **67**, 258-262.

Schuman, B., Persson, M., Landry, R. C., Polakowski, R., Weadge, J. T., Seto, N. O., Borisova, S. N., Palcic, M. M. & Evans, S. V. (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.

Schuttelkopf, A. W. & van Aalten, D. M. (2004). PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr* **60**, 1355-1363.

Sears, V. F. (1992). Neutron scattering lengths and cross sections. *Neutron News* **3**, 26-37.

Seltsam, A., Hallensleben, M., Kollmann, A. & Blasczyk, R. (2003). The nature of diversity and diversification at the ABO locus. *Blood* **102**, 3035-3042.

Seto, N. O., Compston, C. A., Evans, S. V., Bundle, D. R., Narang, S. A. & Palcic, M. M. (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, N. O., Compston, C. A., Szpacenko, A. & Palcic, M. M. (2000). Enzymatic synthesis of blood group A and B trisaccharide analogues. *Carbohydr Res* **324**, 161-169.

Seto, N. O. L., Palcic, M. M., Compston, C. A., Li, H., Bundle, D. R. & Narang, S. A. (1997). Sequential interchange of four amino acids from blood group B to blood group A glycosyltransferase boosts catalytic activity and progressively modifies substrate recognition in human recombinant enzymes. *Journal of Biological Chemistry* **272**, 14133-14138.

Shima, M., Fujimura, Y., Nishiyama, T., Tsujiuchi, T., Narita, N., Matsui, T., Titani, K., Katayama, M., Yamamoto, F. & Yoshioka, A. (1995). Abo Blood-Group Genotype and Plasma Von-Willebrand-Factor in Normal Individuals. *Vox Sang* **68**, 236-240.

Shu, F., Ramakrishnan, V. & Schoenborn, B. P. (2000). Enhanced visibility of hydrogen atoms by neutron crystallography on fully deuterated myoglobin. *Proceedings of the National Academy of Sciences* **97**, 3872-3877.

Sindhuwinata, N., Munoz, E., Munoz, F. J., Palcic, M. M., 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, M. L. (1990). Catalytic Mechanisms of Enzymatic Glycosyl Transfer. *Chem Rev* **90**, 1171-1202.

Sinnott, M. L. & Jencks, W. P. (1980). Solvolysis of D-Glucopyranosyl Derivatives in Mixtures of Ethanol and 2,2,2-Trifluoroethanol. *Journal of the American Chemical Society* **102**, 2026-2032.

Sommer, N., Depping, R., Piotrowski, M. & Ruger, W. (2004). Bacteriophage T4 alpha-glucosyltransferase: a novel interaction with gp45 and aspects of the catalytic mechanism. *Biochem Bioph Res Co* **323**, 809-815.

Soya, N., Fang, Y., Palcic, M.M., Klassen, J. (2010). Detection of Covalent Intermediates Reveals Mechanism of Retaining Glycosyltransferases. *J Am Chem Soc*.

Soya, N., Shoemaker, G. K., Palcic, M. M. & Klassen, J. S. (2009). Comparative study of substrate and product binding to the human ABO(H) blood group glycosyltransferases. *Glycobiology* **19**, 1224-1234.

Taku, A., Hirsch, T. M. & Fan, D. P. (1980). Dissociation and reconstitution of membranes synthesizing the peptidoglycan of Escherichia coli. Lipid dependence of the synthetic enzymes. *J Biol Chem* **255**, 2848-2854.

Tasaki, M., Yoshida, Y., Miyamoto, M., Nameta, M., Cuellar, L. M., Xu, B., Zhang, Y., Yaoita, E., Nakagawa, Y., Saito, K., Yamamoto, T. & Takahashi, K. (2009). Identification and characterization of major proteins carrying ABO blood group antigens in the human kidney. *Transplantation* **87**, 1125-1133.

Tomanicek, S. J., Blakeley, M. P., Cooper, J., Chen, Y., Afonine, P. V. & Coates, L. (2010). Neutron Diffraction Studies of a Class A [beta]-Lactamase Toho-1 E166A/R274N/R276N Triple Mutant. *Journal of Molecular Biology* **396**, 1070-1080.

Umesiri, F. E., Sanki, A. K., Boucau, J., Ronning, D. R. & Sucheck, S. J. (2010). Recent advances toward the inhibition of mAG and LAM synthesis in Mycobacterium tuberculosis. *Med Res Rev* **30**, 290-326.

Unligil, U. M., Zhou, S., Yuwaraj, S., Sarkar, M., Schachter, H. & Rini, J. M. (2000). X-ray crystal structure of rabbit N-acetylglucosaminyltransferase I: catalytic mechanism and a new protein superfamily. *EMBO J* **19**, 5269-5280.

Vetting, M. W., Frantom, P. A. & Blanchard, J. S. (2008). Structural and enzymatic analysis of MshA from Corynebacterium glutamicum: substrate-assisted catalysis. *J Biol Chem* **283**, 15834-15844.

Watkins, W. M. & Morgan, W. T. J. (1957). Specific Inhibition Studies Relating to the Lewis Blood-Group System. *Nature* **180**, 1038-1040.

Weil, R., 3rd, Nozawa, M., Chernack, W., McIntosh, R. & Reemtsma, K. (1973). Effect of Concanavalin A on rat heart allografts. *Surg Forum* **24**, 194-196.

Wennekes, T., van den Berg, R. J., Donker, W., van der Marel, G. A., Strijland, A., Aerts, J. M. & Overkleeft, H. S. (2007). Development of adamantan-1-yl-methoxy-

functionalized 1-deoxynojirimycin derivatives as selective inhibitors of glucosylceramide metabolism in man. *J Org Chem* **72**, 1088-1097.

Werther, J. L., Rivera-MacMurray, S., Bruckner, H., Tatematsu, M. & Itzkowitz, S. H. (1994). Mucin-associated sialosyl-Tn antigen expression in gastric cancer correlates with an adverse outcome. *Br J Cancer* **69**, 613-616.

Werther, J. L., Riveramacmurray, S., Bruckner, H., Tatematsu, M. & Itzkowitz, S. H. (1994). Mucin-Associated Sialosyl-Tn Antigen Expression in Gastric-Cancer Correlates with an Adverse Outcome. *Brit J Cancer* **69**, 613-616.

Wong, S. M., St Michael, F., Cox, A., Ram, S. & Akerley, B. J. (2011). ArcA-Regulated Glycosyltransferase Lic2B Promotes Complement Evasion and Pathogenesis of Nontypeable Haemophilus influenzae. *Infect Immun* **79**, 1971-1983.

Yagi, D., Yamada, T., Kurihara, K., Ohnishi, Y., Yamashita, M., Tamada, T., Tanaka, I., Kuroki, R. & Niimura, N. (2009). A neutron crystallographic analysis of phosphate-free ribonuclease A at 1.7 A resolution. *Acta Crystallographica Section D* **65**, 892-899.

Yamamoto, F., Clausen, H., White, T., Marken, J. & Hakomori, S. (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 B transferases is based on amino acid substitutions. *J Biol Chem* **265**, 19257-19262.

Yamamoto, F., McNeill, P. D., Yamamoto, M., Hakomori, S., Harris, T., Judd, W. J. & Davenport, R. D. (1993). Molecular genetic analysis of the ABO blood group system: 1. Weak subgroups: A3 and B3 alleles. *Vox Sang* **64**, 116-119.

Yamamoto, H., Kaneko, Y., Rebbaa, A., Bremer, E. G. & Moskal, J. R. (1997). alpha 2,6-sialyltransferase gene transfection into a human glioma cell line (U373 MG) results in decreased invasivity. *J Neurochem* **68**, 2566-2576.

Yazer, M. H., Hosseini-Maaf, B. & Olsson, M. L. (2008). Blood grouping discrepancies between ABO genotype and phenotype caused by O alleles. *Current Opinion in Hematology* **15**, 618-624 610.1097/MOH.1090b1013e3283127062.

Yazer, M. H. & Palcic, M. M. (2005). The importance of disordered loops in ABO glycosyltransferases. *Transfus Med Rev* **19**, 210-216.

Yoshida, A., Kobayashi, K., Manya, H., Taniguchi, K., Kano, H., Mizuno, M., Inazu, T., Mitsuhashi, H., Takahashi, S., Takeuchi, M., Herrmann, R., Straub, V., Talim, B., Voit, T., Topaloglu, H., Toda, T. & Endo, T. (2001). Muscular dystrophy and neuronal migration disorder caused by mutations in a glycosyltransferase, POMGnT1. *Dev Cell* **1**, 717-724.

Yoshida, M., Itano, N., Yamada, Y. & Kimata, K. (2000). In vitro synthesis of hyaluronan by a single protein derived from mouse HAS1 gene and characterization of amino acid residues essential for the activity. *Journal of Biological Chemistry* **275**, 497-506.

Zhou, Y.-F., Eng, E. T., Nishida, N., Lu, C., Walz, T. & Springer, T. A. (2011). A pH-regulated dimeric bouquet in the structure of von Willebrand factor. *EMBO J* **30**, 4098-4111.

Appendix I: Permissions

SPRINGER LICENSE
TERMS AND CONDITIONS

Feb 27, 2012

This is a License Agreement between Brock Schuman ("You") and Springer ("Springer") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer, and the payment terms and conditions.

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

License Number	2856771255340
License date	Feb 26, 2012
Licensed content publisher	Springer
Licensed content publication	Springer eBook
Licensed content title	Glycosyltransferase Structure and Function
Licensed content author	Brock Schuman
Licensed content date	Nov 30, 2006
Type of Use	Thesis/Dissertation
Portion	Figures
Author of this Springer article	Yes and you are the sole author of the new work
Order reference number	
Title of your thesis / dissertation	Structural and Functional Analysis of Human ABO Blood Group Glycosyltransferases and Antigens Structural and Functional Analysis of Human ABO Blood Group Glycosyltransferases and Antigens
Expected completion date	Mar 2012
Estimated size(pages)	80
Total	0.00 USD
Terms and Conditions	

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK500727060.
Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

For suggestions or comments regarding this order, contact RightsLink Customer Support:customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.

As an author of an article in Acta Cryst. or other
IUCr journals you can reuse material in your articles without asking further permission, provided you cite the original article. However, I am also providing formal permission below in case this is needed.

Permission is hereby granted, on behalf of the IUCr, for you to reproduce the material specified below, subject to the following conditions:

1. Reproduction is intended in a primary journal, secondary journal, CD-ROM, book or thesis.

2. The original article in which the material appeared is cited.

3. IUCr's copyright permission is indicated next to the Figure in print. In electronic form, this acknowledgement should be visible at the same time as the Figure, and must be hyperlinked to the article (http://dx.doi.org/10.1107/S17443091100512 98?).

Material to be reproduced:
Figures from
Acta Cryst. (2011). F67, 258-262 [doi:10.1107/S1744309110051298]
Preliminary joint neutron time-of-flight and X-ray crystallographic study of human ABO(H) blood group A glycosyltransferase
B. Schuman, S. Z. Fisher, A. Kovalevsky, S. N. Borisova, M. M. Palcic, L.

Coates, P. Langan and S. V. Evans

Best wishes

Peter Strickland

Managing Editor

IUCr Journals

ELSEVIER LICENSE
TERMS AND CONDITIONS

Feb 27, 2012

This is a License Agreement between Brock Schuman ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

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

Supplier	Elsevier Limited The Boulevard,Langford Lane Kidlington,Oxford,OX5 1GB,UK
Registered Company Number	1982084
Customer name	Brock Schuman
Customer address	113 -1718 Newton st Victoria, BC V8R 2R2
License number	2856870192530
License date	Feb 27, 2012
Licensed content publisher	Elsevier
Licensed content publication	Journal of Molecular Biology
Licensed content title	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
Licensed content author	Brock Schuman,Mattias Persson,Roxanne C. Landry,Robert Polakowski,Joel T. Weadge,Nina O.L. Seto,Svetlana N. Borisova,Monica M. Palcic,Stephen V. Evans
Licensed content date	17 September 2010
Licensed content volume number	402
Licensed content issue number	2
Number of pages	13
Start Page	399
End Page	411
Type of Use	reuse in a thesis/dissertation
Intended publisher of new work	other
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
Format	print
Are you the author of this Elsevier article?	Yes
Will you be translating?	No
Order reference number	
Title of your thesis/dissertation	Structural and Functional Analysis of Human ABO Blood Group Glycosyltransferases and Antigens Structural and Functional Analysis of Human ABO Blood Group Glycosyltransferases and Antigens
Expected completion date	Mar 2012
Estimated size (number of pages)	80
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD
VAT/Local Sales Tax	0.0 USD / 0.0 GBP
Total	0.00 USD

Appendix II

X-ray data collection and refinement statistics for evaluated ABO glycosyltransferases.
All crystals belonged to space group $C222_1$ with unit cells dimensions of 52.4-53.5 Å, 149.0-151.9 Å, 78.2-80.2 Å.

Enzyme	**GTB/C80S/C196S[a]**					**GTB/C80S/C196S/C209S[a]**				
substrates	**unliganded**	**H antigen**	**UDP**	**UDP + H**	**UDP-Gal+DA**	**unliganded**	**H antigen**	**UDP**	**UDP + H**	**UDP-Gal+DA**
Resolution (Å)	20 – 1.55 (1.61–1.55)	20 – 1.81 (1.87–1.81)	20 – 1.90 (1.97–1.90)	20 – 1.56 (1.62–1.56)	20 – 1.40 (1.45–1.40)	20 – 1.79 (1.85–1.79)	20 – 1.48 (1.53–1.48)	20 – 1.49 (1.54–1.59)	20 – 1.93 (2.00–1.93)	20 – 1.60 (1.66–1.60)
Unique reflections	43878	27621	24723	43699	58949	28441	51095	49738	22845	39424
Completeness (%)	96.4 (98.1)	95.6 (93.9)	99.0 (99.7)	97.3 (99.6)	95.3 (76.9)	96.2 (98.1)	97.2 (99.4)	96.7 (96.7)	95.8 (98.5)	94.6 (89.5)
R-merge (%)	4.2 (28.8)	7.4 (31.7)	7.8 (31.4)	4.1 (30.0)	3.5 (32.5)	5.1 (29.9)	3.8 (24.9)	3.5 (39.7)	5.1 (29.8)	4.6 (26.9)
I/σ(I)	17.1 (3.8)	12.0 (4.1)	10.5 (4.2)	17.3 (3.6)	19.2 (3.2)	14.0 (4.0)	14.8 (4.1)	17.2 (2.8)	15.2 (4.1)	15.3 (4.1)
R-work (%)	18.98	17.32	18.41	18.85	19.45	16.52	20.19	16.92	17.45	18.40
R free (%)	22.76	22.23	22.94	21.72	22.96	21.97	23.82	22.57	22.81	22.28
bond lengths (Å)	0.028	0.025	0.025	0.030	0.030	0.023	0.027	0.028	0.023	0.030
bond angle (°)	2.305	1.931	2.189	2.247	2.572	1.80	2.178	1.961	2.202	2.447
PDB ID	3I0C	3I0E	3I0D	3I0F	3I0G	3I0H	3I0J	3I0I	3I0K	3I0L

Enzyme	GTA[b]		AABB[b]		ABBA[b]		ABBB[b]		GTB[b]	
Cryoprotectant	**MPD**	**Glycerol**	**MPD**	**Glycerol**	**MPD**	**Glycerol**	**MPD**	**Glycerol**	**MPD**	**Glycerol**
Resolution (Å)	20 – 1.86	20 – 1.49	20 – 1.43	20 – 1.45	20 – 1.47	20 – 1.42	20 – 1.49	20 – 1.50	20-1.55	20 – 1.90
R_{merge} (%)	5.5 (34.0)	4.3 (36.4)	3.1 (30.1)	4.2 (37.1)	5.8 (29.1)	3.0 (29.5)	3.2 (26.5)	3.0 (25.2)	5.5 (42.9)	5.9 (32.5)
Completeness (%)	93.8 (96.2)	96.9 (97.8)	95.6 (89.2)	96.3 (96.0)	96.9 (97.8)	98.2 (95.5)	95.5 (93.5)	96.9 (94.7)	96.9 (94.7)	95.2 (98.2)
Unique Reflections	25287	50161	55579	53446	52036	58371	49296	48983	44205	23770
R_{work} (%)	19.1	19.7	20.7	20.1	19.1	20.5	19.5	20.4	19.2	17.9
R_{free} (%)	24.1	22.7	22.3	22.9	22.3	22.8	22.2	23.2	22.5	22.8
rms bond (Å)	0.026	0.029	0.031	0.031	0.030	0.032	0.032	0.030	0.028	0.026
rms angle (°)	2.114	2.339	2.499	2.560	2.387	2.491	2.517	2.363	2.210	1.952
PDB ID	3SXG	3SXE	3SX5	3SX3	3SX8	3SX7	3SXB	3SXA	3SXD	3SXC

Enzyme	GTA/C209A/E303C	GTB/M214G	GTA/P234S	GTA/L266V
Resolution (Å)	2.00	1.67	1.72	1.55
R_{merge} (%)	5.8 (20.2)	4.4 (20.6)	4.8 (28.2)	4.2 (25.8)
Completeness (%)	98.2 (95.0)	96.9 (99.8)	83.6 (91.1)	97.3 (97.4)
Unique Reflections	21388	35676	27651	193712

[a] 1st reported in (Schuman *et al.*, 2010)
[b] 1st reported in (Johal *et al.*, 2012)
[c] where **A** and **B** represent the A and B antigens
Structures without ordinal indicators are unpublished.