ABO(H) blood group A and B glycosyltransferases recognize substrate via specific conformational changes

Javier A. Alfaro1, Ruixiang Blake Zheng2, Mattias Persson3, James A. Letts1, Robert Polakowski2, Yu Bai2, Svetlana N. Borisova1, Nina O. L. Seto1,4, Todd L. Lowary2, Monica M. Palcic2,3, Stephen V. Evans1§

1 Department of Biochemistry & Microbiology, University of Victoria, Victoria, British Columbia V8W 3P6, Canada.

2 Alberta Ingenuity Centre for Carbohydrate Science and Department of Chemistry, University of Alberta, Edmonton, Alberta, T6G 2G2 Canada.

3 Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Valby, Copenhagen, Denmark.

4 Institute for Biological Sciences, National Research Council of Canada, Ottawa, Ontario K1A 0R6, Canada

Running title: Polypeptide movement in GTA and GTB

* This work was supported in part by grants from the Natural Sciences and Engineering Research Council of Canada, the Danish Natural Science Research Council, Alberta Ingenuity, and the Canadian Institutes of Health Research. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Recipient of a senior scholarship from the Michael Smith Foundation for Health Research. To whom correspondence may be addressed: Dept. of Biochemistry and Microbiology, University of Victoria, Victoria, British Columbia V8W 3P6, Canada. E-mail: svevans@uvic.ca.

The atomic coordinates and structure factors (codes 2RIT, 2RIX, 2RJY, 2RJ8, 2RJ9, 2RIZ, 2RJ0, 2RJ1, 2RJ4, 2RJ5, 2RJ6 and 2RJ7) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).
The final step in the enzymatic synthesis of the ABO(H) blood group A and B antigens is catalyzed by two closely related glycosyltransferases, an \( \alpha-(1\rightarrow3)-N\)-acetylgalactosaminyltransferase (GTA\(^*\)) and an \( \alpha-(1\rightarrow3)\)-galactosyltransferase (GTB). Of their 354 amino acid residues, GTA and GTB differ by only four ‘critical’ residues. High-resolution structures for GTB and the GTA/GTB chimeric enzymes GTB/G176R and GTB/G176R/G235S bound to a panel of donor and acceptor analog substrates reveal ‘open’, ‘semi-closed’ and ‘closed’ conformations as the enzymes go from the unliganded to the liganded states. In the ‘open’ form the internal polypeptide loop (amino acid residues 177–195) adjacent to the active site in the unliganded or H-antigen-bound enzymes is composed of two \( \alpha\)-helices spanning Arg180-Met186 and Arg188-Asp194 respectively. The ‘semi-closed’ and closed forms of the enzymes are generated by binding of UDP or of UDP and H-antigen analogs respectively, and show that these helices merge to form a single distorted helical structure with alternating \( \alpha-3_{10}\)-\( \alpha \) character that partially occludes the active site. The ‘closed’ form is distinguished from the ‘semi-closed’ form by the ordering of the final nine C-terminal residues through the formation of hydrogen bonds to both UDP and H-antigen analogs. The ‘semi-closed’ forms for various mutants generally show significantly more disorder than the ‘open’ forms, while the ‘closed’ forms display little or no disorder depending strongly on the identity of residue 176. Finally, the use of synthetic analogs reveals how H-antigen acceptor binding can be critical in stabilizing the ‘closed’ conformation. These structures demonstrate a delicately-balanced substrate recognition mechanism and give insight on critical aspects of donor and acceptor specificity, on the order of substrate binding, and on the requirements for catalysis.

* Abbreviations: GTA, human ABO(H) blood group A \( \alpha-(1\rightarrow3)-N\)-acetylgalactosaminyltransferase; GTB, human ABO(H) blood group B \( \alpha-(1\rightarrow3)\)-galactosaminyltransferase; HA, H-antigen disaccharide; deoxy-acceptor, DA, 3-deoxy-Gal-H-antigen disaccharide; amino-deoxy-acceptor, ADA, 2-deoxy-Fuc-3-amino-Gal-H-antigen disaccharide.
INTRODUCTION

Glycosyltransferases synthesize carbohydrate moieties of glycoconjugates by catalyzing the sequential addition of monosaccharides from specific donors to specific acceptors. The ubiquitous presence of glycolipids and glycoproteins in all living systems underlines the importance of the glycosyltransferase super family, and the DNA of all domains of life encode for a large number of these enzymes. To date, crystal structures of glycosyltransferases have displayed a high degree of structural similarity even when there is low sequence homology. As such, glycosyltransferases provide an excellent example of the preferential conservation of structural phenotype over the conservation of sequence identity, which indicates that the mechanism of glycosylation, although not yet fully understood, has been conserved. Elucidation of the details of substrate recognition would allow the development of new inhibitors for the treatment of microbial diseases, genetic ailments such as diabetes, and cancer. The generation of inhibitors of the blood group A and B synthesizing glycosyltransferases GTA and GTB have been reported, including an inhibitor-bound structure.

Most glycosyltransferases are observed to lie in one of two major fold families, GT-A and GT-B (not to be confused with the GTA and GTB enzymes discussed here). Structural studies have revealed that specific internal sections of polypeptide adjacent to the active site are often observed to be flexible or completely disordered. These internal loops have been suggested to restrict water access to the active site, as well as act in donor recognition and catalysis, including the inverting enzymes β4Gal-T1, GnT-I, GlcAT-I and GlcAT-P, the retaining enzymes EXTL2, α-(1→3)-GalT, GTA and GTB; the microbial inverting SpsA, CstII; and the retaining microbial enzyme LgtC.

The retaining α-(1→3)-galactosyltransferase α-(1→3)-GalT is the enzyme most homologous to GTA/GB in sequence and structure, and has been reported to display substrate induced conformational changes. This enzyme transfers Gal from UDP-Gal to oligosaccharides terminating in lactose or LacNAc [β-D-Gal-(1→4)-β-D-GlcNAc]. Like GTA and GTB, α-(1→3)-GalT is a retaining enzyme with a GT-A fold. Unlike GTA and GTB, the structure of α-(1→3)-GalT displays a completely ordered internal loop in the unliganded state, which has been reported to lie in different conformations for different mutants and in substrate bound and unbound complexes.
GTA and GTB are responsible for the generation of the human ABO(H) blood group A and B antigens\textsuperscript{25,26}. GTA catalyzes the transfer of GalNAc from UDP-GalNAc to the H antigen acceptor ($\alpha$-L-Fuc-(1$\rightarrow$2)-\beta-D-Gal-OR, where R is glycolipid or glycoprotein) to form the A antigen, while GTB catalyzes the transfer of Gal from UDP-Gal to the H antigen acceptor to form the B antigen\textsuperscript{27,28}. Initial high-resolution structural studies of both GTA and GTB revealed two regions of disordered polypeptide\textsuperscript{19}. One region was comprised of the last 10 residues of the C-terminus, while the other was an internal polypeptide loop composed of residues 177–195. Subsequent studies have shown that part of the disorder of the internal loop was due to the presence of a heavy atom, and that crystals of the mutant enzyme GTB/C209A grown in the absence of heavy atoms display a smaller disordered segment of the internal loop consisting of residues 177 to 187\textsuperscript{20}.

GTA and GTB are the two most homologous glycosyltransferases known that utilize different nucleotide donors, and differ by only four of 354 amino acids: Arg/Gly\textsuperscript{176}, Gly/Ser\textsuperscript{235}, Leu/Met\textsuperscript{266} and Gly/Ala\textsuperscript{268} in GTA and GTB respectively\textsuperscript{29}. The role of each critical residue in donor and acceptor recognition has been studied through the generation of chimeric GTA/GTB enzymes. A nomenclature based on these four critical amino acid residues has been developed to describe GTA and GTB chimera, where GTA can be referred to as AAAA and GTB as BBBB with each letter corresponding to one critical residue in increasing order, such that the ABBB chimera would correspond to the GTB/G176R mutant enzyme, and AABB would correspond to the GTB/G176R/S235G mutant enzyme. Critical residues Leu/Met\textsuperscript{266} and Gly/Ala\textsuperscript{268} have been shown to be responsible for discrimination between the two donor molecules\textsuperscript{30-32}, while Gly/Ser\textsuperscript{235} and Leu/Met\textsuperscript{266} significantly impact acceptor recognition\textsuperscript{33}; however, the function of the conserved mutation Arg/Gly\textsuperscript{176} has been elusive. Structural studies in the past have been hampered by the fact that Arg/Gly\textsuperscript{176} lies at the edge of the internal disordered loop from residues 176–195; however, the development of crystallization conditions for BBBB (GTB), ABBB and AABB in the absence of heavy atoms permits a structural investigation of the influence of residue 176 on loop ordering and substrate binding.

We now report the kinetic characterization of several chimeric enzymes along with high-resolution structures of GTB (BBBB) and the chimeric
enzyme ABBB in their unliganded states, BBBB and AABB in the presence of UDP, BBBB and AABB in the presence of synthetic H-antigen disaccharide α-L-Fucp-(1→2)-β-D-Galp-O(CH₂)₇CH₃, BBBB and ABBB in the presence of UDP and the H-antigen acceptor analog α-L-2-deoxy-Fucp-(1→2)-β-D-3-amino-Galp-O(CH₂)₇CH₃, BBBB and AABB and in the presence of both UDP and H-antigen disaccharide, and AABB in complex with UDP-Gal and the H-antigen acceptor analog α-L-Fucp-(1→2)-β-D-3-deoxy-Galp-O(CH₂)₇CH₃.
Experimental

Construction of the Synthetic Glycosyltransferase Chimeric Genes ABBB and AABB - The synthetic wild-type GTA (designated AAAA, aa 53–354) gene was constructed from synthetic oligonucleotides as described previously. The synthetic gene was designed with unique restriction sites to facilitate mutagenesis. Glycosyltransferase chimeric mutants ABBB and AABB were synthesized by digesting the AAAA gene with KpnI/SphI and ligating in the appropriate oligonucleotides to form the desired gene sequence.

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

All insert and plasmid purifications were made by Qiagen Plasmid Purification System (Qiagen Inc., Chatsworth, CA, USA). All ligations were made by the use of BRL T4 DNA ligase at room temperature for 1 hour. All restriction enzymes were purchased from New England Biolabs.

For GTB R188S, R188K and R188H (aa 63–354) site-directed mutagenesis was carried out using a Quik-Change kit (Strategene). The primers used for mutagenesis were as follows: GTB R188S (sense) 5’-ATG CGT TCC ATG GAA ATG ATC AGC GAC TTC TGC-3’ antisense 5’-CAT TTC CAT GGA ACG CAT GGA AAC GTC CTG CC-3’.

Cloning of R188K (sense) 5’-TGG CAG GAC GTT TCC TGC GTA AAA TGG AAA TGA TCA GCG AC-3’ and antisense 5’-GTC GCT GAT CAT TTC CAT TTT ACG CAT GGA AAC GTC CTG CC-3’ Cloning of R188H (sense) 5’-CAG GAC GTT TCC ATG CGT CAT ATG GAA ATG ATC AGC-3’ and antisense 5’-GCT GAT CAT TTC CAT ATG ACG CAT GGA AAC GTC CTG-3’. The altered nucleotides are shown in bold.

Protein Purification - Mutant enzymes were
purified from *E. coli* by previously described methods\textsuperscript{36} with the exception of R188H and R188K where cells were disrupted at 1.35 kbar with a constant system cell disrupter. Expression levels for mutants were good and the yields of final purified proteins were ABBB 36 mg/L, AABB 50 mg/L, R188S 8 mg/L, R188K 66 mg/L and R188H 15 mg/L.

**Kinetic Characterization** - Kinetics using α-L-Fucp-(1→2)-β-D-Galp-O-R as an acceptor were carried out with a radiochemical assay, where a Sep-Pak reverse-phase cartridge is used to isolate radiolabeled reaction products created when the label is transferred from a radioactive donor to the hydrophobic acceptor\textsuperscript{37}. Assays were performed at 37 °C in a total volume of 12 µl containing substrates and enzyme in 50 mM MOPS buffer pH 7.0, 20 mM MnCl\textsubscript{2} and BSA (1 mg/ml). Seven different concentrations of donor and acceptor were employed, and initial rate conditions were linear with no more than 10 percent of the substrate consumed in the reaction. For the donors, the K\textsubscript{m} values were determined at 1.0 mM acceptor and the K\textsubscript{m} for the acceptor was determined at 1.0 mM donor. The kinetic parameters k\textsubscript{cat} and K\textsubscript{m} were obtained by non-linear regression analysis of the Michaelis Menten equation with the GraphPad PRISM 3.0 program (GraphPad Software, San Diego, CA). Two-substrate kinetic analysis was performed for the AABB and ABBB mutants to obtain K\textsubscript{A} (acceptor K\textsubscript{m}), K\textsubscript{B} (donor K\textsubscript{m}), K\textsubscript{ib} and K\textsubscript{ia} as previously described\textsuperscript{34}. K\textsubscript{ib} is the apparent Michaelis constant for donor that is independent of the concentration of acceptor and thus corresponds to the dissociation constant of the enzyme.UDP-Gal or enzyme.UDP-GalNAc complexes. K\textsubscript{ia} is the dissociation constant of the enzyme.acceptor complex.

**Crystallization** - All proteins were crystallized using conditions different from those reported previously\textsuperscript{10,19,33,35,38,39}. Whereas the first crystals of GTB were grown from relatively low protein concentrations (approximately 8–15 mg/ml) and as a mercury derivative, the crystals in this paper were initially generated from higher protein concentrations (approximately 60–75 mg/ml). The first crystals of the ABBB and AABB mutants grew in stock solutions containing 20 mM 3-(N-morpholino)propanesulfonic buffer (MOPS) PH 7.0, 75 mM NaCl, 15 mM BME, 0.05% NaN\textsubscript{3} and stored at 4°C for several months. Crystals of ABBB and AABB were washed with mother liquor consisting of 7% PEG, 70 mM N-[2-acetamido]-2-iminodiacetic acid buffer, pH 7.5, 30 mM sodium acetate buffer, pH 4.6,
40 mM ammonium sulfate and 5 mM MnCl₂. Crystals of BBBB were obtained by the hanging drop method from 30-40 mg/mL fresh protein solutions containing 1% polyethylene glycol (PEG), 4.5% methylpentanediol (MPD), 0.1 M ammonium sulfate, 0.07 M NaCl, 0.05 M N-[2-acetamido]-2-iminodiacetic buffer, pH 7.5, 5 mM MnCl₂ against a reservoir containing 2.7% PEG, 7% MPD, 0.32 M ammonium sulfate, 0.25 M NaCl and 0.2 M N-[2-acetamido]-2-iminodiaceic buffer, pH 7.5. Crystals of BBBB, ABBB, and AABB in complex with UDP, synthetic H-antigen disaccharide α-L-Fuc-(1→2)-β-D-Gal-O(CH₂)₃CH₃ (HA)⁸, or the analogs α-L-2-deoxy-Fuc-(1→2)-β-D-3-amino-Galp-O(CH₂)₃CH₃ (ADA)⁴⁰ and α-L-Fucp-(1→2)-β-D-3-deoxy-Galp-O(CH₂)₃CH₃ (DA)⁴¹ in various combinations were obtained by soaking substrate into the unliganded crystals. Crystals were washed with mother liquor consisting of 7% PEG, 70 mM N-[2-acetamido]-2-iminodiaceic acid buffer, pH 7.5, 30 mM sodium acetate buffer, pH 4.6, 40 mM ammonium sulfate and 5 mM MnCl₂. The concentration of UDP was usually 25 mM, but as little as 10 mM was often sufficient and 50 mM was used for BBBB+UDP. The H antigen acceptor analogs HA, DA and ADA concentrations ranged from 10 to 20 mM. The concentration of UDP-Gal ranged from 35 to 50 mM. The concentration of MnCl₂ was 5 mM. All substrates were added incrementally over a period of a few minutes to a few hours so as to prevent crystal fracture. In the case of AABB+UDP-Gal+DA, additional UDP-Gal was added to the crystal minutes before freezing in order to minimize the extent of UDP-Gal hydrolysis. No UDP was added to AABB+UDP, as the UDP appeared to follow the protein through the purification process. The UDP was removed to generate the AABB+H structure by washing the crystal with 10 mM EDTA to remove the manganese that bound the UDP to the protein.

Data collection & Reduction - X-ray diffraction data was collected at -160 °C for all crystals using a CryoStream 700 crystal cooler. Each crystal was incubated with a cryoprotectant solution that consisted of mother liquor with 30% (v/v) glycerol replacing a corresponding volume of water, except AABB+UDP-Gal+DA where a corresponding volume of MPD was used. Data were collected on a Rigaku R-AXIS IV++ area detector at distances of 72 mm and exposure times between 4.0 and 7.0 min for 0.5° oscillations. X-rays were produced by an MM-002 generator (Rigaku/MSC, College Station, TX) coupled to Osmic...
‘Blue’ confocal x-ray mirrors with power levels of 30 watts (Osmic, Auburn Hills, MI). The data were scaled, averaged and integrated using d*trek and CrystalView\textsuperscript{42}.

\textit{Structure determination} – Although the structures were nearly isomorphous, for completeness all structures were solved by molecular replacement using the CCP4 module MOLREP\textsuperscript{33,44} with the structure of wild type GTB as a starting model (PDB accession code 1LZ7), and subsequently refined using the CCP4 module REFMAC\textsuperscript{55}. All figures were produced using Setor\textsuperscript{46} and SetoRibbon (unpublished).
RESULTS

The details of data collection and refinement for the enzyme complexes are provided in Table 1a for BBBB structures and Table 1b for ABBB and AABB structures. In the 12 structures determined, BBBB, ABBB, and AABB crystals were soaked with combinations of UDP, UDP-Gal, H-antigen disaccharide (HA), 3-deoxy-Gal-H-antigen disaccharide (deoxy-acceptor, DA), and 2-deoxy-Fuc-3-amino-Gal-H-antigen disaccharide (amino-deoxy-acceptor, ADA). These 12 structures are labeled as BBBB, BBBB+UDP, BBBB+HA, BBBB+UDP+HA, BBBB+UDP+ADA, ABBB, ABBB+UDP, ABBB+UDP+HA, ABBB+UDP+ADA, AABB+UDP, AABB+HA, AABB+UDP-Gal+DA. The maximum resolution of the diffraction data collected varied from 1.75 Å to 1.41 Å with a final $R_{work}$ ranging from 19.4% to 22.3% and an $R_{free}$ ranging from 21.7% to 24.6%.

The primary distinguishing characteristic among the structures of the liganded and unliganded forms of BBBB, ABBB and AABB can be found in the two regions of polypeptide observed to be completely disordered in the original structures of GTA and GTB$^{19}$. In general, the internal loop of the BBBB, ABBB and AABB structures show fewer disordered residues than the corresponding region in the heavy atom structures$^{19}$. A summary of the observed electron density surrounding the internal loop (residues 176–195) and the C-terminus (residues 346 to 354) for all structures is given in Table 2. Without exception, structures containing Arg176 show significantly more order than the corresponding structure containing Gly176.

All ABBB and AABB structures display large portions of the internal loop, which is observed to adopt an ‘open’ conformation when unliganded, a ‘semi-closed’ conformation when bound to UDP and a ‘closed’ conformation when bound to UDP or UDP-Gal and acceptor, Fig 1a,b.

For all structures, the internal loop itself can be divided into two portions. The first consists of residues 175–188 that shows significant flexibility and contains an $\alpha$-helix consisting of residues 180–187. The second consists of residues 189–195 that adopts an $\alpha$-helical conformation similar to that observed in a mutant GTB structure$^{20}$. The nine C-terminal residues remain disordered in the ‘open’ or ‘semi-closed’ states.
display various levels of order in the ‘closed’ conformation depending on the presence of substrate and on the identity of Arg-176. Structures soaked with UDP sometimes show partial occupancy, while all structures soaked with H antigen analogs display a fully occupied acceptor binding site.

**Disorder in BBBB, ABBB and AABB** - The identity of residue 176 (arginine in AXXX enzymes and glycine in BXXX enzymes) is not only strongly correlated with the level of order observed in the internal polypeptide loop of which it is a part, but with that of the C-terminal residues as well. Given that the ABBB and AABB mutants display remarkably higher levels of order and detail than the wild type BBBB enzyme, the structures of the mutants will be discussed first and then compared with the wild type.

**ABBB and AABB structures** – A comparison of electron density observed in the different complexes of ABBB reveals that the level of order in the internal and C-terminal loops changes significantly with different substrate groupings, Table 2. The unliganded ABBB structure displays excellent electron density along almost the entire length of the polypeptide (including almost the entire internal loop), with the nine C-terminal residues completely disordered. The ABBB+UDP and AABB+UDP structures show electron density corresponding to a partially-occupied UDP molecule and a level of order comparable to the unliganded structure; however, there is clear evidence for two alternate conformations of approximately equal weight for residues 176–188, Fig. 1c,d. In one conformation the loop follows the path observed in the unliganded structure, while in the second conformation these same residues move toward the UDP molecule to partially occlude the active site to form the ‘semi-closed’ state. The AABB+HA structure displays a similar degree of order as the ABBB unliganded structure (having complete disorder only for Ala177 and Lys179 in the internal loop); however, there is no evidence of order in the C-terminal region. Interestingly, attempts were made to crystallize ABBB in the presence of UDP-Gal; however, the structure displays only low UDP occupancy in the active site and no conformational shift, indicating that the Gal moiety is disordered or that the UDP-Gal has hydrolyzed (structure not included).

The most striking structures in this series are ABBB+UDP+HA and AABB+UDP-Gal+DA, which both display excellent electron density for almost the entire polypeptide chain (including the C-terminal residues) and unambiguous electron
density for both the UDP and the HA. The structure of AABB+UDP-Gal+DA displays electron density corresponding to UDP-Gal and a fully occupied DA. The internal loop is not disordered over two conformations, but shows a 100% conformational change in that corresponds to the ‘semi-closed’ state in ABBB+UDP. Together, the conformational shift in the internal loop and ordering of the C-terminus result in the completely occluded active site of the ‘closed’ conformation of the enzyme, Table 2, Fig. 1a. The C-terminal residue His-348 forms unambiguous hydrogen bonds with the O2- and O3-hydroxyl groups of the α-L-Fucp moiety on the acceptor molecule, Fig. 2a. In contrast, the structure of ABBB in complex with ADA (remembering that ADA is the H-antigen acceptor analog that lacks the O2-hydroxyl group) shows complete electron density only for C-terminal residues Lys346 and Asn347, main chain density only for His348 to Arg352, and complete disorder for residues His348, Asn353 and Pro354, Table 2, Fig 2b.

ABBB+UDP+HA shows a fully-occupied glycerol molecule (cryoprotectant) in the donor binding site of the enzyme, Fig. 2c. A comparison of this structure with AABB+UDP-Gal+DA shows that the glycerol molecule is positioned to mimic the interaction of the galactosyl residue with Arg188 from the internal loop, Fig. 2c,d.

**BBBB structures** – The BBBB structures show the same two major regions of disorder. Unlike the ABBB and AABB structures, the internal polypeptide loop cannot be clearly divided into a flexible and an ordered region for many of the structures, as all the residues from 176 to 195 can display disorder, Table 2. Notably, only the unliganded BBBB and fully liganded BBBB+UDP+HA structure show significant order, and the binding of a single ligand, either UDP or the H-antigen, results in appreciably more disorder. There are only 4 residues (Ala177 to Arg180) disordered in the internal loop of BBBB, while BBBB+UDP and BBBB+HA show disorder in 9 and 11 residues, respectively (Ala177 to Ser-185 in BBBB+UDP and Gly176 to Met186 in BBBB+HA). The region 189 to 195 observed to be an α-helix in ABBB and AABB, as well as the GTB/C209A mutant20, also displays somewhat more disorder upon the binding of either substrate alone. Binding of both substrates in the BBBB+UDP+HA complex causes most of the polypeptide main chain of the internal and C-terminal loops to become ordered and so form the ‘closed’ state. The main chain carbonyls of Val351 and Arg352 interact with the O4-carbonyl of the uracil moiety of the UDP through a bridging water molecule,
and the side chain of Arg352 forms salt bridges with both phosphates moieties of UDP; however, electron density corresponding to the last three C-terminal residues is absent. There is no evidence of glycerol in the donor binding site of BBBB+UDP+HA.

As found in the ABBB structures, increased disorder in the C-terminal residues of the corresponding BBBB structures is observed when the ADA is substituted for the HA or DA acceptor analogs. While BBBB+UDP+HA shows the ‘closed’ form of the enzyme with electron density corresponding to Lys346-Val351, the BBBB+UDP+ADA structure displays significantly more disorder in the C-terminal region with electron density corresponding only to residues Lys346 and Asn347, Fig 2e,f.

It is important to note the large degree of movement of polypeptide that was permitted in the crystalline state of BBBB, ABBB and AABB, and that great care had to be taken to add substrate slowly in order to prevent crystal cracking.

**Kinetic Parameters** – Kinetic constants for wild type GTA, GTB and mutant enzymes are given in Table 3. The chimeric enzymes ABBB and AABB show tight binding of UDP-Gal with dissociation constants (Kib) of 1.6 and 1.1 μM, respectively. Acceptor binding is also tighter for these mutants with Kia values of 3.8 and 0.94 μM. The kcat for UDP-Gal for ABBB is comparable to that of GTB but is reduced from 5.1 to 2.2 s⁻¹ for AABB. There was a marginal increase in kcat for UDP-GalNAc from 0.41 for GTB to 0.65 and 0.60 s⁻¹ for ABBB and AABB confirming the dominance of Leu/Met266 and Gly/Ala268 in donor discrimination. The binding of the alternate donor UDP-GalNAc was also tighter for the chimeric enzymes than for GTB with dissociation constants of 9.2 and 44 μM compared to 69 μM; however, donor binding is weaker than that of GTA which has a Kib of 3 μM. The importance of the interaction between Arg188 and donor is evident from the dramatic reduction in kcat for the R188S and R188K mutants.
DISCUSSION

The ‘open’ conformation for the enzymes – In the absence of donor or acceptor, BBBB and ABBB crystallize in the ‘open’ form, where the nine C-terminal residues are disordered, and a major portion of the internal loop is disordered or lies in a conformation that leaves the donor and acceptor binding sites exposed to solvent. The effect of Arg176 on internal loop structure is clearly evident, as a substantial portion of the internal loop is disordered in BBBB while most of the loop is ordered in ABBB, where it consists of two helical segments joined at Arg187, Table 2, Fig 1d. This ‘open’ form is likely due to the mutual repulsion of many positively charged residues, such as internal loop residues Lys179, Arg180, and Arg188 as well as C-terminal residues Arg352 and Lys346, Fig. 3a.

UDP binding induces a ‘semi-closed’ conformation – The ABBB+UDP and AABB+UDP structures reveal a fascinating transition between the ‘open’ and ‘semi-closed’ states as both structures display clear evidence that residues 176–188 are disordered over both conformations, Fig. 1c. The ‘semi-closed’ state has the helix formed by residues 176–188 moving as much as 6 Å toward the UDP molecule to partially occlude the active site without forming any new hydrogen bonds to the UDP moiety, Fig1a,b. The change to the ‘semi-closed’ form results in the first helix of the internal loop moving into alignment with the second, Fig. 1d. In this shift a new main chain hydrogen bond forms between Arg187-N and Asp-183-O and a main chain hydrogen bond between Glu190-N and Met186-O transfers to between Met189-N and Met186-O, such that two helices become linked by a single turn of a 310 helix, Fig1d. The result is a distorted helical structure with mixed α-310-α character that partially occludes the active site. The mutual repulsion of positively-charged residues Lys179, Arg180 and Arg188 that held the enzyme in the ‘open’ state are likely overcome to form the ‘semi-closed’ conformation through electrostatic interactions with the negatively charged pyrophosphate moiety of bound UDP20. Interestingly, despite high concentrations of UDP, both the ABBB+UDP and AABB+UDP structures show electron density corresponding to approximately 50% occupancy which correlates to the occupancy in the two observed conformations of the internal loop. In contrast, BBBB+UDP does not show clear evidence of a split between its ‘open’ and ‘semi-closed’ states. Indeed, there is significantly higher thermal motion in the semi-closed form seen in BBBB+UDP.
compared to the unliganded form. Although only a few residues in the internal loop of BBBB+UDP can be seen in the electron density maps it is clear that these at least have moved to positions that correspond to the ‘semi-closed’ conformation; however, the remainder of the internal loop in BBBB displays a great number or even a continuum of conformations between the two states.

*ABBB+UDP+HA and BBBB+UDP+HA display a ‘closed’ conformation* – The fully liganded ABBB enzyme shows ordering of almost all previously disordered residues in both the C-terminus and internal loop. This ‘closed’ conformation in ABBB and BBBB is achieved only in the presence of both UDP and acceptor. Those residues of the internal loop that are ordered are in the same conformation as observed in the ‘semi-closed’ state of ABBB+UDP, and the mutual repulsion observed among the positively-charged residues in the internal loop has been fully overcome by the combination of UDP binding and the interaction of the UDP with the newly-ordered C-terminus, Fig 3b.

Significantly, six of the nine C-terminal residues of the protein form a short α-helix (residues 347–352) that makes contact with residues in the active site, with UDP, with the α-L-Fucp moiety of the acceptor, and completes the sequestration of the substrates from solvent. The side chain of Lys346 extends into the active site to from a salt bridge with the β-phosphate of UDP and the side chain of the third residue (Asp213) of the DXD motif.

Although relative levels of disorder and thermal motion clearly show that the internal loop is stabilized by the ordering of the C-terminal loop, there are no direct hydrogen bonds between these two flexible regions. Instead this stabilization occurs through a number of bridging interactions moderated by UDP moiety and three water molecules. The only direct contact between the internal loop and the C-terminus occurs through a stacking interaction between Trp-181 and Arg352, Fig 3c.

*Effect of cryoprotectant* – A fully-occupied glycerol molecule is seen in the acceptor binding site of each structure in the absence of acceptor; however, given that this molecule does not contact either mobile polypeptide loop, and given that it is displaced by even modest concentrations of acceptor, it is unlikely to influence the conformation of these loops.

ABBB+UDP+HA is the only structure to display a glycerol molecule in the donor binding site, where it may contribute to the observed formation of the ‘closed’ state mimicking (through a bridging water molecule) the interaction of the
galactosyl moiety in AABB+UDP-Gal+DA with Arg188 of the internal mobile loop, Fig. 2c,d. Both BBBB+UDP+HA and ABBB+UDP+ADA form the ‘closed’ conformation without any indication of glycerol in the donor binding site.

**AABB and binding of UDP-Gal** - Crystals of AABB soaked with UDP-Gal and DA revealed a highly-occupied donor and acceptor in the active site cleft with the enzyme in a ‘closed’ conformation. This represents the fully-liganded state required for turnover of the enzyme and, in combination with other structures that bind the active acceptor disaccharide analog HA, a complete schematic of substrate recognition can be drawn, Fig 4. The donor sugar is bound in the classic ‘folded back’ conformation observed for other glycosyltransferases, Fig. 3d. The shift to the ‘closed’ conformation does bring Ser185 and Arg188 into the donor-sugar binding site, but the donor displays a somewhat different hydrogen bond pattern for the α-Gal moiety than predicted47.

While the hydrogen bonds between Asp211 and the O3-hydroxyl group and between Asp302 and the O4-hydroxyl group are observed, the predicted interaction between Ser185 and the O6-hydroxyl group is not observed. Instead, hydrogen bonds are found between Arg188 and the O3-hydroxyl group, and between His-301 and the O6-hydroxyl group, Fig 3e. Although it does not participate in the active recognition of the donor sugar galactosyl residue, Ser185 is positioned to provide a steric barrier to the binding of UDP-Glc and accounts for this aspect of donor specificity. The ability of Ser185 to exclude UDP-Glc had been predicted, leading to speculation that an appropriate mutation at position 185 could allow the GTB to transfer glucose to the H-antigen48. Further, the observed position of the fully-occupied nucleotide donor confirms the mechanism by which Met266 and Ala268 in BBBB distinguish between acetamido and hydroxyl groups present on the UDP-GalNAc and UDP-Gal donors, respectively31, Fig 3f.

Recent NMR studies reported a similar conformation for the UDP-Gal bound to GTB49. However, the observed contact by Lys346 to the β-phosphate is missing in the NMR structure, and the position of the pyrophosphate group differs significantly, which can be attributed to the absence of donor 1H-NMR signals in this region. It is known that the O3 and O4 hydroxyl groups of UDP-Gal are particularly important in donor substrate recognition50.

The observation that UDP is carried through the purification process by ABBB and AABB, as well as the observation of
intact bound donor only for AABB correlates with the observed Kib (enzyme.donor dissociation constant) values for these enzymes, Table 3. ABBB has among the lowest observed Kib value (1.6 μM) for any GTA/GTB mutant studied while AABB has the lowest Kib (1.1 μM) for UDP-Gal observed for any mutant.

Acceptor recognition and conformational change – It has been known for some time that GTA and GTB do not efficiently transfer to acceptors that lack the terminal non-reducing α-L-Fucp moiety. When the original structure of GTB was solved in complex with the H-antigen disaccharide it was observed that the α-L-Fucp only provided a single contact to the enzyme (a hydrogen bond between the O4-hydroxyl and the side chain of amino acid residue Asp326). Structures of GTA and GTB reported in complex with seven different fragments and analogs of the H-antigen acceptors revealed many novel aspects of acceptor recognition by these homologous enzymes; however, the absolute necessity of the α-L-Fucp for catalytic activity was obscure.

In the present structures, the C-terminal residues recognize the acceptor via two hydrogen bonds to the α-L-Fucp residue with no contacts observed between the C-terminal residues and the β-Gal residue, Fig. 4. Given that the presence of both HA and UDP in the active site is required for the closed conformation in BBBB and ABBB, and that the C-terminal residues contact the H-antigen only through the α-L-Fuc moiety, and that the α-L-Fuc moiety is required for efficient catalysis, it can be concluded that closed conformation is likely required for efficient catalysis.

The stabilizing effect that O2-hydroxyl group of the α-L-Fucp residue imparts on the C-terminal region can be seen in structures of BBBB and ABBB in the presence of UDP and the ADA acceptor analog which lacks this important hydroxyl group. In general, the C-terminal residues in both ADA structures display considerably higher levels of disorder than in the analogous structures soaked with HA, having complete main chain and side chain electron density corresponding only to residues 346 in BBBB and residues 346 and 347 in ABBB, Table 2. Significantly, both structures displayed complete disorder for the side chains of His348 (involved in fucose recognition) and Arg352 (involved in UDP stabilization), Fig. 2b,f.

Effects of loop mutations on enzyme activity – Kinetic studies completed on mutants of internal loop residue 188 of GTA and GTB
can now be rationalized on the basis of the current structures, Table 3, Fig 4. For example, the BBBB/R188S and R/188K mutants demonstrate increases in Km for the donor and large decreases in kcat, which is consistent with its role in donor sugar recognition and turnover. BBBB/R188H had a specific activity 3% that of BBBB/R188K and was not further characterized.

Comparison with α-(1→3)-GalT – The most closely-related CAZy family 6 glycosyltransferase to GTA and GTB that has been structurally-characterized is bovine α-(1→3)-GalT. A mutant of α-(1→3)-GalT has recently been crystallized in the presence of the donor analog UDP-2-fluoro-Gal. With a sequence similarity of only 45%, it is not surprising that bovine α-(1→3)-GalT displays significant differences with GTB and the GTB/GTA chimera. Firstly, unlike BBBB, the corresponding internal loop in α-(1→3)-GalT has always been observed to be ordered in the wild-type enzyme. Further, whereas the generation of ‘semi-closed’ form involves a complete ordering of the internal loop in BBBB and a conformational change in the loop in ABBB, the binding of UDP in wild-type α-(1→3)-GalT or of UDP-2-fluoro-Gal in its R365K mutant results in 5 residues of the internal loop changing from random coil to an extra turn of a helix. Also unlike BBBB, ABBB and AABB, the C-terminal region of the R365K mutant structure bound to UDP-2-fluoro-Gal does not close, but is observed to curve away from the active site. Like UDP-Gal in AABB+UDP-Gal+DA, the donor sugar nucleotide UDP-2-fluoro-Gal lies in a ‘tucked-under’ conformation (Fig. 3d), and the C-terminus makes significant contact with the pyrophosphate moiety of the UDP. However unlike BBBB, ABBB and AABB, the C-terminus of α-(1→3)-GalT does not contact the acceptor18, and the presence of UDP is sufficient to fully order its C-terminus17.

The apparent ordering of the internal mobile loop in BBBB and ABBB only in the presence of both acceptor and UDP suggests a more elaborate recognition mechanism. Firstly, a significant array of positive charges from residues in the internal loop, its nascent helix dipole, and from residues in the C-terminal region are brought into proximity by the pyrophosphate moiety of bound UDP, Fig.3b. Secondly, the presence of acceptor stabilizes the C-terminus through direct hydrogen bonds to its α-L-Fucp residue, and brings it into proximity with the internal loop where there is a stacking interaction between Arg352 and Trp181, Fig 3c, and a bridging water molecule between
the main-chain amide group of Trp181 and the side chain of Asn353.

**Multiple substrate binding and the catalytic cycle of GTA/GTB** – The increased ease of formation of the closed form of the enzyme is apparent in moving from BBBB to ABBB or AABB. BBBB adopts the closed conformation in the presence of UDP and H antigen while manifesting higher levels of disorder and higher temperature factors overall. ABBB shows the ‘closed’ conformation in the presence of UDP and H antigen with lower temperature factors and AABB is able to close in the presence of UDP-Gal and DA.

Although the acceptor binding does stabilize the conformation of the C-terminal residues in BBBB, ABBB and AABB it is clear that this does not happen in the absence of UDP, and therefore the formation of the closed conformation induced by UDP binding is a requirement for the stable association of the acceptor disaccharide. NMR studies have suggested a catalytic cycle for BBBB that involves long-lived UDP-Gal binding complex and rapid on-off kinetics for acceptor binding. The structures presented here demonstrate that donor binding is a critical component in generating the closed conformation and therefore in acceptor disaccharide stabilization. The reductions in the mobility of both mobile loops upon donor binding would stabilize the UDP-Gal and explain the observed long life of the UDP-Gal-enzyme complex observed by NMR to which the more labile acceptor could subsequently bind. Together with the observation that bound acceptor would present a steric barrier to the binding of UDP-Gal, the evidence suggests that UDP-donor binding precedes the H-antigen acceptor binding to bring about the conformational changes required for catalysis.

The STD-NMR studies also suggested that GTB discriminates between the C4 epimers UDP-Gal and UDP-Glc by a ‘tweezers’ model, where the recognition of Gal-O4 by Asp302 would lead to a reactive conformation but a similar recognition of Glc-O4 would not be possible. These results contrast with the structure of AABB+UDP-Gal+DA that shows that Glc-O4 would actually be more favorably positioned to contact Asp302, but that the shift to the closed conformation would require a steric collision with internal loop residue Ser185. It is likely that the ‘open’ form of the enzyme and all intermediate conformations en route to the ‘semi-closed’ and ‘closed’ conformations are capable of binding both UDP-Glc and UDP-Gal, and that selection for the galactose over glucose occurs only in the brief period before the
fully closed conformation is achieved.

**The role of critical amino acid Arg/Gly-176** – Kinetic characterization of the chimeric enzyme ABBB (GTB/G176R) reveals a $k_{cat}$ for UDP-Gal that is not significantly different from the wild type GTB enzyme (Table 3); however, the chimeric enzyme BAAA (GTA/R176G) displays a 3-fold increase in $k_{cat}$ for UDP-GalNAc over the wild type GTA enzyme, showing that Arg/Gly176 affects in enzyme turnover in GTA but not in GTB. The increased flexibility of the internal loop and resulting higher mobility of the C-terminal residues afforded by the R176G mutation may result in faster product release and substrate exchange.

The current structures clearly show that Arg/Gly176 greatly influences the level of flexibility in both the internal loop and the C-terminal region. While the C-terminal residues remain disordered in the unliganded BBBB and ABBB structures, the internal loop in the unliganded BBBB structure is almost entirely disordered and unliganded ABBB exhibits disorder only for Lys189, Ala177 and Arg176 itself. Similar trends are seen for the corresponding UDP-bound and H-antigen-bound structures. The pattern is repeated for the BBBB and ABBB structures in complex with UDP and HA, which both show the ‘closed’ conformation; however, BBBB+UDP+HA displays much higher thermal motion than ABBB+UDP+HA in both mobile loops. It is interesting to note that the side chain of Arg176 itself does not make any contact with any other part of the enzyme in any structure, and is observed to be partially or fully disordered in all structures. This supports the hypothesis that its contribution to substrate turnover may be steric in that glycine has a greater freedom of rotation about its main chain dihedral angles.

**Conclusions** - The BBBB, ABBB and AABB enzymes have demonstrated unambiguous conformational changes upon substrate binding. Three distinct states have been observed for the enzymes: an ‘open’ conformation, a ‘closed’ conformation and a ‘semi-closed’ conformation. These crystal structures reveal a novel role for the first critical amino acid residue Arg/Gly176 in determining the flexibility of the internal mobile loop. The conformational changes also provide insight into both H-antigen and UDP-Gal recognition, demonstrate the necessity of ordered substrate binding, and suggest necessary events in the catalytic cycle.
References


Table 1a
Data collection and refinement results for crystal structures of BBBB

<table>
<thead>
<tr>
<th></th>
<th>BBBB</th>
<th>BBBB + UDP</th>
<th>BBBB + HA</th>
<th>BBBB + UDP</th>
<th>BBBB + UDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution (Å)</td>
<td>20-1.43</td>
<td>20-1.75</td>
<td>20-1.55</td>
<td>20-1.69</td>
<td>20-1.69</td>
</tr>
<tr>
<td>Space Group</td>
<td>C222₁</td>
<td>C222₁</td>
<td>C222₁</td>
<td>C222₁</td>
<td>C222₁</td>
</tr>
<tr>
<td>A (Å)</td>
<td>52.54</td>
<td>52.54</td>
<td>52.56</td>
<td>52.45</td>
<td>52.43</td>
</tr>
<tr>
<td>B (Å)</td>
<td>149.91</td>
<td>149.93</td>
<td>150.25</td>
<td>149.60</td>
<td>149.93</td>
</tr>
<tr>
<td>C (Å)</td>
<td>79.21</td>
<td>79.28</td>
<td>79.38</td>
<td>78.74</td>
<td>78.92</td>
</tr>
<tr>
<td>R&lt;sub&gt;merge&lt;/sub&gt; (%)&lt;sup&gt;a,b&lt;/sup&gt;</td>
<td>0.034(.322)</td>
<td>0.051 (.338)</td>
<td>0.037(.318)</td>
<td>0.048 (.324)</td>
<td>0.044 (.325)</td>
</tr>
<tr>
<td>Completeness (%)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>98.1 (96.9)</td>
<td>98.5 (96.2)</td>
<td>98.2 (96.0)</td>
<td>99.5 (99.9)</td>
<td>97.0 (98.4)</td>
</tr>
<tr>
<td>Unique Reflections</td>
<td>57,030</td>
<td>31,525</td>
<td>45,171</td>
<td>34,923</td>
<td>34,204</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R&lt;sub&gt;work&lt;/sub&gt; (%)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>21.2</td>
<td>19.6</td>
<td>21.1</td>
<td>19.8</td>
<td>20.0</td>
</tr>
<tr>
<td>R&lt;sub&gt;free&lt;/sub&gt; (%)&lt;sup&gt;c,d&lt;/sup&gt;</td>
<td>22.1</td>
<td>23.1</td>
<td>23.5</td>
<td>22.2</td>
<td>23.2</td>
</tr>
<tr>
<td>No. water</td>
<td>236</td>
<td>213</td>
<td>240</td>
<td>200</td>
<td>196</td>
</tr>
<tr>
<td>r.m.s. bond (Å)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.009</td>
<td>0.014</td>
<td>0.011</td>
<td>0.014</td>
<td>0.017</td>
</tr>
<tr>
<td>r.m.s. angle (°)&lt;sup&gt;e&lt;/sup&gt;</td>
<td>1.28</td>
<td>1.44</td>
<td>1.28</td>
<td>1.41</td>
<td>1.61</td>
</tr>
<tr>
<td>PDB ID</td>
<td>2RIT</td>
<td>2RIX</td>
<td>2RIY</td>
<td>2RJ8</td>
<td>2RJ9</td>
</tr>
</tbody>
</table>

<sup>a</sup> R-merge, \( \frac{\sum |I_{obs} - I_{ave}|}{\sum I_{ave}} \).

<sup>b</sup> Values in parentheses represent highest resolution shell.

<sup>c</sup> R-work, \( \frac{\sum |F_o| - |F_c|}{\sum |F_o|} \).

<sup>d</sup> 10% of reflections were omitted for R-free calculations.

<sup>e</sup> r.m.s. root-mean-square
Table 1b

Data collection and refinement results for crystal structures of ABBB and AABB

<table>
<thead>
<tr>
<th></th>
<th>ABBB</th>
<th>ABBB+UDP</th>
<th>ABBB+UDP+HA</th>
<th>ABBB+UDP+ADA</th>
<th>AABB+UDP</th>
<th>AABB+UDP+HA</th>
<th>AABB+UDP+DA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution (Å)</td>
<td>20.1-1.45</td>
<td>20-1.52</td>
<td>20-1.55</td>
<td>20-1.47</td>
<td>20-1.45</td>
<td>20-1.41</td>
<td>20-1.70</td>
</tr>
<tr>
<td>Space Group</td>
<td>C222₁</td>
<td>C222₁</td>
<td>C222₁</td>
<td>C222₁</td>
<td>C222₁</td>
<td>C222₁</td>
<td>C222₁</td>
</tr>
<tr>
<td>A (Å)</td>
<td>52.53</td>
<td>52.53</td>
<td>52.45</td>
<td>52.48</td>
<td>52.59</td>
<td>52.52</td>
<td>52.36</td>
</tr>
<tr>
<td>B (Å)</td>
<td>149.58</td>
<td>149.35</td>
<td>149.12</td>
<td>149.74</td>
<td>149.02</td>
<td>149.06</td>
<td>148.65</td>
</tr>
<tr>
<td>C (Å)</td>
<td>79.64</td>
<td>79.65</td>
<td>79.64</td>
<td>79.79</td>
<td>79.65</td>
<td>79.61</td>
<td>79.52</td>
</tr>
<tr>
<td>R_merge (%)</td>
<td>2.5 (23)</td>
<td>4.2 (36)</td>
<td>4.6 (33)</td>
<td>3.4(28.9)</td>
<td>2.8(20.7)</td>
<td>3.7(31.5)</td>
<td>5.7 (31.0)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>91.5 (86.1)</td>
<td>99.9</td>
<td>97.1</td>
<td>98.6</td>
<td>97.1</td>
<td>96.2</td>
<td>95.3</td>
</tr>
<tr>
<td>Unique Reflections</td>
<td>51,207</td>
<td>48,558</td>
<td>44,381</td>
<td>53,049</td>
<td>54,190</td>
<td>58,216</td>
<td>32,932</td>
</tr>
</tbody>
</table>

Refinement

<table>
<thead>
<tr>
<th></th>
<th>ABBB</th>
<th>ABBB+UDP</th>
<th>ABBB+UDP+HA</th>
<th>ABBB+UDP+ADA</th>
<th>AABB+UDP</th>
<th>AABB+UDP+HA</th>
<th>AABB+UDP+DA</th>
</tr>
</thead>
<tbody>
<tr>
<td>R_work (%)</td>
<td>20.9</td>
<td>22.3</td>
<td>20.5</td>
<td>20.6</td>
<td>20.3</td>
<td>21.6</td>
<td>19.4</td>
</tr>
<tr>
<td>R_free (%)</td>
<td>22.8</td>
<td>24.6</td>
<td>22.8</td>
<td>22.6</td>
<td>21.7</td>
<td>23.7</td>
<td>21.7</td>
</tr>
<tr>
<td>No. water</td>
<td>236</td>
<td>214</td>
<td>200</td>
<td>230</td>
<td>205</td>
<td>237</td>
<td>194</td>
</tr>
<tr>
<td>r.m.s. bond (Å)</td>
<td>0.009</td>
<td>0.010</td>
<td>0.011</td>
<td>0.009</td>
<td>0.009</td>
<td>0.008</td>
<td>0.014</td>
</tr>
<tr>
<td>r.m.s. angle (°)</td>
<td>1.29</td>
<td>1.32</td>
<td>1.35</td>
<td>1.28</td>
<td>1.24</td>
<td>1.23</td>
<td>1.43</td>
</tr>
</tbody>
</table>

PDB ID

<table>
<thead>
<tr>
<th></th>
<th>2RIZ</th>
<th>2RJ0</th>
<th>2RJ1</th>
<th>2RJ4</th>
<th>2RJ5</th>
<th>2RJ6</th>
<th>2RJ7</th>
</tr>
</thead>
</table>

- **a** R-merge, \( \frac{\sum |I_{obs} - I_{ave}|}{\sum I_{ave}} \).
- **b** Values in parentheses represent highest resolution shell.
- **c** R-work, \( \frac{\sum |F_o| - |F_c|}{\sum |F_o|} \).
- **d** 10% of reflections were omitted for R-free calculations.
- **e** r.m.s. root-mean-square
Table 2
Loop ordering in BBBB, ABBB and AABB

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Internal loop</th>
<th>C-terminal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>176 181 186 191 196</td>
<td>346 351</td>
</tr>
<tr>
<td><strong>Open</strong></td>
<td><strong>two α-helices - some disorder – active site open</strong></td>
<td><strong>Disordered</strong></td>
</tr>
<tr>
<td>BBBB</td>
<td>EV Gaykr WQDVS MRRME MISDF CERR</td>
<td>VP knhqa vrnq</td>
</tr>
<tr>
<td>BBBB+HA</td>
<td>EV gaykr wQdVS MRRME MISDF CERR</td>
<td>VP knhqa vrnq</td>
</tr>
<tr>
<td>ABBB</td>
<td>EV RAYKR WQDVS MRRME MISDF CERR</td>
<td>VP knhqa vrnq</td>
</tr>
<tr>
<td>ABBB+HA</td>
<td>EV RAYKR WQDVS MRRME MISDF CERR</td>
<td>VP knhqa vrnq</td>
</tr>
<tr>
<td>ABBB+UDP (1)</td>
<td>EV Raykr wQdVS MRRME MISDF CERR</td>
<td>VP knhqa vrnq</td>
</tr>
<tr>
<td>ABBB+UDP (1)</td>
<td>EV RAYKR WQDVS MRRME MISDF CERR</td>
<td>VP knhqa vrnq</td>
</tr>
<tr>
<td><strong>Semi-closed</strong></td>
<td><strong>one helical structure – more disorder – active site closing</strong></td>
<td><strong>Disordered</strong></td>
</tr>
<tr>
<td>BBBB+UDP</td>
<td>EV Gaykr wQdVS MRRME MISDF CERR</td>
<td>VP knhqa vrnq</td>
</tr>
<tr>
<td>BBBB+UDP+ADA</td>
<td>EV GAYK wQDVS MRRME MISDF cERR</td>
<td>VP KNhqa vrnq</td>
</tr>
<tr>
<td>ABBB+UDP (2)</td>
<td>EV Raykr wQDVS MRRME MISDF CERR</td>
<td>VP knhqa vrnq</td>
</tr>
<tr>
<td>ABBB+UDP (2)</td>
<td>EV RAYKR WQDVS MRRME MISDF CERR</td>
<td>VP knhqa vrnq</td>
</tr>
<tr>
<td><strong>Closed</strong></td>
<td><strong>one helical structure - little disorder – active site closed</strong></td>
<td><strong>Ordered</strong></td>
</tr>
<tr>
<td>BBBB+UDP+HA</td>
<td>EV GAYK wQDVS MRRME MISDF CERR</td>
<td>VP KNHQA Vrnq</td>
</tr>
<tr>
<td>ABBB+UDP+HA</td>
<td>EV RAYKR WQDVS MRRME MISDF CERR</td>
<td>VP KNHQA Vrnq</td>
</tr>
<tr>
<td>ABBB+UDP+ADA</td>
<td>EV RAYKR WQDVS MRRME MISDF CERR</td>
<td>VP KNHQA Vrnq</td>
</tr>
<tr>
<td>ABBB+UDP-Gal+DA</td>
<td>EV RAYKR WQDVS MRRME MISDF CERR</td>
<td>VP KNHQA Vrnq</td>
</tr>
</tbody>
</table>

Black one-letter amino acid codes correspond to unambiguous electron density for main chain and side chain atoms, green correspond to unambiguous electron density for main chain atoms only; red letters correspond to weak or ambiguous electron density for main chain and side chain atoms. The internal loops of ABBB+UDP and ABBB+UDP are disordered over two conformations corresponding to the open (1) and closed (2) forms. Residues involved in helices are underlined. Residues observed to move in going from the ‘open’, to ‘semi-closed’ or ‘closed’ form are outlined in yellow. Residues with one-letter amino acid codes in lower case have not been included in the refined models. Substrate moieties that exhibit partial occupancy are shown in green.
Table 3

Kinetic constants for BBBB, ABBB, AABB and AAAA
as well as internal mutants of BBBB

<table>
<thead>
<tr>
<th>enzyme</th>
<th>UDP-GalNAc</th>
<th></th>
<th>UDP-Gal</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>KA (μM)</td>
<td>KB (μM)</td>
<td>Kib (μM)</td>
<td>k_{cat} (s⁻¹)</td>
</tr>
<tr>
<td>BBBB&lt;sup&gt;a&lt;/sup&gt;</td>
<td>180</td>
<td>138</td>
<td>69</td>
<td>0.42</td>
</tr>
<tr>
<td>ABBB</td>
<td>75</td>
<td>44</td>
<td>9.2</td>
<td>0.65</td>
</tr>
<tr>
<td>AABB</td>
<td>13</td>
<td>37</td>
<td>44</td>
<td>0.60</td>
</tr>
<tr>
<td>BAAA&lt;sup&gt;b&lt;/sup&gt;</td>
<td>48</td>
<td>35</td>
<td>ND</td>
<td>48</td>
</tr>
<tr>
<td>AAAA&lt;sup&gt;a&lt;/sup&gt;</td>
<td>9.9</td>
<td>8.7</td>
<td>3</td>
<td>17.5</td>
</tr>
<tr>
<td>BBBB R188S</td>
<td>(too slow)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BBBB R188K</td>
<td>40</td>
<td>86</td>
<td>ND</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

<sup>a</sup> from (38). <sup>b</sup> from (35). ND = Not determined.
Legends for Figures:

**Fig. 1:** Conformational changes associated with substrate binding. (a) Superimposition of unliganded ABBB in the ‘open’ form (white) with AABBB+UDP-Gal+DA in the ‘closed’ form (yellow/red) showing the internal and C-terminal loops (red), UDP-Gal and DA (orange) and Mn$^{2+}$ (blue), and the location of R176. (b) Expanded view about the active site with an arrow indicating the movement of the internal loop toward the donor in going from the ‘open’ state to both the ‘semi-closed’ or ‘closed’ states, and showing the ordering of the C-terminal residues to form the ‘closed’ state. (c) Stereoview of electron density corresponding to the internal loop in ABBB+UDP showing two distinct conformations of the enzyme (at 50% occupancy) corresponding to the ‘open’ (yellow) and ‘semi-closed’ (green) forms of the enzyme. The disorder converges at Met189 (gray). (d) The transformation of the internal loop (residues 176-195) from the ‘open’ (left) to the ‘semi-closed’ (right) conformation is accomplished by the merger of two alpha helices (Arg180-Met186 and Arg187-Asp194) into a distorted helical structure with alternating $\alpha$-310-$\alpha$ character. The pivot point is indicated by a star.

**Fig. 2:** Effects of different substrate analogs on ABBB, AABB and BBBB. The presence and nature of the acceptor and donor substrate analogs has a significant effect on the level of ordering of polypeptide chain. (a) Unambiguous electron density is visible for all nine C-terminal residues of ABBB in complex with UDP and H-antigen disaccharide, while (b) the substitution of ADA for the H-antigen results in significant disorder. (c) Electron density about the UDP molecule in ABBB+UDP+H showing a fully occupied glycerol molecule bridging through a water molecule to Arg188 in the internal loop to generate the ‘closed’ form. (d) Electron density about UDP-Gal in AABB+UDP-Gal+DA showing Gal-O3 forming a hydrogen bond directly to Arg-188 in the internal loop to generate the ‘closed’ form. (e) Electron density for the C-terminal loop of BBBB+UDP+HA shows significantly more order than seen in (f) BBBB+UDP+ADA. All electron density diagrams are 2Fo-Fc maps contoured at 1σ.

**Fig. 3:** Substrate binding and the ‘closed’ conformation. The two helices observed in the internal loop of ‘open’ form (a) of ABBB have several positively-charged side chains that are brought into proximity with other positively-charged side chains on the C-terminus and the nascent helix dipole to form the ‘closed’ conformation (b) upon substrate binding in ABBB+UDP+HA. (c) The only direct contact between the internal and C-terminal loops in the ‘closed’ conformation is a stacking interaction between Trp181 and Arg352. (d)
Superimposition of UDP-Gal bound the active site of AABB (green), GT7 (red) and GT43 (blue). (e) Specific hydrogen bonds involved in the recognition of the Gal moiety of UDP-Gal in AABB+UDP-Gal+DA. (f) Model of UDP-GalNAc shown in the active site of AABB in the same manner as UDP-Gal demonstrates that UDP-GalNAc could not bind as it would require approach within the van der Waals contact radius of critical residues Met266 and Ala268. Hydrogen atoms (green) are included for clarity.

**Fig. 4:** Schematic representation of donor and acceptor recognition in GTB. The chimeric enzyme AABB displays the ‘closed’ form when bound to UDP-Gal and DA, which allows for a complete characterization of substrate recognition. The acceptor Gal-O3 (*) is modeled, and does not appear in the 3-deoxy acceptor DA.
Fig. 1.
Fig. 2.
Fig. 3.
Figure 3 (cont):
ABO(H) blood group A and B glycosyltransferases recognize substrate via specific conformational changes
Javier A. Alfaro, Ruixiang Blake Zheng, Mattias Persson, James A. Letts, Robert Polakowski, Yu Bai, Svetlana N. Borisova, Nina O. L. Seto, Todd L Lowary, Monica M. Palcic and Stephen V. Evans

J. Biol. Chem. published online January 11, 2008

Access the most updated version of this article at doi: 10.1074/jbc.M708669200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts