THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 283, NO. 15, pp. 10097–10108, April 11, 2008
© 2008 by The American Society for Biochemistry and Molecular Biology, Inc. Printed in the U.S.A.

ABO(H) Blood Group A and B Glycosyltransferases Recognize Substrate via Specific Conformational Changes*

Received for publication, October 19, 2007, and in revised form, January 7, 2008 Published, JBC Papers in Press, January 11, 2008, DOI 10.1074/jbc.M708669200

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[‡1]

From the [‡]Department of Biochemistry and Microbiology, University of Victoria, Victoria, British Columbia V8W 3P6, Canada, [§]Alberta Ingenuity Centre for Carbohydrate Science and Department of Chemistry, University of Alberta, Edmonton, Alberta T6G 2G2, Canada, [¶]Carlsberg Laboratory, Gamle Carlsberg Vej 10, DK-2500 Valby, Copenhagen, Denmark, and [‖]Institute for Biological Sciences, National Research Council of Canada, Ottawa, Ontario K1A 0R6, Canada

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 α-(1→3)-*N*-acetylgalactosaminyltransferase (GTA) and an α-(1→3)-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 α-helices spanning Arg^{180}–Met^{186} and Arg^{188}–Asp^{194}, 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 α-3_{10}-α 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, whereas 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.

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 glycosyltransferases superfamily, and the DNA of all domains of life encode for a large number of these enzymes (1). To date, crystal structures of glycosyltransferases have displayed a high degree of structural similarity even when there is low sequence homology (2–4). As such, glycosyltransferases provide an excellent example of the preferential conservation of structural phenotype over the conservation of sequence identity (2), 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 (5), genetic ailments such as diabetes (6), and cancer (7). The generation of inhibitors of the blood group A and B synthesizing glycosyltransferases GTA[2] and GTB have been reported (8, 9), including an inhibitor-bound structure (10).

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) (2, 3, 11). 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 (3), including the inverting enzymes β4Gal-T1 (12), GnT-I (13), GlcAT-I (14), and GlcAT-P (15); the retaining enzymes EXTL2 (16) α-(1→3)-GalT (17, 18), GTA, and GTB (19, 20); the microbial inverting SpsA (21) and CstII (22); and the retaining microbial enzyme LgtC (23).

The retaining α-(1→3)-galactosyltransferase (α-(1→3)-GalT) is the enzyme most homologous to GTA/GTB in sequence and structure, and it has been reported to display substrate-induced conformational changes (18). This enzyme

* 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.

The atomic coordinates and structure factors (code 2RIT, 2RIX, 2RIY, 2RJ8, 2RJ9, 2RIZ, 2RJ0, 2RJ1, 2RJ4, 2RJ5, 2RJ6, 2RJ7) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] Recipient of a senior scholarship from the Michael Smith Foundation for Health Research. To whom correspondence should be addressed. E-mail: svevans@uvic.ca.

[2] The abbreviations used are: GTA, human ABO(H) blood group A α-(1→3)-*N*-acetylgalactosaminyltransferase; GTB, human ABO(H) blood group B α-(1→3)-galactosaminyltransferase; HA, H antigen disaccharide; DA, deoxy-acceptor, 3-deoxy-Gal-H antigen disaccharide; ADA, amino-deoxy-acceptor, 2-deoxy-Fuc-3-amino-Gal-H antigen disaccharide; MOPS, 3-(*N*-morpholino)propanesulfonic acid; PEG, polyethylene glycol; Ada, *N*-(2-acetamido)-2-iminodiacetic acid; MPD, methyl-pentanediol.

Downloaded from http://www.jbc.org/ by guest on July 12, 2019

transfers Gal from UDP-Gal to oligosaccharides terminating in lactose or LacNAc (β-D-Gal-(1→4)-β-D-GlcNAc) (24). Like GTA and GTB, α-(1→3)-GalT is a retaining enzyme with a GTA 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 (17, 18).

GTA and GTB are responsible for the generation of the human ABO(H) blood group A and B antigens (25, 26). GTA catalyzes the transfer of GalNAc from UDP-GalNAc to the H antigen acceptor (α-L-Fuc-(1→2)-β-D-Gal-*O*-R, where R is glycolipid or glycoprotein) to form the A antigen, whereas GTB catalyzes the transfer of Gal from UDP-Gal to the H antigen acceptor to form the B antigen (27, 28). Initial high resolution structural studies of both GTA and GTB revealed two regions of disordered polypeptide (19). One region consisted of the last 10 residues of the C terminus, whereas 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 because of 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 177–187 (20).

GTA and GTB are the two most homologous glycosyltransferases known that utilize different nucleotide donors and differ by only 4 of 354 amino acids as follows: Arg/Gly^{176}, Gly/Ser^{235}, Leu/Met^{266}, and Gly/Ala^{268} in GTA and GTB, respectively (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^{266} and Gly/Ala^{268} have been shown to be responsible for discrimination between the two donor molecules (30–32), whereas Gly/Ser^{235} and Leu/Met^{266} significantly impact acceptor recognition (33); however, the function of the conserved mutation Arg/Gly^{176} has been elusive. Structural studies in the past have been hampered by the fact that Arg/Gly^{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 ABBB in the presence of UDP, BBBB and AABB in the presence of synthetic H antigen disaccharide α-L-Fuc*p*-(1→2)-β-D-Gal*p*-*O*$(CH_2)_7CH_3$, BBBB and ABBB in the presence of UDP, and the H antigen acceptor analog α-L-2-deoxy-Fuc*p*-(1→2)-β-D-3-amino-Gal*p*-*O*$(CH_2)_7CH_3$, BBBB, and ABBB 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-Fuc*p*-(1→2)-β-D-3-deoxy-Gal*p*-*O*$(CH_2)_7CH_3$.

EXPERIMENTAL PROCEDURES

Construction of the Synthetic Glycosyltransferase Chimeric Genes ABBB and AABB—The synthetic wild-type GTA (designated AAAA, amino acids 53–354) gene was constructed from synthetic oligonucleotides as described previously (34). 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 (amino acids 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, and 72 °C, 1 min) for 30 cycles. After gel purification, the PCR products were digested with EcoRI and HindIII for 2 h at 37 °C and were ligated into 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, Chatsworth, CA). All ligations were made by the use of T4 DNA ligase (Invitrogen) at room temperature for 1 h. All restriction enzymes were purchased from New England Biolabs.

For GTB R188S, R188K, and R188H (amino acids 63–354), site-directed mutagenesis was carried out using a QuikChange kit (Stratagene). The primers used for mutagenesis were as follows: GTB R188S (sense) 5′-ATG CGT **TCC** ATG GAA ATG ATC AGC GAC TTC TGC-3′ and antisense 5′-CAT TTC CAT **GGA** ACG CAT GGA AAC GTC CTG CC-3′. Cloning of R188K was as follows: (sense) 5′-TGG CAG GAC GTT TCC TGC GT**A** **AA**A TGG AAA TGA TCA GCG AC-3′ and antisense 5′-GTC GCT GAT CAT TTC CA**T** **TT**T ACG CAT GGA AAC GTC CTG CC-3′; and cloning of R188H was as follows: (sense) 5′-CAG GAC GTT TCC ATG CGT C**A**T ATG GAA ATG ATC AGC-3′ and antisense 5′-GCT GAT CAT TTC CAT A**T**G ACG CAT GGA AAC GTC CTG-3′. The altered nucleotides are shown in boldface.

Protein Purification—Mutant enzymes were purified from *Escherichia coli* by methods described previously (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/liter, AABB 50 mg/liter, R188S 8 mg/liter, R188K 66 mg/liter, and R188H 15 mg/liter.

Kinetic Characterization—Kinetics using α-L-Fuc*p*-(1→2)-β-D-Gal*p*-*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

Downloaded from http://www.jbc.org/ by guest on July 12, 2019

acceptor (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 $MnCl_2$, and bovine serum albumin (1 mg/ml). Seven different concentrations of donor and acceptor were employed, and initial rate conditions were linear with no more than 10% of the substrate consumed in the reaction. For the donors, the K_m values were determined at 1.0 mM acceptor, and the K_m for the acceptor was determined at 1.0 mM donor. The kinetic parameters k_{cat} and K_m were obtained by nonlinear regression analysis of the Michaelis-Menten equation with the Graph Pad PRISM 3.0 program (GraphPad Software, San Diego). Two-substrate kinetic analysis was performed for the AABB and ABBB mutants to obtain K_A (acceptor K_m), K_B (donor K_m), K_{ib}, and K_{ia}, as described previously (34). K_{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_{ia} is the dissociation constant of the enzyme·acceptor complex.

Crystallization—All proteins were crystallized using conditions different from those reported previously (10, 19, 33, 35, 38, 39). Whereas the first crystals of GTB were grown from relatively low protein concentrations (~8–15 mg/ml) and as a mercury derivative, the crystals in this paper were initially generated from higher protein concentrations (~60–75 mg/ml). The first crystals of the ABBB and AABB mutants grew in stock solutions containing 20 mM MOPS, pH 7.0, 75 mM NaCl, 15 mM β-mercaptoethanol, 0.05% NaN_3 and stored at 4 °C for several months. Crystals of ABBB and AABB were washed with mother liquor consisting of 7% PEG-4000, 70 mM Ada buffer (*N*-(2-acetamindo)iminodiacetic acid), pH 7.5, 30 mM sodium acetate buffer, pH 4.6, 40 mM ammonium sulfate, and 5 mM $MnCl_2$. Crystals of BBBB were obtained by the hanging drop method from 30 to 40 mg/ml fresh protein solutions containing 1% PEG, 4.5% methyl-pentanediol (MPD), 0.1 M ammonium sulfate, 0.07 M NaCl, 0.05 M Ada buffer, pH 7.5, 5 mM $MnCl_2$ against a reservoir containing 2.7% PEG-4000, 7% MPD, 0.32 M ammonium sulfate, 0.25 M NaCl, and 0.2 M Ada buffer, pH 7.5. Crystals of BBBB, ABBB, and AABB in complex with UDP, synthetic H antigen disaccharide α-L-Fuc*p*-(1→2)-β-D-Gal*p*-*O*$(CH_2)_7CH_3$ (HA) (8), or the analogs α-L-2- deoxy-Fuc*p*-(1→2)-β-D-3-amino-Gal*p*-*O*$(CH_2)_7CH_3$ (ADA) (40) and α-L-Fuc*p*-(1→2)-β-D-3-deoxy-Gal*p*-*O*$(CH_2)_7CH_3$ (DA) (41) in various combinations were obtained by soaking substrate into the unliganded crystals. Crystals were washed with mother liquor consisting of 7% PEG-4000, 70 mM Ada buffer, pH 7.5, 30 mM sodium acetate buffer, pH 4.6, 40 mM ammonium sulfate, and 5 mM $MnCl_2$. 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 $MnCl_2$ 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 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 and Reduction—X-ray diffraction data were 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^{2+} 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 (42).

Structure Determination—Although the structures were nearly isomorphous, for completeness all structures were solved by molecular replacement using the CCP4 module MOLREP (43, 44) with the structure of wild-type GTB as a starting model (Protein Data Bank accession code 1LZ7), and subsequently refined using the CCP4 module REFMAC5 (45). All figures were produced using Setor (46) and SetoRibbon.[3]

RESULTS

The details of data collection and refinement for the enzyme complexes are provided in Table 1 for BBBB structures and Table 2 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 (aminodeoxy-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, and AABB+UDP-Gal+DA. The maximum resolution of the diffraction data collected varied from 1.75 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 346–354) for all structures is given in Table 3. Without exception, structures containing Arg^{176} show significantly more order than the corresponding structure containing Gly^{176}. All ABBB and AABB structures display large portions of the internal loop, which is observed to adopt an "open" conformation when unli-

[3] S. V. Evans, unpublished.

Downloaded from http://www.jbc.org/ by guest on July 12, 2019

TABLE 1
Data collection and refinement results for crystal structures of BBBB

	BBBB	BBBB+UDP	BBBB+HA	BBBB+UDP+HA	BBBB+UDP+ADA
Resolution (Å)	20-1.43	20-1.75	20-1.55	20-1.69	20-1.69
Space group	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$
A (Å)	52.54	52.54	52.56	52.45	52.43
B (Å)	149.91	149.93	150.25	149.60	149.93
C (Å)	79.21	79.28	79.38	78.74	78.92
R_{merge} (%)[a,b]	0.034 (0.322)	0.051 (0.338)	0.037 (0.318)	0.048 (0.324)	0.044 (0.325)
Completeness (%)[b]	98.1 (96.9)	98.5 (96.2)	98.2 (96.0)	99.5 (99.9)	97.0 (98.4)
Unique reflections	57,030	31,525	45,171	34,923	34,204
Refinement					
R_{work} (%)[c]	21.2	19.6	21.1	19.8	20.0
R_{free} (%)[c,d]	22.1	23.1	23.5	22.2	23.2
No. of waters	236	213	240	200	196
r.m.s. bond (Å)[e]	0.009	0.014	0.011	0.014	0.017
r.m.s. angle (°)[e]	1.28	1.44	1.28	1.41	1.61
Protein Data Bank ID	2RIT	2RIX	2RIY	2RJ8	2RJ9

[a] R_{merge}, $\Sigma|I_{obs} - I_{ave}|/\Sigma I_{ave}$.
[b] Values in parentheses represent highest resolution shell.
[c] R_{work}, $\Sigma||F_o| - |F_o||/\Sigma|F_o|$.
[d] 10% of reflections were omitted for R_{free} calculations.
[e] r.m.s. is root mean square.

TABLE 2
Data collection and refinement results for crystal structures of ABBB and AABB

	ABBB	ABBB+UDP	ABBB+UDP+HA	ABBB+UDP+ADA	AABB+UDP	AABB+HA	AABB+UDP-Gal+DA
Resolution (Å)	20-1.45	20-1.52	20-1.55	20-1.47	20-1.45	20-1.41	20-1.70
Space group	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$	$C222_1$
A (Å)	52.53	52.53	52.45	52.48	52.59	52.52	52.36
B (Å)	149.58	149.35	149.12	149.74	149.02	149.06	148.65
C (Å)	79.64	79.65	79.64	79.79	79.65	79.61	79.52
R_{merge} (%)[a,b]	2.5 (23.0)	4.2 (36.0)	4.6 (33.0)	3.4 (28.9)	2.8 (20.7)	3.7 (31.5)	5.7 (31.0)
Completeness (%)[b]	91.5 (86.1)	99.9 (99.4)	97.1 (94.2)	98.6 (97.8)	97.1 (96.6)	96.2 (84.6)	95.3 (98.9)
Unique reflections	51,207	48,558	44,381	53,049	54,190	58,216	32,932
Refinement							
R_{work} (%)[c]	20.9	22.3	20.5	20.6	20.3	21.6	19.4
R_{free} (%)[c,d]	22.8	24.6	22.8	22.6	21.7	23.7	21.7
No. of waters	236	214	200	230	205	237	194
r.m.s. bond (Å)[e]	0.009	0.010	0.011	0.009	0.009	0.008	0.014
r.m.s. angle (°)[e]	1.29	1.32	1.35	1.28	1.24	1.23	1.43
Protein Data Bank ID	2RIZ	2RJ0	2RJ1	2RJ4	2RJ5	2RJ6	2RJ7

[a] R_{merge}, $\Sigma|I_{obs} - I_{ave}|/\Sigma I_{ave}$.
[b] Values in parentheses represent highest resolution shell.
[c] R_{work}, $\Sigma||F_o| - |F_c||/\Sigma|F_o|$.
[d] 10% of reflections were omitted for R_{free} calculations.
[e] r.m.s. is root mean square.

ganded, a "semi-closed" conformation when bound to UDP, and a "closed" conformation when bound to UDP or UDP-Gal and acceptor (Fig. 1, *a* and *b*). Although the structures of the BBBB enzymes display significant levels of disorder in the mobile polypeptide loops, the relative movement of the observed residues indicates that they undergo similar conformational shifts upon substrate binding.

For all structures, the internal loop itself can be divided into two portions. The first structure consists of residues 175–188 that shows significant flexibility and contains an α-helix consisting of residues 180–187. The second structure consists of residues 189–195 that adopts an α-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 but 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, whereas 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 A*XXX* enzymes and glycine in B*XXX* 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 3). 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 with the unliganded structure; however, there is clear evidence for two alternative conformations of approximately equal weight for residues 176–188 (Fig. 1, *c* and *d*). In one conformation the loop follows the path observed in the unliganded structure, whereas in the second conformation these same residues

Downloaded from http://www.jbc.org/ by guest on July 12, 2019

TABLE 3

Loop ordering in BBBB, ABBB, and AABB

Black one-letter amino acid codes correspond to unambiguous electron density for main chain and side chain atoms; green letters 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 AABB+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 from the open, to semi-closed, or closed form are outlined in yellow. Residues with one-letter amino acid codes in lowercase have not been included in the refined models. Substrate moieties that exhibit partial occupancy are shown in green.

Enzyme	Internal loop 176 181 186 191 196	C-terminal 346 351
Open	*two α-helices - some disorder – active site open*	*Disordered*
BBBB	EV Gaykr WQDVS MRRME MISDF CERR	VP knhqa vrnp
BBBB+HA	EV gaykr wqdvS MRRME MISDF CERR	VP knhqa vrnp
ABBB	EV RAYKR WQDVS MRRME MISDF CERR	VP knhqa vrnp
AABB+HA	EV RAYKR WQDVS MRRME MISDF CERR	VP knhqa vrnp
ABBB+UDP (1)	EV Raykr wqdVS MRRME MISDF CERR	VP knhqa vrnp
AABB+UDP (1)	EV RAYKR WQDVS MRRME MISdf CERR	VP knhqa vrnp
Semi-closed	*one helical structure – more disorder – active site closing*	*Disordered*
BBBB+UDP	EV GAykr wqdvs MRRME MISDF CERR	VP knhqa vrnp
BBBB+UDP+ADA	EV GAYKr wQDVS MRRME MISdf cerR	VP KNhqa vrnp
ABBB+UDP (2)	EV RaykR WQDVS MRRME MISDF CERR	VP knhqa vrnp
AABB+UDP (2)	EV RAYKR WQDVS MRRME MISdf CERR	VP knhqa vrnp
Closed	*one helical structure - little disorder – active site closed*	*Ordered*
BBBB+UDP+HA	EV GAYKr wQDVS MRRME MISdf CERR	VP KNHQA Vrnp
ABBB+UDP+HA	EV RAYKR WQDVS MRRME MISdf CERR	VP KNHQA VRNP
ABBB+UDP+ADA	EV RAYKR WQDVS MRRME MISdF CERR	VP KNHQA VRnp
AABB+UDP-Gal+DA	EV RAYKR WQDVS MRRME MISdF CERR	VP KNHQA VRNP

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 Ala^{177} and Lys^{179} 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 it 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 3 and Fig. 1*a*). The C-terminal residue His^{348} forms unambiguous hydrogen bonds with the O-2- and O-3-hydroxyl groups of the α-L-Fuc*p* moiety on the acceptor molecule (Fig. 2*a*). In contrast, the structure of ABBB in complex with ADA (remembering that ADA is the H antigen acceptor analog that lacks the O-2-hydroxyl group) shows complete electron density only for C-terminal residues Lys^{346} and Asn^{347}, main chain density only for His^{348} to Arg^{352}, and complete disorder for residues His^{348}, Asn^{353}, and Pro^{354} (Table 3 and Fig. 2*b*).

ABBB+UDP+HA shows a fully occupied glycerol molecule (cryoprotectant) in the donor binding site of the enzyme (Fig. 2*c*). 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 Arg^{188} from the internal loop (Fig. 2, *c* and *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

Downloaded from http://www.jbc.org/ by guest on July 12, 2019

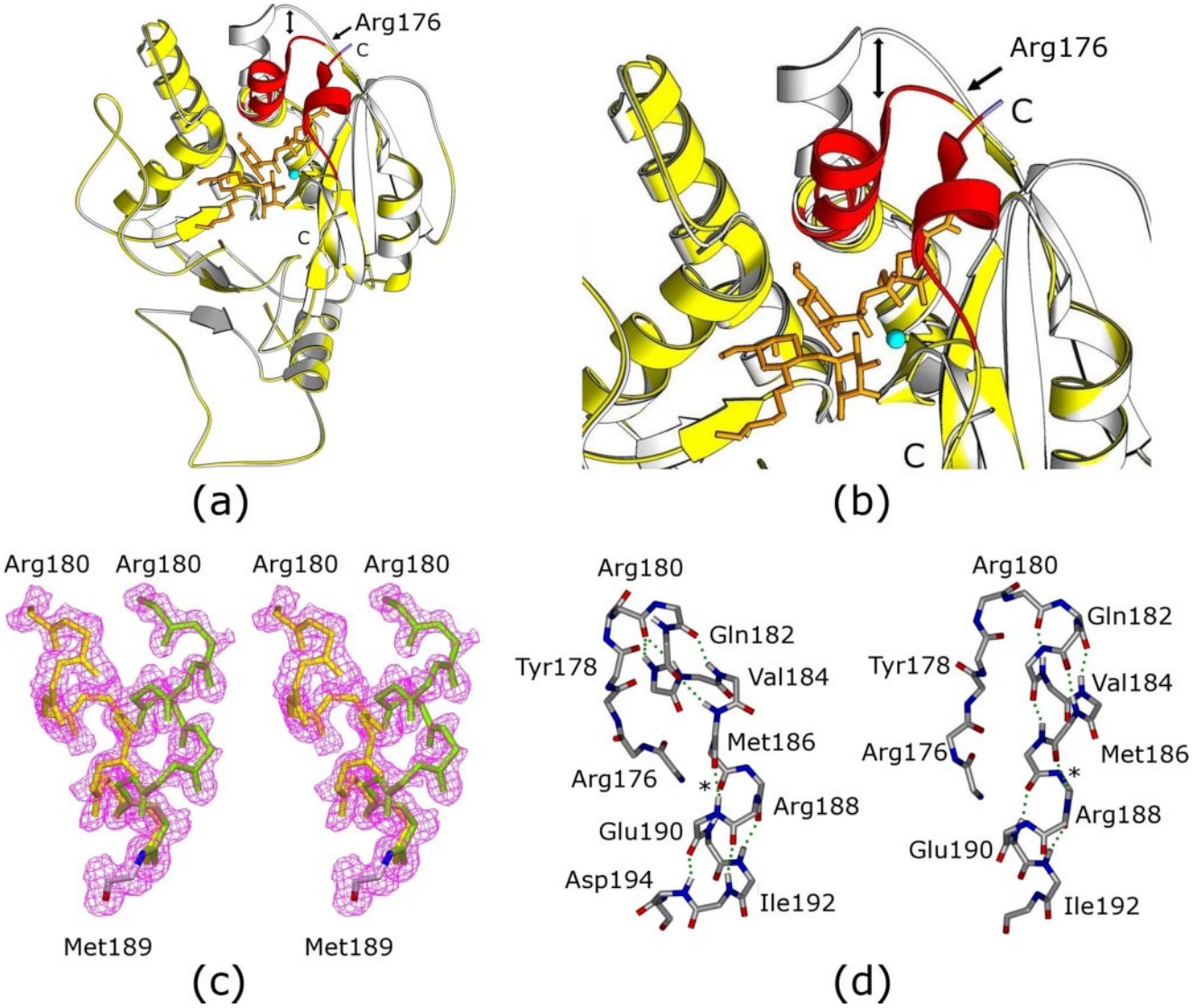

FIGURE 1. **Conformational changes associated with substrate binding.** *a,* superimposition of unliganded ABBB in the open form (*white*) with AABB+UDP-Gal+DA in the closed form (*yellow/red*) showing the internal and C-terminal loops (*red*), UDP-Gal and DA (*orange*), and Mn^{2+} (*blue*), and the location of Arg^{176}. *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 AABB+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 Met^{189} (*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 α-helices (Arg^{180}–Met^{186} and Arg^{187}–Asp^{194}) into a distorted helical structure with alternating α-3_{10}-α character. The pivot point is indicated by a *star*.

into a flexible and an ordered region for many of the structures, as all the residues from 176 to 195 can display disorder (Table 3). 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 (Ala^{177} to Arg^{180}) disordered in the internal loop of BBBB, whereas BBBB+UDP and BBBB+HA show disorder in 9 and 11 residues, respectively (Ala^{177} to Ser^{185} in BBBB+UDP and Gly^{176} to Met^{186} in BBBB+HA). The region 189–195 observed to be an α-helix in ABBB and AABB, as well as the GTB/C209A mutant (20), 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 Val^{351} and Arg^{352} interact with the O-4-carbonyl of the uracil moiety of the UDP through a bridging water molecule, and the side chain of Arg^{352} forms salt bridges with both phosphate 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. Although BBBB+UDP+HA shows the closed form of the enzyme with electron density corresponding to Lys^{346}–Val^{351}, the BBBB+UDP+ADA structure displays significantly

Downloaded from http://www.jbc.org/ by guest on July 12, 2019

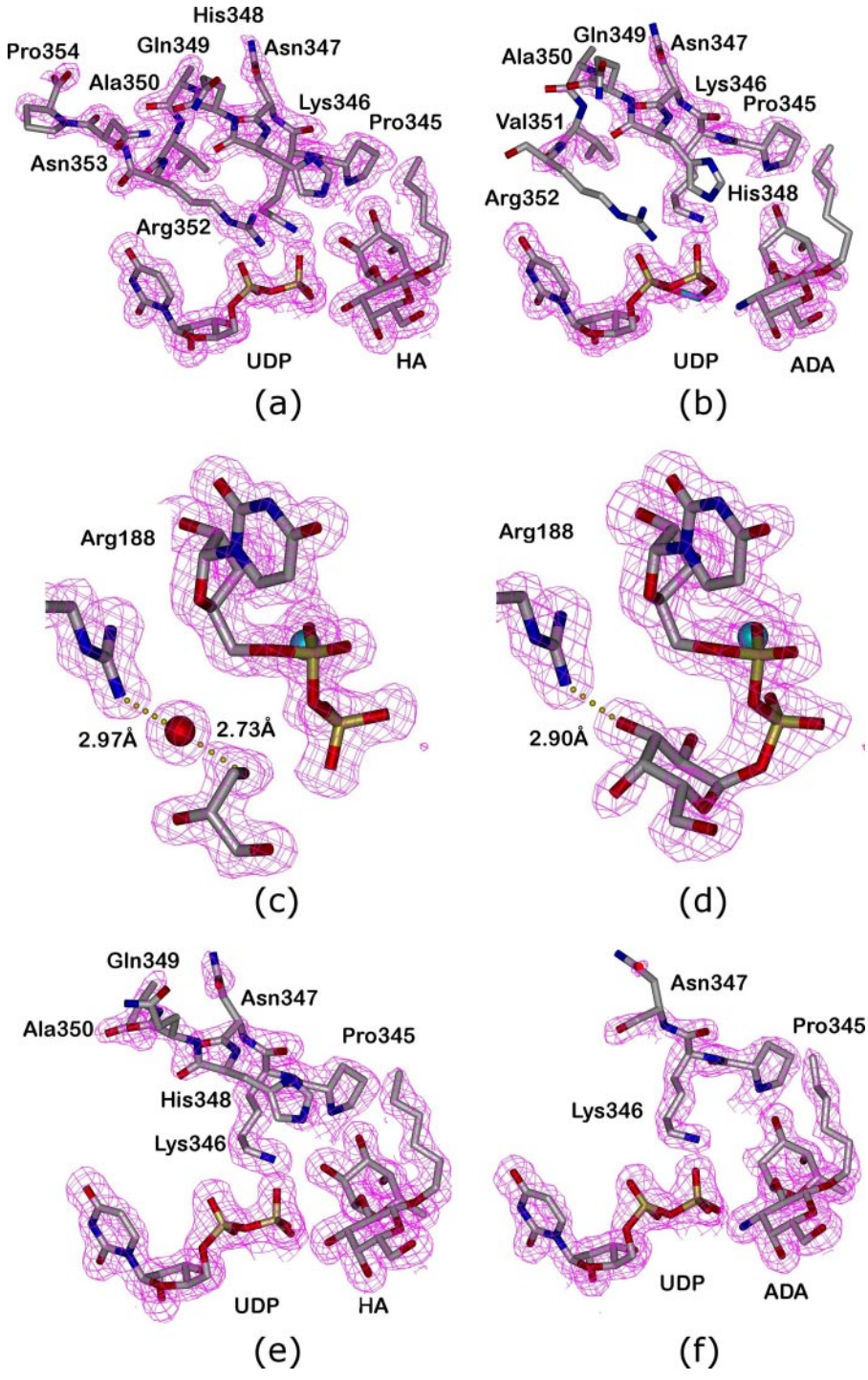

FIGURE 2. **Effects of different substrate analogs on ABBB, AABB, and BBBB.** The presence and nature of the acceptor and donor substrate analogs have a significant effect on the level of ordering of polypeptide chain. *a,* unambiguous electron density is visible for all nine C-terminal residues of ABBB in complex with UDP and H antigen disaccharide, whereas (*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 Arg^{188} in the internal loop to generate the closed form. *d,* electron density about UDP-Gal in AABB+UDP-Gal+DA showing Gal-O-3 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 BBBB+UDP+ADA (*f*). All electron density diagrams are $2F_o - F_c$ maps contoured at 1σ.

more disorder in the C-terminal region with electron density corresponding only to residues Lys^{346} and Asn^{347} (Fig. 2, *e* and *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 to prevent crystal cracking.

Kinetic Parameters—Kinetic constants for wild-type GTA, GTB and mutant enzymes are given in Table 4. The chimeric enzymes ABBB and AABB show tight binding of UDP-Gal with dissociation constants (K_{ib}) of 1.6 and 1.1 μM, respectively. Acceptor binding is also tighter for these mutants with K_{ia} values of 3.8 and 0.94 μM. The k_{cat} value for UDP-Gal for ABBB is comparable with that of GTB but is reduced from 5.1 to 2.2 s^{-1} for AABB. There was a marginal increase in k_{cat} 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/Met^{266} and Gly/Ala^{268} 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 with 69 μM; however, donor binding is weaker than that of GTA, which has a K_{ib} of 3 μM. The importance of the interaction between Arg^{188} and donor is evident from the dramatic reduction in k_{cat} 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 Arg^{176} on internal loop structure is clearly evident, as a substantial portion of the internal loop is disordered in BBBB whereas most of the loop is ordered in ABBB, where it consists of two helical segments joined at Arg^{187} (Table 3 and Fig. 1*d*). This open form is likely because of the mutual repulsion of many positively charged residues, such as internal loop residues Lys^{179}, Arg^{180}, and Arg^{188} as well as C-terminal residues Arg^{352} and Lys^{346} (Fig. 3*a*).

Downloaded from http://www.jbc.org/ by guest on July 12, 2019

TABLE 4
Kinetic constants for BBBB, ABBB, AABB, and AAAA as well as internal mutants of BBBB

Enzyme	UDP-GalNAc				UDP-Gal			
	K_A	K_B	K_{ib}	k_{cat}	K_A	K_B	K_{ib}	k_{cat}
	μM	μM	μM	s^{-1}	μM	μM	μM	s^{-1}
BBBB[a]	180	138	69	0.42	88	27	13	5.1
ABBB	75	44	9.2	0.65	12	5.2	1.6	5.4
AABB	13	37	44	0.60	1.6	1.8	1.1	2.2
BAAA[b]	48	35	ND	48	80	29	ND[b]	0.14
AAAA[a]	9.9	8.7	3	17.5	67	3.2	3.7	0.088
BBBB R188S	(too slow)				74	158	ND	0.000004
BBBB R188K	40	86	ND	0.0002	150	140	ND	0.0006

[a] Data are from Ref. 35.
[b] ND indicates not determined.

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. 1*c*). 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 (Fig. 1, *a* and *b*). The change to the semi-closed form results in the first helix of the internal loop moving into alignment with the second (Fig. 1*d*). In this shift a new main chain hydrogen bond forms between Arg^{187}(N) and Asp^{183}(O) and a main chain hydrogen bond between Glu^{190}(N) and Met^{186}(O) transfers to between Met^{189}(N) and Met^{186}(O), such that two helices become linked by a single turn of a 3_{10} helix (Fig. 1*d*). The result is a distorted helical structure with mixed α-3_{10}-α character that partially occludes the active site. The mutual repulsion of positively charged residues Lys^{179}, Arg^{180}, and Arg^{188} 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 UDP (20). Interestingly, despite high concentrations of UDP, both the ABBB+UDP and AABB+UDP structures show electron density corresponding to ~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 with 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. 3*b*).

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-Fuc*p* moiety of the acceptor, and completes the sequestration of the substrates from solvent. The side chain of Lys^{346} extends into the active site to form a salt bridge with the β-phosphate of UDP and the side chain of the third residue (Asp^{213}) of the D*X*D 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 though 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 Arg^{352} (Fig. 3*c*).

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 Arg^{188} of the internal mobile loop (Fig. 2, *c* and *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. 3*d*). The shift to the closed conformation does bring Ser^{185} and Arg^{188} into the donor sugar binding site, but the donor displays a somewhat different hydrogen bond pattern for the α-Gal moiety than predicted (47).

Although the hydrogen bonds between Asp^{211} and the O-3-hydroxyl group and between Asp^{302} and the O-4-hydroxyl group are observed, the predicted interaction between Ser^{185} and the O-6-hydroxyl group is not observed. Instead, hydrogen bonds are found between Arg^{188} and the O-3-hydroxyl group, and between His^{301} and the O-6-hydroxyl group (Fig. 3*e*). Although it does not participate in the active recognition of the donor sugar galactosyl residue, Ser^{185} is positioned to provide a steric barrier to the binding of UDP-Glc and accounts for this aspect of donor specificity. The ability of Ser^{185} 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 antigen (48). Furthermore, the

Downloaded from http://www.jbc.org/ by guest on July 12, 2019

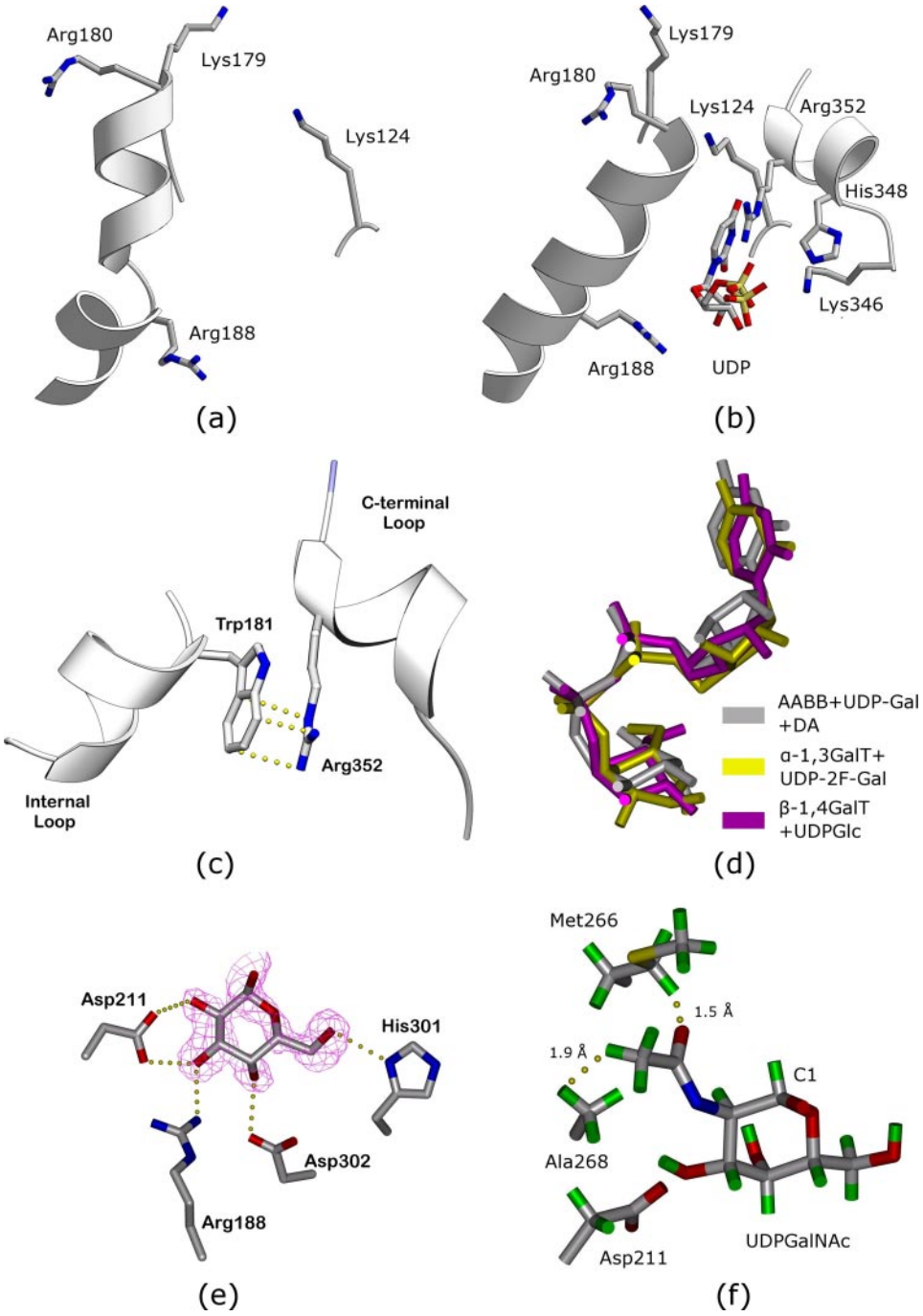

FIGURE 3. **Substrate binding and the closed conformation.** The two helices observed in the internal loop of open form (*a*) of ABBB have several positively charged side chains that are brought into proximity with other positively charged side chains on the C terminus and the nascent helix dipole to form the closed conformation (*b*) upon substrate binding in ABBB+UDP+HA. *c,* only direct contact between the internal and C-terminal loops in the closed conformation is a stacking interaction between Trp^{181} and Arg^{352}. *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 Met^{266} and Ala^{268}. Hydrogen atoms (*green*) are included for clarity.

observed position of the fully occupied nucleotide donor confirms the mechanism by which Met^{266} and Ala^{268} in BBBB distinguish between acetamido and hydroxyl groups present on the UDP-GalNAc and UDP-Gal donors, respectively (31) (Fig. 3*f*).

Recent NMR studies reported a similar conformation for the UDP-Gal bound to GTB (49). However, the observed contact by Lys^{346} 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 O-3- and O-4-hydroxyl groups of UDP-Gal are particularly important in donor substrate recognition (50).

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 K_{ib} (enzyme. donor dissociation constant) values for these enzymes (Table 4). ABBB has among the lowest observed K_{ib} value (1.6 μM) for any GTA/GTB mutant studied, whereas AABB has the lowest K_{ib} (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 nonreducing α-L-Fuc*p* moiety (51). When the original structure of GTB was solved in complex with the H antigen disaccharide, it was observed that the α-L-Fuc*p* only provided a single contact to the enzyme (a hydrogen bond between the O-4-hydroxyl and the side chain of amino acid residue Asp^{326}) (19). 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 (33); however, the absolute necessity of the α-L-Fuc*p* for catalytic activity was obscure.

In the present structures, the C-terminal residues recognize the acceptor via two hydrogen bonds to the α-L-Fuc*p* 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

Downloaded from http://www.jbc.org/ by guest on July 12, 2019

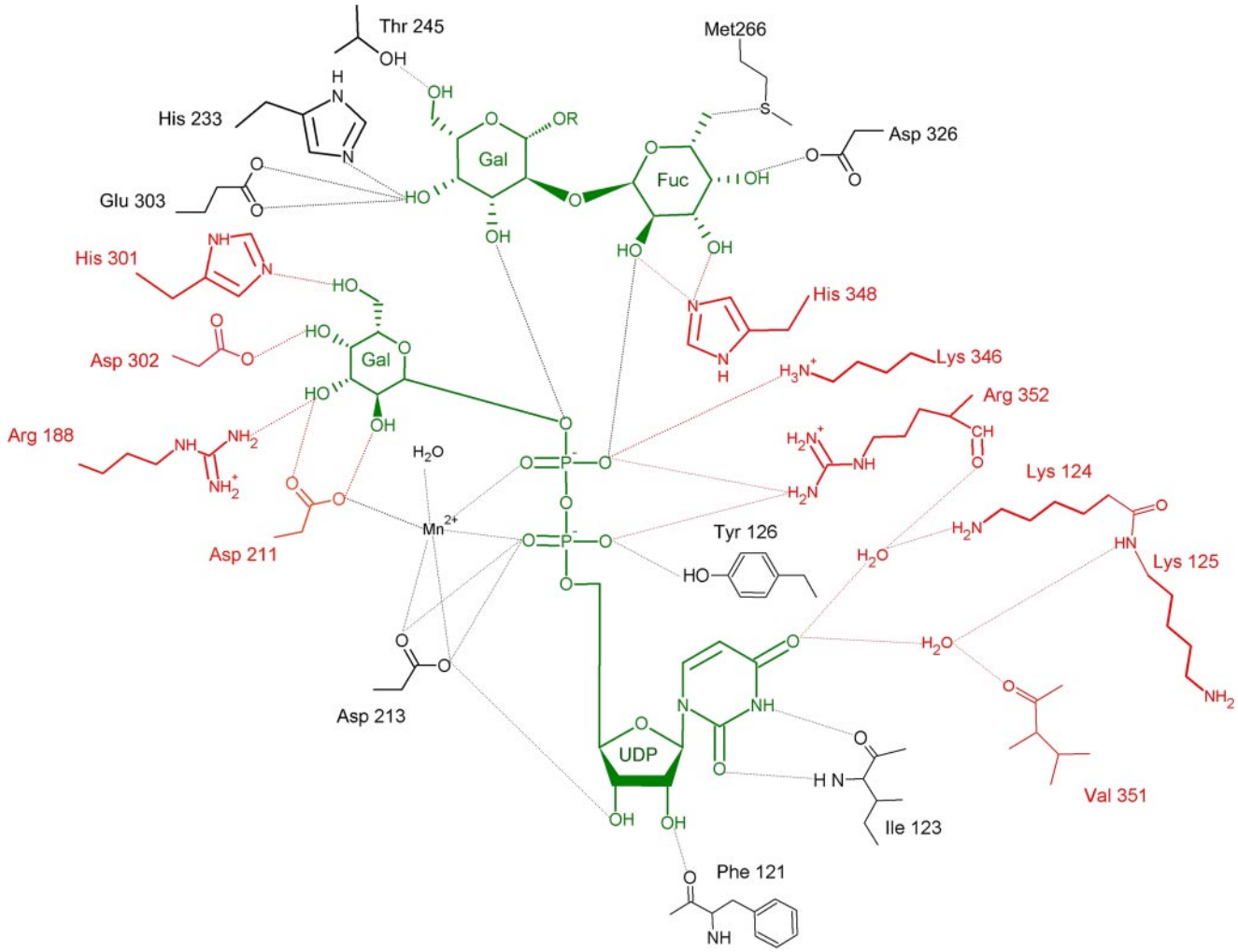

FIGURE 4. **Schematic representation of donor and acceptor recognition in GTB.** The chimeric enzyme AABB displays the closed form when bound to UDP-Gal and DA, which allows for a complete characterization of substrate recognition. The acceptor Gal-O-3 is modeled and does not appear in the 3-deoxy acceptor DA.

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 O-2-hydroxyl group of the α-L-Fuc*p* 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 3). Significantly, both structures displayed complete disorder for the side chains of His^{348} (involved in fucose recognition) and Arg^{352} (involved in UDP stabilization) (Fig. 2, *b* and *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 4 and Fig. 4). For example, the BBBB/R188S and BBBB/R188K mutants demonstrate increases in K_m values for the donor and large decreases in k_{cat}, 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 (18). 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. First, unlike BBBB, the corresponding internal loop in α-(1→3)-GalT has always been observed to be ordered in the wild-type enzyme. Furthermore, 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 five 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-un-

Downloaded from http://www.jbc.org/ by guest on July 12, 2019

der" conformation (Fig. 3*d*), 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 acceptor (18), and the presence of UDP is sufficient to fully order its C terminus (17).

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. First, 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. 3*b*). Second, the presence of acceptor stabilizes the C terminus through direct hydrogen bonds to its α-L-Fuc*p* residue, and brings it into proximity with the internal loop where there is a stacking interaction between Arg^{352} and Trp^{181} (Fig. 3*c*), and a bridging water molecule between the main chain amide group of Trp^{181} and the side chain of Asn^{353}.

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. The wild type enzyme 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 (49). 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 C-4 epimers UDP-Gal and UDP-Glc by a "tweezers" model, where the recognition of Gal-O-4 by Asp^{302} would lead to a reactive conformation, but a similar recognition of Glc-O-4 would not be possible (49). These results contrast with the structure of AABB+UDP-Gal+DA that shows that Glc-O-4 would actually be more favorably positioned to contact Asp^{302} but that the shift to the closed conformation would require a steric collision with internal loop residue Ser^{185}. 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} value for UDP-Gal that is not significantly different from the wild-type GTB enzyme (Table 4); however, the chimeric enzyme BAAA (GTA/R176G) displays a 3-fold increase in k_{cat} values for UDP-GalNAc over the wild-type GTA enzyme, showing that Arg/Gly^{176} affects 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/Gly^{176} greatly influences the level of flexibility in both the internal loop and the C-terminal region. Although 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 Lys^{189}, Ala^{177}, and Arg^{176} 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 Arg^{176} itself does not make any contact with any other part of the enzyme in any structure, and it 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 from the unliganded or HA bound structures to UDP or UDP and acceptor analog bound structures. Three distinct states have been observed for the enzymes as follows: 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/Gly^{176} 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

1. Davies, G. J., Gloster, T. M., and Henrissat, B. (2005) *Curr. Opin. Struct. Biol.* **6,** 637–645
2. Hu, Y., and Walker, S. (2002) *Chem. Biol.* **9,** 1287–1296
3. Qasba, P. K., Ramakrishnan, B., and Boeggeman, E. (2005) *Trends Biochem. Sci.* **30,** 53–62
4. Breton, C., Snajdrová, L., Jeanneau, C., Koca, J., and Imberty, A. (2006) *Glycobiology* **16,** R29–R37
5. Hu, Y., Helm, J. S., Chen, L., Ginsberg, C., Gross, B., Kraybill, B., Tiyanont, K., Fang, X., Wu, T., and Walker, S. (2004) *Chem. Biol.* **5,** 703–711
6. Somsák, L., Nagya, V., Hadady, Z., Docsa, T., and Gergely, P. (2003) *Curr. Pharm. Des.* **9,** 1177–1189
7. Greenwell, P. (1997) *Glycoconj. J.* **14,** 159–173
8. Lowary, T. L., and Hindsgaul, O. (1994). *Carbohydr. Res.* **251,** 33–67
9. Laferté, S., Chan, N. W., Sujino, K., Lowary, T. L., and Palcic, M. M. (2000)

Downloaded from http://www.jbc.org/ by guest on July 12, 2019

Eur. J. Biochem. **267,** 4840–4849
10. Nguyen, H. P., Seto, N. O. L., Cai, Y., Leinala, E. K., Borisova, S. N., Palcic, M. M., and Evans, S. V. (2003) *J. Biol. Chem.* **278,** 49191–49195
11. Bourne, Y., and Henrissat, B. (2001) *Curr. Opin. Struct. Biol.* **11,** 593–600
12. Ramakrishnan, B., Balaji, P. V., and Qasba, P. K. (2002) *J. Mol. Biol.* **318,** 491–502
13. Unligil, U. M., Zhou, S., Yuwaraj, S., Sarkar, M., Schachter, H., and Rini, J. M. (2000) *EMBO J.* **19,** 5269–5280
14. Pedersen, L. C., Tsuchida, K., Kitagawa, H., Sugahara, K., Darden, T. A., and Negishi, M. (2000) *J. Biol. Chem.* **275,** 34580–34585
15. Kakuda, S., Shiba, T., Ishiguro, M., Tagawa, H., Oka, S., Kajihara, Y., Kawasaki, T., Wakatsuki, S., and Kato, R. (2004) *J. Biol. Chem.* **279,** 22693–22703
16. Pedersen, L. C., Dong, J., Taniguchi, F., Kitagawa, H., Krahn, J. M., Pedersen, L. G., Sugahara, K., and Negishi, M. (2003) *J. Biol. Chem.* **278,** 14420–14428
17. Boix, E., Zhang, Y., Swaminathan, G. J., Brew, K., and Acharya, K. R. (2002) *J. Biol. Chem.* **277,** 28310–28318
18. Jamaluddin, H., Tumbale, P., Withers, S. G., Acharya, K. R., and Brew, K. (2007) *J. Mol. Biol.* **369,** 1270–1281
19. Patenaude, S. I., Seto, N. O. L., Borisova, S. N., Szpacenko, A., Marcus, S. L., Palcic, M. M., and Evans, S. V. (2002) *Nat. Struct. Biol.* **9,** 685–690
20. Letts, J. A., Persson, M., Schuman, B., Borisova, S. N., Palcic, M. M., and Evans, S. V. (2007) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **63,** 860–865
21. Charnock, S. J., and Davies, G. J. (1999) *Biochemistry* **38,** 6380–6385
22. Chiu, C. P., Watts, A. G., Lairson, L. L., Gilbert, M., Lim, D., Wakarchuk, W. W., Withers, S. G., and Strynadka, N. C. (2004) *Nat. Struct. Mol. Biol.* **11,** 163–170
23. Persson, K., Ly, H. D., Dieckelmann, M., Wakarchuk, W. W., Withers, S. G., and Strynadka, N. C. (2001) *Nat. Struct. Biol.* **8,** 166–175
24. Blanken, W. M., and van den Eijnden, D. H. (1985) *J. Biol. Chem.* **260,** 12927–12934
25. Hearn, V. M., Smith, Z. G., and Watkins, W. M. (1968) *Biochem. J.* **109,** 315–317
26. Kobata, A., Grollman, E. F., and Ginsburg, V. (1968) *Biochem. Biophys. Res. Commun.* **32,** 272–277
27. Watkins, W. M., and Morgan, W. T. (1959) *Vox Sang.* **4,** 97–119
28. Kabat, E. A. (1956) *Blood Group Substances: Their Chemistry and Immunochemistry,* pp. 135–139, Academic Press, New York
29. Yamamoto, F., Clausen, H., White, T., Marken, J., and Hakomori, S. (1990) *Nature* **345,** 229–233
30. Yamamoto, F., and Hakomori, S. (1990) *J. Biol. Chem.* **265,** 19257–19262
31. Seto, N. O. L., Compston, C. A., Evans, S. V., Bundle, D. R., Narang, S. A., and Palcic, M. M. (1999) *Eur. J. Biochem.* **259,** 770–775
32. Kamath, V. P., Seto, N. O. L., Compston, C. A., Hindsgaul, O., and Palcic, M. M. (1999) *Glycoconj. J.* **16,** 599–606
33. Letts, J. A., Rose, N. L., Fang, Y. R., Barry, C. H., Borisova, S. N., Seto, N. O. L., Palcic, M. M., and Evans, S. V. (2006) *J. Biol. Chem.* **281,** 3625–3632
34. Seto, N. O. L., Palcic, M. M., Compston, C. A., Li, H., Bundle, D. R., and Narang, S. A. (1997) *J. Biol. Chem.* **272,** 14133–14138
35. Lee, H. J., Barry, C. H., Borisova, S. N., Seto, N. O. L., Zheng, R. B., Blancher, A., Evans, S. V., and Palcic, M. M. (2005) *J. Biol. Chem.* **280,** 525–529
36. Seto, N. O. L., Compston, C. A., Szpacenko, A., and Palcic, M. M. (2000) *Carbohydr. Res.* **324,** 161–169
37. Palcic, M. M., Heerze, L. D., Pierce, M., and Hindsgaul, O. (1988) *Glycoconj. J.* **5,** 49–63
38. Marcus, S. L., Polakowski, R., Seto, N. O. L., Leinala, E., Borisova, S. N., Blancher, A., Roubinet, F., Evans, S. V., and Palcic, M. M. (2003) *J. Biol. Chem.* **278,** 12403–12405
39. Persson, M., Letts, J. A., Hosseini-Maaf, B., Borisova, S. N., Palcic, M. M., Evans, S. V., and Olsson, M. L. (2007) *J. Biol. Chem.* **282,** 9564–9570
40. Bai, Y., Lin, S. J., Qi, G., Palcic, M. M, and Lowary, T. L. (2006). *Carbohydr. Res.* **341,** 1702–1707
41. Lowary, T. L., and Hindsgaul, O. (1993) *Carbohydr. Res.* **249,** 163–195
42. Pflugrath, J. W. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55,** 1718–1725
43. Vagin, A., and Teplyakov, A. (2000) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **56,** 1622–1624
44. Vagin, A. A., and Isupov, M. N. (2001) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **57,** 1451–1456
45. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50,** 760–763
46. Evans, S. V. (2003) *J. Mol. Graphics* **11,** 134–138
47. Heissigerova, H., Breton, C., Moravcova, J., and Imberty, A. (2003) *Glycobiology* **13,** 377–386
48. Nakahara, T., Hindsgaul, O., Palcic, M. M., and Nishimura, S. (2006) *Protein Eng. Des. Sel.* **19,** 571–578
49. Angulo, J., Langpap, B., Blume, A., Biet, T., Meyer, B., Krishna, N. R., Peters, H., Palcic, M. M., and Peters, T. (2006) *J. Am. Chem. Soc.* **128,** 13529–13538
50. Sujino, K., Uchiyama, T., Hindsgaul, O., Seto, N. O. L., Wakarchuk, W. W., and Palcic, M. M. (2000) *J. Am. Chem. Soc.* **122,** 1261–1269
51. Schwyzer, M., and Hill, R. L. (1977) *J. Biol. Chem.* **252,** 2346–2355

Downloaded from http://www.jbc.org/ by guest on July 12, 2019

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. 2008, 283:10097-10108.
doi: 10.1074/jbc.M708669200 originally 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

This article cites 50 references, 15 of which can be accessed free at
http://www.jbc.org/content/283/15/10097.full.html#ref-list-1

Downloaded from http://www.jbc.org/ by guest on July 12, 2019