# 2.07 X-Ray Crystal Structures of Glycosyltransferases

**P. K. Qasba**, CCR, NCI-Frederick, Frederick, MD, USA
**B. Ramakrishnan**, CCR, NCI-Frederick, and SAIC-Frederick, Inc., Frederick, MD, USA
Published by Elsevier Ltd.



## 2.07.1 Introduction

Glycosyltransferases synthesize the carbohydrate moieties of the glycoproteins and glycolipids, which play a very important role in several cellular functions.[1] It is estimated that nearly 1.5–2% of the human genome is dedicated to glycosyltransferases. They exist as a large superfamily of enzymes. Mutations in glycosyltransferases are known to cause diseases in humans.[2–4] Gene knockout studies have shown that many of these transferases are essential for the survival of the organism.[5–7] Therefore, the structure and function of these transferases are important to understanding their role.

Many glycosyltransferases reside in the Golgi apparatus of a cell. Most are type II membrane proteins, consisting of four domains: a short N-terminal cytoplasmic domain, a transmembrane domain, followed by a stem region and a C-terminal catalytic domain that faces the lumen. Although the functions of the cytoplasmic domain and the stem region are not well understood, the characteristic transmembrane is responsible for the retention of these enzymes in the Golgi apparatus; and the C-terminal catalytic domain carries out the catalytic activity.[8] Often they are also glycosylated at the stem and/or at the catalytic domains.

Glycosyltransferases transfer sugar moiety from their activated donor substrate to an acceptor substrate such as proteins, or lipids or DNA, or to the sugar moiety of glycoproteins or glycolipids. Besides residing in the Golgi apparatus, they are also found in the endoplasmic reticulum (ER). The Golgi-resident enzymes use nucleotide-based activated sugars, such as nucleotide-diphosphate-α-sugar (NDP-α-sugar), as the donor substrates, whereas the ER glycosyltransferases, in addition to using NDP-sugars, also use glycolipid-based donor substrates, such as dolichol-phosphate-α-mannose.[9]

Glycosyltransferases are named based on the sugar transferred and the stereochemistry of the linkage formed. The anomeric C-atom of the donor sugar is linked either in an α- or β-configuration to one of the positions of: (1) the acceptor sugar; (2) the side-chain hydroxyl group of Ser/Thr residue of a protein; or (3) to an hydroxyl group in the polar end of a lipid. For example, the β-1,4-galactosyltransferase family transfers Gal from UDP-α-Gal in a β-configuration to the C4 hydroxyl group of the GlcNAc sugar moiety (which is present at the nonreducing end of a glycoprotein). This results in the formation of the disaccharide, Galβ1-4GlcNAc.

Glycosyltransferase is a superfamily of enzymes; each family, such as the galactosyltransferase family, consists of subfamily members that create the same linkage with the same sugar donor but transfer to different sugar acceptors. Of the galactosyltransferase family, β-1,4-galactosyltransferase, β-1,3-galactosyltransferase, and α-1,3-galactosyltransferase subfamilies inherently form variable sugar linkages as represented by their respective names; however, each of these three subfamilies transfer the same sugar moiety Gal from the donor substrate UDP-α-Gal (UDP-Gal). Although β-1,4-galactosyltransferase and β-1,3-galactosyltransferase transfer Gal to the same sugar acceptor GlcNAc, α-1,3-galactosyltransferase transfers Gal to Galβ1-4GlcNAc. While there is no protein sequence similarity among the families or the subfamilies of the glycosyltransferases, there is high protein sequence homology among members of each subfamily; for example, β-1,4-galactosyltransferase (β1,4Gal-T) has seven members, has high sequence homology (30–68%) among the members, but no homology with the β-1,3-galactosyltransferase subfamily that has five members (which have also high sequence homology among each other).

Glycosyltransferases, which are metal-ion dependent (such as $Mn^{2+}$ or $Mg^{2+}$) for their catalytic activity, have a common metal-binding motif, DxD or DxH, in their protein sequence.[10] Furthermore, these enzymes are classified based on the stereochemistry of the donor sugar and the linkage they create; for example, when the stereochemistry of the donor sugar configuration is changed from α to β or vice versa, then the enzyme is called 'the inverting enzyme', whereas if the configuration is unchanged, then it is called 'the retaining enzyme'.

Only in recent years have the crystal structures of the catalytic domain of many glycosyltransferases begun to emerge. Of the 17 structures known to date, seven are from bacteria, one from yeast, and the remaining nine from humans.[11] Lack of many crystal structures of human glycosyltransferases is mainly due to the fact that these native membrane-bound glycoproteins are found only in small quantities in biological systems. Recombinant protein expression is the only way to make large quantities of homogenous protein, which are needed for the single crystal structure studies. When expressed in *Eschericia coli*, they often form a denatured, inactive form of the protein, known as inclusion bodies. Of the nine human glycosyltransferases whose structures were solved, one was expressed in the baculovirus expression system, while two are in the yeast and the remaining six in *E. coli*. Five of the six glycosyltransferases are expressed in *E. coli* as soluble active proteins. The remaining one has been refolded *in vitro* from inclusion bodies, generating an active protein. Thus, it seems that availability of the recombinant protein in large quantities is the limiting factor for determining the crystal structures of these enzymes.

The overall fold of these proteins falls into either of two classes, called GTA and GTB fold,[12] and there seems to be no pattern between the type of the enzyme and its fold. Although most of these enzymes are classified into single-domain proteins, they seem to consist of two subdomains, in which one corresponds to the nucleotide-diphosphate-sugar-binding domain. Strikingly, many proteins that bind to UDP-sugar show high structural similarity in this domain, particularly around the metal-binding motif, even though there is no evolutionary relationship among them. However, more structural data are needed to understand the relationship among these transferases.

Due to its availability in large quantities in milk, extensive kinetics and biochemical studies on the bovine and human β-1,4-galactosyltransferase have been carried out in the past 30 years. Therefore, the recent crystal structure studies on β-1,4-galactosyltransferase, together with the previous studies, the structure, and the function of the enzyme, can well be understood. Thus, in order to understand their structure and function, the structural features of β-1,4-galactosyltransferase have been used as a model for comparing the structures of other glycosyltransferases in this chapter. The structural features, such as the evidence for the conformational change upon the donor substrate binding, and the characteristic of such conformational changes in terms of acceptor and metal ion bindings, have been discussed.

## 2.07.2 Crystal Structure β-1,4-Galactosyltransferase-I (β4Gal-T1)

β4Gal-T1 enzyme (EC 2.4.1.90/38), in the presence of $Mn^{2+}$, catalyzes the transfer of the galactosyl residue from UDP-α-galactose (UDP-Gal) to *N*-acetylglucosamine (GlcNAc) in either its free form or linked to an oligosaccharide.[13] α-Lactalbumin (α-LA), a mammary gland-specific protein that is expressed in large quantities during lactation of mammals, interacts with β4Gal-T1, forming a 1:1 complex called lactose synthase (LS) (EC 2.4.1.22) complex. α-LA has sequence and structural homology to lysozyme, yet does not have any lysozyme activity.[14–17] In the LS complex, α-LA modifies the acceptor specificity of β4Gal-T1 such that it transfers Gal from UDP-α-Gal to glucose to synthesize lactose, Galβ1-4Glc, the 'milk sugar'. In the absence of α-LA, β4Gal-T1 has low affinity for Glc, with the $K_m$ of ~2 M. In the presence of α-LA, however, $K_m$ for Glc is reduced by 1000-fold so that its physiological concentrations are adequate for the LS activity.

The catalytic domain of the bovine β4Gal-T1 (residues 129–402) has been crystallized in an apo-form, without any substrates,[18] and as a holo-form with bound $Mn^{2+}$-(UDP-Gal) (**Figures 1a** and **b**).[19,20] Comparison of the crystal structures of the apo- and holo-forms of β4Gal-T1 suggests that the enzyme undergoes conformational changes upon binding $Mn^{2+}$ and UDP-Gal, from an open to a closed conformation involving two flexible loops. It appears from the crystal structure studies that first a small loop, residues 313–316 (GWGG), undergoes conformational change by moving the Trp314 side chain from the outside to the inside of the catalytic pocket. This is followed by a change in the conformation of the long loop, residues 346–365, from an open to a closed conformation, which also creates α-helical structure for the region between the residues 359 and 365. In the holo-structure when UDP-Gal is placed inside the catalytic pocket, the Trp314 side chain not only forms a hydrogen bond with the β-phosphate oxygen atom but also interacts with the Gal moiety of UDP-Gal and the acceptor substrate through the hydrophobic interactions. Similarly, in the closed conformation, the long flexible loop not only creates the acceptor- and α-LA-binding site but also buries the bound UDP-Gal molecule.[20,21] Thus, both loops act as lids of the catalytic pocket and expose only the sugar acceptor site to the solvent environment, facilitating acceptor binding to it.

Although the Trp314 side chain in the open conformation causes steric hindrance to the long flexible loop in changing the conformation from open to closed, mutation of Trp314 to Ala314 does not seem to help the long loop to change its conformation readily from the open to closed one.[22,23] On the other hand, such mutation helps to stabilize the long flexible loop in the open conformation.[22] Molecular dynamic studies suggest that during the conformational change, the side chain of Trp314 assists the long flexible loop to change its conformation through a transient interaction with the side chain of the residue Arg359.[23] Thus, it seems that these two loops change their conformation in a coordinated way from an open to a closed one.

It has been shown that the binding of $Mn^{2+}$ to β4Gal-T1 is essential for the binding of UDP-Gal.[24,25] CD spectroscopic studies have shown that the sugar moiety of the UDP-Gal is important for inducing conformational changes, since UDP alone cannot induce it effectively.[26] The enzyme kinetics analysis shows that when UDP-Gal is used as the donor substrate, the dissociation constant of the UDP-Gal-$Mn^{2+}$-β4Gal-T1 complex, the $K_{ia}$, approaches zero, whereas with UDP-Glc as the donor substrate, the dissociation constant, $K_{ia}$, of the UDP-Glc-$Mn^{2+}$-β4Gal-T1 complex is very significant and clearly measurable.[25,27] Since in the closed conformation the UDP-sugar is buried by the flexible loops, the only way the UDP-sugar can dissociate from the enzyme-$Mn^{2+}$-(UDP-sugar) complex is by changing the closed conformation of the loops back to an open conformation. Thus, these enzyme kinetics studies, together with the crystallographic studies, suggest that only UDP-Gal can induce a stable closed conformation, and other UDP-sugar substrates, such as UDP-Glc, do not. Cross-linking studies between α-LA and β4Gal-T1 show that even the monosaccharide acceptor, GlcNAc, can induce the conformational change; however, it does so at a 100-fold higher concentration than is required when UDP-Gal is present, and an α-LA, by interacting only with the closed conformation of β4Gal-T1, stabilizes this complex.[20,28] This property is utilized for the purification of β4Gal-T1 using α-LA-agarose column, where β4Gal-T1 binds to the α-LA-agarose column in the presence of GlcNAc, and is eluted from the column in the absence of GlcNAc. Crystallographic studies show that α-LA binds to β4Gal-T1 at the C-terminal α-helical region of the long flexible loop and forms a stable complex, and it only binds to β4Gal-T1 when it is in the closed conformation.[21]

### 2.07.2.1 Metal Ion Binding and Its Specificity

The residue Asp254 of the metal-binding motif, $D^{252}xD^{254}$, of β4Gal-T1 forms a coordination bond with the $Mn^{2+}$ (**Figure 1c**) while the second residue, Asp252, forms a hydrogen bond with the donor sugar moiety. Furthermore, Asp252 interacts with the Asp254 residue through the metal ion-bound water molecule. There are two other residues, Met344 and His347, that coordinate with the $Mn^{2+}$.[29] This metal ion-binding site is in the N-terminal-binding region

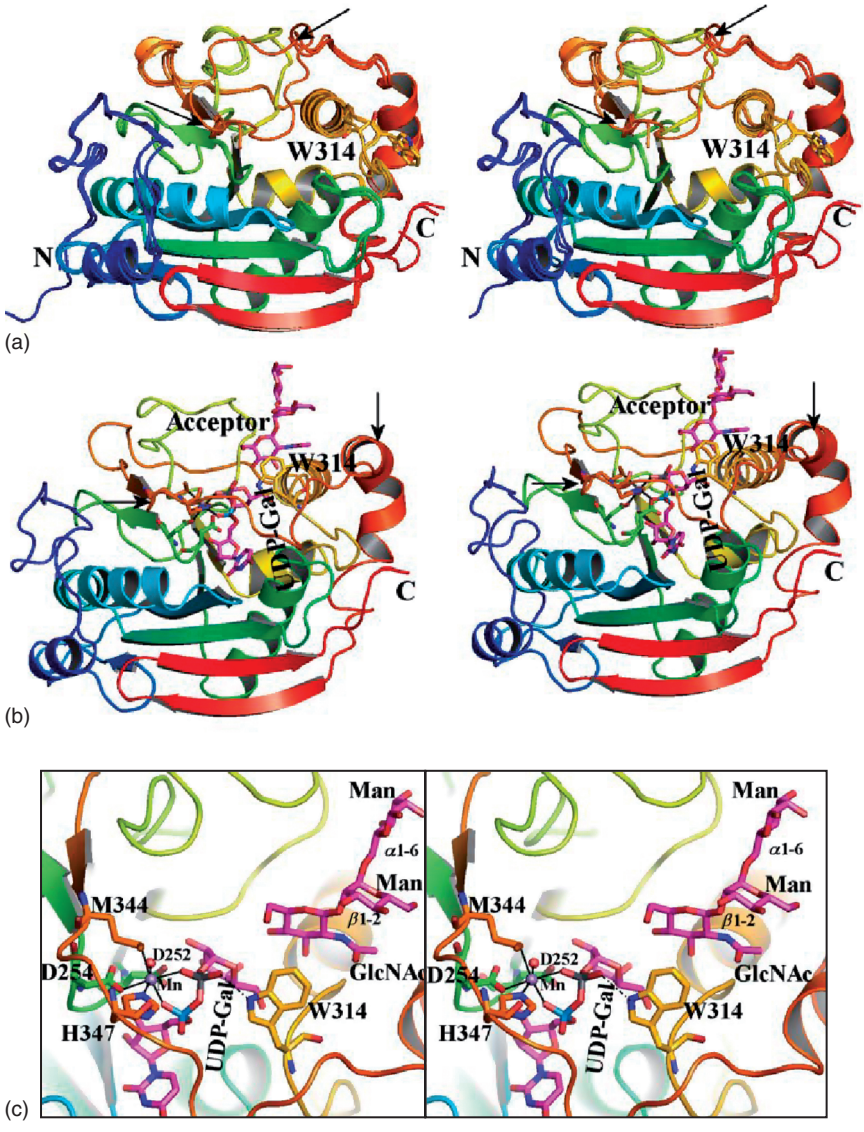


**Figure 1** a, Stereo view of the superposition of three open conformation crystal structures: bovine $\beta$4Gal-T1, bovine W314A-Gal-T1, and human M344H-Gal-T1. The arrows indicate the ends of the region that could not be traced in each of these structures. b, Stereo view of the structure of $\beta$4Gal-T1 in closed conformation with $Mn^{2+}$, UDP-Gal, and the trisaccharide acceptor, GlcNAc$\beta$1-2Man$\alpha$1-6Man. The structures are shown together, each taken from separately bound structures. The arrows point to the ends of the long flexible loop. c, Stereo view of the structure of the $Mn^{2+}$-binding region of $\beta$4Gal-T1 in closed conformation. The UDP-Gal and the acceptor, GlcNAc$\beta$1-2Man$\alpha$1-6Man structures, are shown together, each taken from separately bound structures.

of the long flexible loop, where the residues Met344 and H347 flank the hinge residue Ile345. Site-directed mutagenesis studies show that although the mutation of the Asp254 to Ala254 prevented the $Mn^{2+}$ binding to the protein, mutation to Asn254, while reducing the catalytic activity by 10 000-fold, did not inhibit the $Mn^{2+}$ binding.[29] On the other hand, the mutation of His347 not only reduced the catalytic activity but also weakened the $Mn^{2+}$ binding. In the apo-structures, which are in open conformation, the residue His347 is placed away from the metal-binding site. It coordinates with the metal ion only in the closed conformation, which is essential for the stabilization of the long flexible loop. Thus, the mutation of His347 is expected to abolish the catalytic activity of the enzyme. Met is generally considered to be a hydrophobic residue and not involved in metal ion coordination. However, so far only two crystal structures, cytochrome *C* and azurin, have Met residues that coordinate with metal ion.[30,31] When Met344 is

mutated to Ala344 in β4Gal-T1, a water molecule is found to coordinate with the $Mn^{2+}$ and the mutant Ala344-β4Gal-T1 exhibits nearly 55% of the catalytic activity, suggesting that Met344 is only weakly contributing toward the binding of $Mn^{2+}$. However, when the Met344 is mutated to other residues, such as Ser, Gln, and Glu, the enzyme progressively loses its activity in the presence of $Mn^{2+}$. Surprisingly, however, when Met344 is mutated to His344, the mutant shows high catalytic activity with $Mg^{2+}$ instead of $Mn^{2+}$. Other alkali-earth metal ions such as $Ca^{2+}$ can also bind the enzyme and activate the mutant.[32] It has been shown that only transition metal ions bind to and activate the wild-type enzyme, while alkali-earth metal ions neither bind nor activate the wild-type β4Gal-T1.[24] Considering the fact that the M344H-Gal-T1 mutant can bind both the transition- and alkali-earth metal ions, the specificity of the wild-type enzyme toward the transition metal ion must be due to the residue Met at position 344.

### 2.07.2.2 Dual Role for the Metal Ion

When a metal ion cofactor is necessary for the catalysis of an enzyme, it is expected to play an important role in the transition-state stabilization. Many glycosyltransferases use NDP-sugar as the donor substrate, and a metal ion, preferably $Mn^{2+}$, for the catalysis. $Mn^{2+}$ binds to the phosphate groups of NDP-sugar and, during catalysis, when the donor sugar separates from the NDP-sugar, the $Mn^{2+}$ neutralizes the negative charge on the NDP molecule. In addition to this role of $Mn^{2+}$, it also influences the dynamics of the conformational change in β4Gal-T1, which also contributes to the turnaround number, $k_{cat}$, of the enzyme.[32] The enzyme catalytic cycle involves, first, binding of $Mn^{2+}$, followed by UDP-Gal binding and change in the conformation of the flexible loops, from open to closed conformation, where the UDP-Gal molecule is buried inside the catalytic pocket, simultaneously facilitating the acceptor binding. After catalysis, the enzyme reverts back to open conformation, releasing the UDP molecule, and starting a new cycle.[33,34] The more frequently the enzyme goes through this conformational cycle in a second, the higher the $k_{cat}$ value. Thus, the conformational dynamics influence the $k_{cat}$ value. When the Met residue is present at position 344, as in the wild-type enzyme, the enzyme has a high $k_{cat}$ value in the presence of $Mn^{2+}$ with a $K_{dI}$ value for $Mn^{2+}$ of 16 μM, while the $Mg^{2+}$ ion does not even bind to the enzyme. However, when a His residue is present at position 344, as in the mutant M344H-Gal-T1, even though the $K_{dI}$ value of $Mn^{2+}$ decreases by about 10-fold, that is, the affinity for $Mn^{2+}$ is increased by about 10-fold, the $k_{cat}$ of mutant is decreased by 40-fold. On the other hand, the $K_{dI}$ value for the $Mg^{2+}$ is lower than $Mn^{2+}$ by 100-fold but the mutant shows a high $k_{cat}$ value. Although both $Mn^{2+}$ and $Mg^{2+}$ are divalent metal ions, they may participate differently in the stabilization of the transition-state complex that is reflected in the $k_{cat}$ values.[32]

In addition to the conformational dynamics and the transition-state stabilization, the product-release phase also influences the turnaround number, $k_{cat}$, in β4Gal-T1. In the product release state, the conformation of the flexible loops may not readily reverse back to open conformation in the M344H-Gal-T1 mutant to release the UDP from the enzyme-$Mn^{2+}$-UDP complex and, thus, facilitate the formation of the enzyme-$Mn^{2+}$-UDP-acceptor dead-end complex. The formation of such a complex is reflected in the noncompetitive inhibition of the acceptor substrate at high acceptor concentrations, which affects the $k_{cat}$ of the reactions.[32] In the M344H-Gal-T1, such inhibition is observed when $Mn^{2+}$ is used in the reaction;[32] stronger the acceptor binding lower the inhibition concentration required. However, when $Mg^{2+}$ is used in the same catalytic reaction, no such inhibition is observed. Thus, the type of metal ion binding at the hinge region of the long flexible loop influences the conformational dynamics of the long flexible loop during the conformational cycle, thereby affecting the $k_{cat}$ of the enzyme. $Mn^{2+}$, in combination with a Met residue at position 344, or $Mg^{2+}$ in combination with His residue at position 344, determines the high $k_{cat}$ of the enzyme. The inability of the M344H-Gal-T1 mutant to return to the open conformation in the presence of UDP and $Mn^{2+}$ has enabled us to crystallize the mutant in complex with various oligosaccharides.

### 2.07.2.3 The Snapshots of Crystal Structures along the Enzyme Kinetics Pathway

Among the several enzyme kinetics pathways proposed for β4Gal-T1,[35,36] the sequential ordered mechanism fits very well to the crystal structural data. $Mn^{2+}$ binds first to the enzyme, followed by the UDP-Gal. Upon the UDP-Gal binding, the enzyme-$Mn^{2+}$-(UDP-Gal) forms a stable complex to which the acceptor binds. After the enzyme catalysis, the product disaccharide is released, followed by UDP and $Mn^{2+}$. The enzyme then starts a new cycle. Previously, the crystal structures of the apo-enzyme and its complex with $Mn^{2+}$-(UDP-Gal) in the closed conformations were available.[18,19] However, to investigate the metal binding to the enzyme followed by the UDP-Gal binding in the open conformation, we used the M344H-Gal-T1 enzyme, which exhibits 10-fold higher affinity for the $Mn^{2+}$

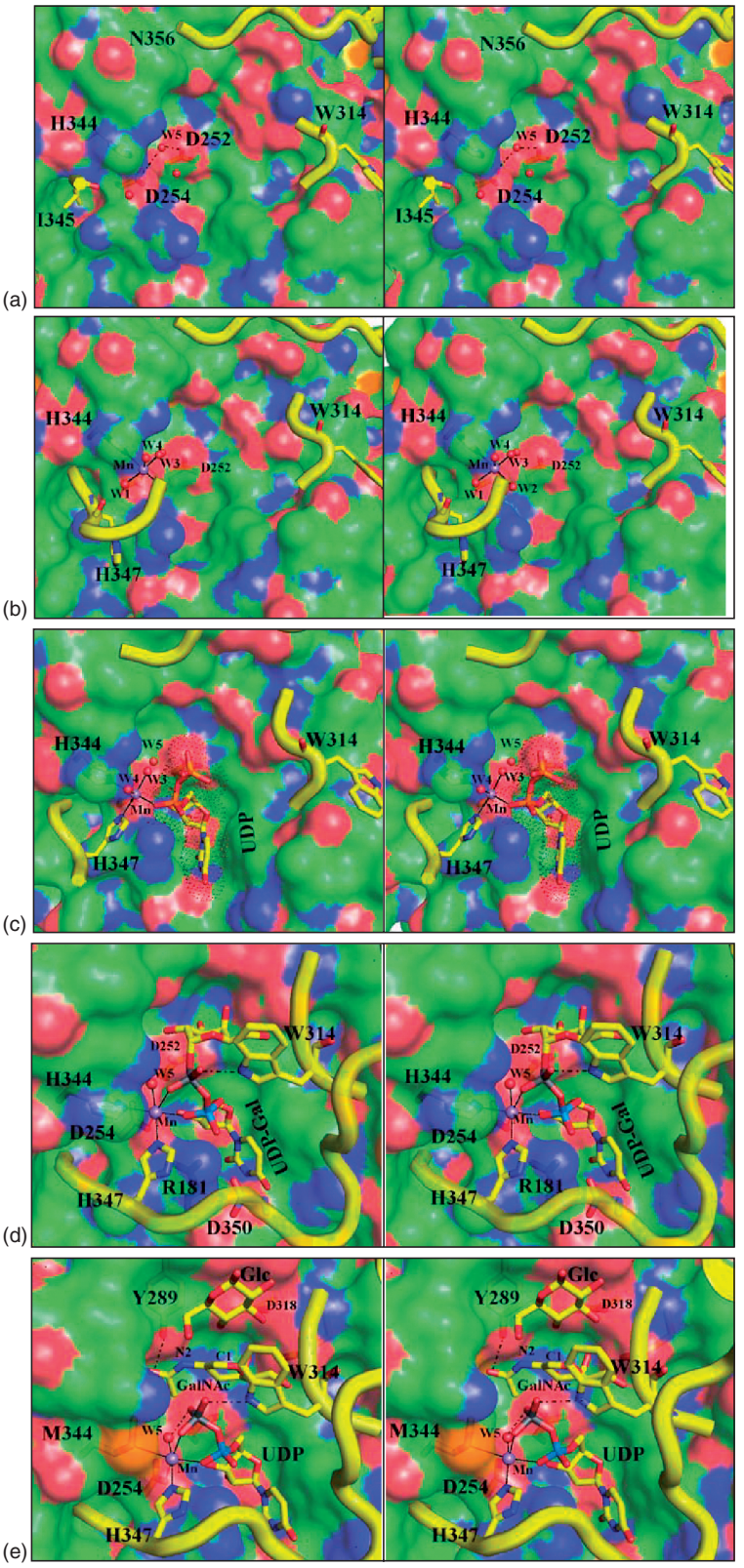


**Figure 2** (continued)

than for the wild type. In order to capture the catalytic intermediate state, we used the LS reaction, which in the presence of α-LA, UDP-GalNAc, instead of UDP-Gal, was used a donor substrate and Glc as the acceptor substrate. The $k_{cat}$ value of the LS reaction, where UDP-Gal is used as the donor substrate, is nearly 4 molecules $s^{-1}$; each catalytic cycle takes 0.25 s. However, when UDP-GalNAc is used as the donor substrate, the $k_{cat}$ value is more than 1000-fold less than that of UDP-Gal, suggesting that a single enzyme cycle takes at least 4 min to complete. Such a slow reaction is the most suitable for successfully obtaining a pentenary complex, β4Gal-T1-$Mn^{2+}$-(UDP-GalNAc)-Glc-α-LA, in the crystals.

In the apo-M344H-Gal-T1 crystal structure and in the wild-type crystal structure, the metal-binding site is simply hydrated (**Figure 2a**). A water molecule, W5, is found to bridge the Asp residues of the DxD motif and is found conserved in all the crystal structures. The residues H347 to N352 were found disordered and could not be traced. When M344H-Gal-T1 was crystallized with 10 mM $MnCl_2$, a single $Mn^{2+}$ ion was found near the $Mn^{2+}$-binding site, coordinating with the Asp254 and His344 residues; the remaining four coordinations are with four water molecules, W1, W2, W3, and W4 (**Figure 2b**). The residues His347, Ser348, and Arg349 are found ordered, and the side chain of His347 is facing away from the metal ion, not coordinating with it. In the M344H-Gal-T-$Mn^{2+}$-UDP-Gal-complex in the open conformation, the Gal moiety is disordered and cannot be located and the β-phosphate group is only partially observed (**Figure 2c**). Due to the open conformation, the Trp314 side chain is facing away from the catalytic pocket, which creates a void near the β-phosphate-binding site. In fact, mutation of Trp314 to Ala314 shows that the mutant A314-Gal-T1 binds weakly to the UDP-agarose column, suggesting the importance of the Trp314 side chain for the binding of the UDP moiety in the closed conformation. In the crystal structure of M344H-Gal-T-$Mn^{2+}$-UDP-Gal, the His347 side chain has moved toward the $Mn^{2+}$ and coordinates with it by replacing the W1 water molecule. Similarly, the W2 water molecule is replaced by the α-phosphate oxygen atom of the UDP moiety; while the conserved water molecule W5 is still present, it is 3.4 Å away from the $Mn^{2+}$. In the closed conformation of M344H-Gal-T1-$Mn^{2+}$-UDP-Gal complex, the Trp314 side chain is placed inside the catalytic pocket, forming a hydrogen bond with the β-phosphate oxygen atom (**Figure 2d**). This results in placing the β-phosphate closer to the $Mn^{2+}$ ion. Upon replacing the W3 water molecule, the β-phosphate oxygen coordinates with the $Mn^{2+}$. The coordinating water molecule, W4, found in the open conformation; is missing; instead, the W5 coordinates with the $Mn^{2+}$, fulfilling the six coordinations for the $Mn^{2+}$. This suggests that during the conformational change, due to the positioning of the long flexible loop on top of the bound UDP-Gal, the $Mn^{2+}$ coordination undergoes significant change. For its sixth coordination, $Mn^{2+}$ picks up water molecule W5 in exchange for the water molecule W4. Since $Mn^{2+}$ cannot support more than six coordinations, the W4 water molecule was not observed in all the $Mn^{2+}$-bound closed conformational structures. Since the $Ca^{2+}$ ion can support seven coordinations, the $Ca^{2+}$ ion-bound structure in the closed conformation captured both the W4 and W5 water molecules simultaneously bound to the $Ca^{2+}$ ion (data not shown).[25] Also, this exchange of coordination ligands in $Mn^{2+}$ is not due to the binding of the β-phosphate to the metal ion, since in the crystal structure of the M344H-Gal-T1-$Mn^{2+}$-GlcNAc-α-LA complex in the absence of the donor or UDP substrate, the M344H-Gal-T1 is found in the closed conformation and $Mn^{2+}$ coordinates with the W5 water molecule. In this crystal structure, in place of phosphate oxygen atoms, water molecules W2 and W3 were bound to the $Mn^{2+}$ ion, along with W5 (not W4) (unpublished results).

Therefore, it is clear from these structures that during conformational changes, even $Mn^{2+}$ coordination undergoes significant changes to coordinate with the W5 water molecule. Since the next step in the enzyme kinetics pathway is the acceptor binding and W5 is away from the acceptor-binding site, it is possible that the W5 water molecule may play

**Figure 2** a, The human Met344His-Gal-T1 apo-structure, where the metal binding site is seen hydrated. A water molecule, W5, is found interacting with the Asp252 residues of the DxD motif. b, $Mn^{2+}$ bound to human Met344His-Gal-T1 structure. The $Mn^{2+}$ is coordinated with the residues Asp254 and His344, and the remaining four coordinations are with the water molecules, W1, W2, W3, and W4. The metal-binding residue His347 is found ordered but does not coordinate with the $Mn^{2+}$. c, The human Met344His-Gal-T1 structure in open conformation with the bound $Mn^{2+}$-UDP-Gal. Due to the open conformation, the Trp314 side chain is found outside the catalytic pocket, which creates a void near the β-phosphate group. Because of this void, the Gal moiety is flexible and cannot be observed. The His347 residue coordinates with the $Mn^{2+}$ by replacing the W1 water molecule, and α-phosphate of UDP replaces the W2 water molecule. d, The human Met344His-Gal-T1 structure in closed conformation. The Trp314 side chain is placed inside the catalytic pocket displacing the β-phosphate toward the $Mn^{2+}$. By replacing the W3 water molecule, β-phosphate forms a coordination bond with the $Mn^{2+}$. The long loop is placed on the UDP-Gal. The sixth coordinating water molecule, W4, is not found; instead, W5 forms the sixth coordination with the $Mn^{2+}$. e, A pentenary complex of $Mn^{2+}$-UDP-GalNAc-Glc-Gal-T1-α-LA. The GalNAc moiety is cleaved off the UDP-GalNAc molecule and is displaced toward the Glc molecule. The β-phosphate has rotated away from the GalNAc moiety to form a hydrogen bond with the W5 water molecule, preventing the cleaved GalNAc from re-forming UDP-GalNAc. Due to the steric hindrance between the residue Tyr289 and the *N*-acetyl group of GalNAc, GalNAc has not formed the product.

an important role during the catalysis. In the crystal structure of the pentenary complex, β4Gal-T1-$Mn^{2+}$-(UDP-GalNAc)-Glc-α-LA, the GalNAc moiety is cleaved from the UDP-GalNAc and observed to be 3.6 Å away from the β-phosphate oxygen atom of the UDP moiety and toward the acceptor Glc molecule (**Figure 2e**). The increase in the distance between the β-phosphate oxygen atom and the C1 atom of the cleaved GalNAc moiety is not solely due to the change in the puckering of GalNAc from $^4C_1$ to $^4H_3$. Comparison of the β-phosphate orientation of the cleaved UDP and the uncleaved UDP-GalNAc from the crystal structure of the β4Gal-T1-$Mn^{2+}$-(UDP-GalNAc)-α-LA[37] conforms that the β-phosphate oxygen atoms have rotated about their phosphor-diester bonds by an average of 26°, away from the C1 atoms of the GalNAc moiety and toward the W5 water molecule, forming a hydrogen bond with it (distance of 2.7 Å). This rotation not only increases the distance between the phosphate oxygen atom to the C1 atom of the GalNAc moiety, but these two atoms are also no longer aligned to reform the broken glycosidic bond (**Figure 2e**). In the uncleaved UDP-GalNAc bound crystal structure, the W5 water molecule is found 3.5 Å away from the glycosidic oxygen atom. This clearly suggests that the role of the W5 water molecule is to rotate the β-phosphate oxygen atom upon the cleavage of the sugar moiety, thereby preventing the reformation of the UDP-sugar and ensuring that the catalytic reaction proceeds only in the forward direction.[37]

In the pentenary crystal structure, β4Gal-T1-$Mn^{2+}$-(UDP-GalNAc)-Glc-α-LA, the GalNAc sugar donor is not transferred, and no disaccharide is detected. GalNAc is found 2.6 Å away from the O4 atom of the Glc (**Figure 2e**). The steric hindrance caused by the Tyr289 residue, which interacts with the *N*-acetyl group of the GalNAc moiety, prevents the donor sugar from being transferred. In a separate study, when Tyr289 is mutated to a less bulky side chain amino acid, Leu289, this mutant, as expected, demonstrates similar catalytic activity with both donor substrates, UDP-GalNAc and UDP-Gal.[37] Furthermore, in the crystal structure of the pentenary complex employing UDP-GalNAc as the donor substrate (β4Gal-T1-$Mn^{2+}$-(UDP-GalNAc)-Glc-α-LA), the C1 atom of the GalNAc moiety has only two covalent bonds with the nonhydrogen atoms O5 and C2 and exists in nearly $^4H_3$ conformation. This suggests that it may exist as an oxocarbenium ion or in an *N*-acetylgalactal form, concurring with the $S_N2$ mechanism for the enzyme catalysis.[38–40]

### 2.07.2.4 α-LA Binding and Modulation of the Acceptor Specificity

Among the placenta-based mammals, α-LA and β4Gal-T1 from any mammal can form an efficient LS complex that can synthesize lactose. In fact, recent studies have shown that α-LA from mammals can form a LS complex with β4Gal-T1 from even nonmammalian vertebrates such as chicken.[41] However, such a combination seems not possible among marsupial LS complexes, where α-LA and β4Gal-T1 from the same species can only make an efficient LS complex.[42] Although bovine α-LA forms an efficient LS complex with the bovine β4Gal-T1 and both recombinant proteins are available, they do not crystallize easily. However, recombinant mouse α-LA crystallizes efficiently with the recombinant bovine β4Gal-T1 as the LS complex (**Figure 3a**). Biochemical studies have shown that LS complex is formed only in the presence of its substrates. Similarly, the two proteins form LS complex and crystallize only in the presence of various substrates of β4Gal-T1 such as GlcNAc, Glc, $Mn^{2+}$-(UDP-Gal), and also with the less preferred donor substrates $Mn^{2+}$-(UDP-Glc)[27] and $Mn^{2+}$-(UDP-GalNAc).[37] The mouse α-LA structure in the LS complex is very similar to the published structure of free α-LA from other species.[43] α-LA has a flexible region, residues 105–111, which adopts a loop or helix conformation, depending upon the pH of the crystallization medium.[44] This flexible region is also known to be important for its binding to β4Gal-T1.[45] When the flexible region of mouse α-LA adopts a helix conformation and β4Gal-T1 adopts a closed conformation; they are bound together in LS complex. As detailed above, it is only in the closed conformation of β4Gal-T1 that the α-LA binding site is available to form an LS complex. α-LA interacts with the β4Gal-T1 in the vicinity of its oligosaccharide acceptor-binding site, directly interacts with the monosaccharide acceptor molecule, and does not make any direct interactions with the donor substrate (**Figure 3a**). α-LA binds to the long flexible loop at its C-terminal α-helix region, which is created in the closed conformation, thereby stabilizing the closed conformation of β4Gal-T1. When the disaccharide chitobiose or glycoprotein ovalbumin is used as the acceptor, α-LA acts as a competitive inhibitor.[46] This suggests that α-LA competes for the oligosaccharide acceptor-binding site in the β4Gal-T1.[46] The extended sugar moiety GlcNAc of chitobiose makes hydrophobic interactions with the Tyr286 residue,[47] and in the crystal structure of the trisaccharide from the N-glycan complex with β4Gal-T1, the core mannose residue interacts with the side chain of Tyr286. In the crystal structure of the LS complex, the side chain of the residue Phe31 in α-LA makes hydrophobic interactions with the side chain of the residue Tyr286 in β4Gal-T1, explaining the competitive binding of α-LA to the oligosaccharide-binding site of β4Gal-T1.

In the closed conformation of β4Gal-T1, the C-terminal region of the long flexible loop forms an α-helix, where the side chain of the residues Arg359, Phe360, and Ile363 forms a hydrophobic pocket that facilitates the binding of the

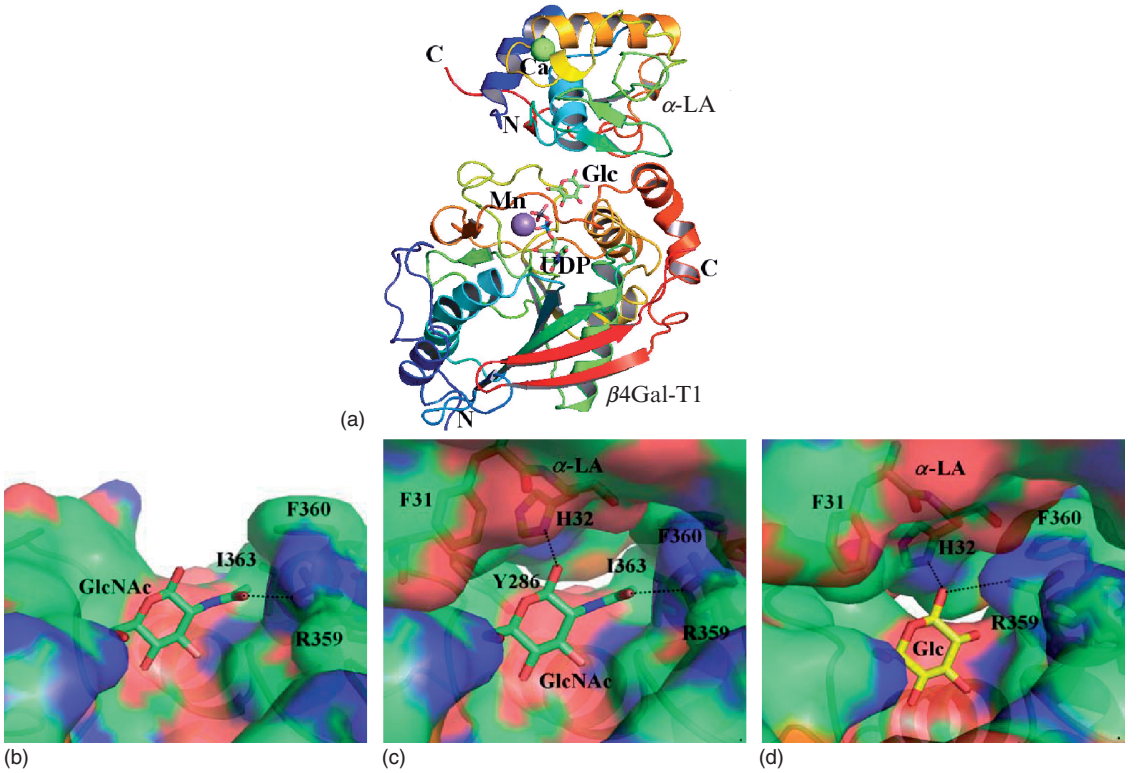


**Figure 3** a, The LS complex (a 1:1 complex between α-LA and β4Gal-T1), with the bound $Mn^{2+}$, UDP, and Glc. The α-LA interacts with the β4Gal-T1 only in the acceptor-binding site and makes hydrogen bonds and hydrophobic interactions with the acceptor substrate. b, Binding of GlcNAc to β4Gal-T1. The side chain of the residues Arg359, Phe360, and Ile363 of the C-terminal end of the long flexible loop form a hydrophobic pocket, which facilitates the binding of the *N*-acetyl moiety of GlcNAc. The C1 and O1 atoms of the acceptor do not make contacts with the protein. c, Binding of GlcNAc to the LS complex. In addition to the interaction of GlcNAc with β4Gal-T1, α-LA forms a hydrogen bond with the O1 hydroxyl group of the GlcNAc. This additional interaction decreases the $K_m$ for GlcNAc by nearly 10-fold. d, Binding of Glc to the LS complex. α-LA holds the Glc molecule in the acceptor-binding site of β4Gal-T1, through a hydrogen bond between the residue His32 and the O1 hydroxyl group of Glc. The side-chain conformation of Arg359 changes in such a way that it maximizes its interactions with the Glc molecule, thereby closing the hydrophobic pocket utilized for the binding of *N*-acetyl group of GlcNAc.

*N*-acetyl moiety of the acceptor GlcNAc. When GlcNAc binds to β4Gal-T1, the *N*-acetyl moiety interacts with this hydrophobic pocket. Besides this interaction, GlcNAc also makes a number of hydrogen bonds with the protein molecule with its exocyclic atoms, such as N2, O3, and O4, and the carbonyl oxygen atom (O7), while the C6 and O6 atoms make only van der Waals interactions with the protein molecule (**Figure 3b**). The C1 and O1 atoms of GlcNAc do not make any significant contact with β4Gal-T1.[20,21] However, in the crystal structure of the LS complex in the presence of GlcNAc, although interactions of GlcNAc with β4Gal-T1 have not altered, it makes additional interactions with α-LA. Through its highly conserved residues Phe31 and His32, α-LA makes hydrophobic and hydrogen-bonding interactions with the GlcNAc molecule (**Figure 3c**). The side chain of the residue Phe31of α-LA not only makes hydrophobic interactions with the residue Try286 of β4Gal-T1 but also with the O5 atom of GlcNAc, and the side chain of the His32 residue in α-LA forms a hydrogen bond with the O1 hydroxyl group of GlcNAc. These additional interactions observed in the presence of α-LA are responsible for the lowering of the $K_m$ value (nearly 10-fold) for the GlcNAc in the presence of α-LA in the β4Gal-T1 catalytic activity. Thus, α-LA binding restructures the acceptor binding site of β4Gal-T1 in such a way that the affinity for the acceptor substrate is increased. In the absence of α-LA, Glc is a very poor acceptor for β4Gal-T1; therefore, we can only model the binding to β4Gal-T1, based on the GlcNAc binding. In the model, the *N*-acetyl group binding pocket in β4Gal-T1is not occupied by the Glc molecule; therefore, its binding to β4Gal-T1 is expected to be much weaker than GlcNAc. However, in the crystal structure of the LS complex with Glc, the His32 residue of α-LA is observed to form a hydrogen bond with the O1 hydroxyl group

of Glc molecule, and in the acceptor-binding site of β4Gal-T1, Arg359 of β4Gal-T1 has changed its side chain conformation in such a way that the hydrophobic pocket that binds the *N*-acetyl moiety of GlcNAc is absent (**Figure 3d**). Thus, α-LA modulates the acceptor specificity of β4Gal-T1 by binding to its extended sugar acceptor-binding site, and holding the monosaccharide Glc molecule in the acceptor-binding site of β4Gal-T1 by a hydrogen bond between the His32 residue and the O1 hydroxyl group of Glc. This enables β4Gal-T1 to restructure its acceptor-binding site by changing the Arg359 side chain conformation to maximize its interactions with Glc molecule, thereby reducing its $K_m$ by nearly 1000-fold.[20,21]

Although the UDP-sugar donors other than UDP-Gal bind to β4Gal-T1, they do not, however, produce a stable β4Gal-T1-substrate complex. They dissociate easily. Due to this reason, it was impossible to crystallize these substrates in complex with β4Gal-T1. In contrast, α-LA not only binds to β4Gal-T1 in the closed conformation but also stabilizes the closed conformation with the bound donor substrate. This unique feature, which is not available with other glycosyltransferases, enabled us to crystallize β4Gal-T1 with various less preferred substrates in the presence of α-LA.[27,37] Such studies have led to a better understanding of the nature of interactions between these less-preferred sugar donor substrates, such as UDP-GalNAc, UDP-Glc, and UDP-Man with β4Gal-T1, at the molecular level. Based on these structures, design of mutant enzymes with enhanced preferences for these substrates was possible.[37,48]

## 2.07.3 Crystal Structure of the β-1,2-*N*-Acetylglucosaminyltransferase-1 (β2GlcNAc-T1)

β2GlcNAc-T1 transfers GlcNAc from UDP-α-GlcNAc to the acceptor Manα1-3Manβ-R arm (3-arm) of the $Man_5$-$GlcNAc_2$ acceptor, creating the β-linked GlcNAcβ1-2Manα1-3Manβ-R product.[49,50] This enzyme is at the gateway from oligomannose structures to hybrid and complex N-glycan structures of glycoproteins.[50] High-resolution crystal structure of the catalytic domain of the rabbit β2GlcNAc-T1 (residues 106–447) has been determined as an apo-enzyme without any substrates and as a holo-enzyme with $Mn^{2+}$ and UDP-GlcNAc (**Figures 4a** and **4b**).[51] In the apo-crystal structure, a small region, residues 318–330, present on the surface of the protein could not be traced in the crystal structure (**Figure 4a**).[51] However, in the holo-crystal structure, where $Mn^{2+}$ and UDP-GlcNAc were co-crystallized with the protein molecule, this region could be clearly traced (**Figure 4b**).[51] The conformation of this loop is such that it buries the bound $Mn^{2+}$ and UDP-GlcNAc, suggesting that this flexible loop acts like a lid to the $Mn^{2+}$- and the donor substrate-binding pocket, and only upon their binding is its conformation stabilized in such a way that it buries them. Thus, the binding of $Mn^{2+}$ and the UDP-GlcNAc induces a conformational change in this flexible loop, from a disordered open conformation to an ordered, closed conformation. In the closed conformation, the region comprising residues 324–330 forms an α-helix, and the residues Phe326 and Leu331 make hydrophobic interactions with the *N*-acetyl group of the donor sugar moiety, GlcNAc.[51] This is somewhat similar to β4Gal-T1, where the newly formed α-helix forms a hydrophobic pocket for the binding of the *N*-acetyl moiety of the acceptor substrate, GlcNAc.

The protein has an ExD as the metal-binding motif, in which the residue Asp213 forms a single coordination with the $Mn^{2+}$, whereas the residue Glu211 does not form a direct coordination bond with the metal ion (**Figure 4c**). On the contrary, as in the β4Gal-T1, Glu211 interacts with the metal ion through a water molecule. The $Mn^{2+}$ exhibits an octahedral coordination with three bound water molecules. Although the metal ion is bound to the protein through only one residue, it interacts with the backbone of the flexible loop through its water molecule, which indirectly stabilizes the flexible loop in the closed conformation (**Figure 4c**). The fact that only one Asp residue, Asp213, is interacting with the metal ion explains the requirement of a high concentration of $Mn^{2+}$ for the maximum activation of the enzyme. Furthermore, the enzyme also may not have a specific requirement for a transition metal ion for its activation. In fact, it has been shown that nearly 20 mM concentration of $MnCl_2$ is required to observe a maximum activity of the enzyme; in addition, alkali-earth metal ions such as $Mg^{2+}$, $Ca^{2+}$, and $Ba^{2+}$ activate the enzyme.[52] However, they exhibit very little activity. This requirement for high concentration of $Mn^{2+}$ for the maximum activity of β2GlcNAc-T1 is in contrast to β4Gal-T1, where three protein residues coordinate with the metal ion and only 5 mM $MnCl_2$ is enough for maximum activation of the enzyme.

Although the acceptor sugar, mannose, binding is not known, in the closed conformation its binding site can be clearly observed with Asp291 as the catalytic base.[51] However, since the protein exhibits a high acceptor preference toward the (1-3)-arm of the N-glycan,[52] an extended sugar-binding site which facilitates the binding of at least one more sugar with an α1-3 linkage might be present on the surface of the protein. This crystal structure has been used as a model system to examine the catalytic mechanism of the inverting enzymes through quantum energy calculations.[53]