NMR structural characterization of substrates bound to the alpha-2, 6-sialyltransferase, ST6Gal-I

Shan Liu, **Lu Meng**, **Kelley W. Moremen**, and **James H. Prestegard**[*]
Complex Carbohydrate Research Center, University of Georgia, Athens, GA30602, USA

Abstract

The alpha-2,6-sialyltransferase, ST6Gal-I, is a key enzyme that regulates the distribution of sialic acid-containing molecules on mammalian cell surfaces. However, the fact that its native form is membrane-bound and glycosylated has made structural characterization by X-ray crystallography of this eukaryotic protein difficult. Its large size (~40 KDa for just the catalytic domain) also poses a challenge for complete structure determination by nuclear magnetic resonance (NMR). However, even without complete structure determination, there are NMR strategies that can return targeted information about select regions of the protein, including information about the active site as seen from the perspective of its bound ligands. Here, in a continuation of a previous study, a spin-labeled mimic of a glycan acceptor ligand is used to identify additional amino acids located in the protein active site. In addition, the spin-labeled donor is used to characterize the relative placement of the two bound ligands. The ligand conformation and protein-ligand contact surfaces are studied by transferred nuclear Overhauser effects (trNOEs) and saturation transfer difference (STD) experiments. The data afforded by the above methods lead to a geometric model of the bound substrates that in many ways carries an imprint of the ST6Gal-I binding site.

Nuclear magnetic resonance (NMR) methods, focused on bound ligand properties rather than complete protein structures, can provide a very detailed view of some of the most critical regions of protein systems, including the active sites of enzymes (1–4). Previously, in one type of NMR experiment, we reported using a spin-labeled donor analog of a sugar nucleotide, CMP-4-O-[4-carboxy-2,2,6,6-tetramethylpiperidine-1-oxyl] (CMP-4carboxyTEMPO) to map the positions of isotopically labeled phenylalanines relative to the position of the analog in the active site of the alpha-2,6-sialyltransferase (ST6Gal-I) (5). The paramagnetic TEMPO derivative perturbs NMR resonances of the protein in a manner dependent on the inverse sixth power of the distances between the unpaired electron of the nitroxide moiety in TEMPO and the nuclei giving rise to the NMR resonances. Sixteen NMR cross-peaks were observable in ^{1}H-^{15}N HSQC spectra when the protein was expressed in HEK 293 cells grown on media containing ^{15}N labeled phenylalanine; the sites were ranked in the order of their distance to the bound donor mimic based on perturbation of cross-peak intensity. Glycosyltransferases, however, have two ligands, a donor and an acceptor, offering the opportunity to use similar ligand-based NMR methods to map a more extended binding site (1,6). Here previous studies are first extended using a spin-labeled acceptor analog to map distances to additional labeled amino acid sites. Then related paramagnetic strategies are used to characterize the position of the sugar acceptor relative to the donor analog while both are simultaneously bound to ST6Gal-I. When combined with more traditional transferred NOE (trNOE) (7) and saturation transfer

SUPPORTING INFORMATION AVAILABLE

Data on STD experiments using the ligands CMP-NeuAc and CMP-3FNeuAc are presented in two figures. This material is available free of charge via the Internet at http://pubs.acs.org.

difference (STD) (8) methodology, the later methods provide an imprint of the active site geometry without the need for a structure determination of the protein.

ST6Gal-I is a structurally uncharacterized glycosyltransferase of biomedical importance. It catalyzes the transfer of sialic acid residues from CMP-NeuAc to the terminal galactose residues of glycan chains of glycolipids and glycoproteins on the surfaces of mammalian cells. Variation in the amount of sialic acid incorporated is correlated with colon cancer, brain tumors and immune regulation in the human body (9–10). The cDNA encoding ST6Gal-I was first cloned from rat liver (11) and biochemically characterized in 1982 (12–13). However, the structure of ST6Gal-I remains a challenge because of the normal membrane association, native glycosylation, and poor expression in bacterial hosts. Crystal structures of four sialyltransferases of bacterial origin have been reported. They include the 2,3-/2,8-sialyltransferase (CstII) from *C. jejuni* (OH4384) (14), the 2,3-sialyltransferase (Cst-I) from *C. jejuni* (15), the 2,3/6- sialyltransferase (Δ24PmST1) from *P. multocida* strain P-1059 (16–17) and, most recently, the α 2,6-sialyltransferase (JT-ISH-224) from *Vibrionaceae Photobacterium* sp. (18). In several cases, these structures have been produced in complex with the inactive donor analog, CMP-3FNeuAc, or the product, CMP; in the two latter cases, a potential acceptor, lactose, has also been included. Unfortunately, there are no overall sequence similarities between these bacterial sialyltransferases and the mammalian sialyltransferases. The sequence identities for ST6Gal-I to CstII, Cst-I, Δ24PmST1 and JT-ISH-224 are 11%, 14%, 15% and 16% respectively. Currently no information is available on the relative orientation of donor and acceptor co-occupying the active site of ST6Gal-I. Also, the differences in reactions carried out by the bacterial sialyltransferases, which include 2–3 and 2–6 donor activities as well as broader acceptor specificities, make it risky to draw conclusions from the position of lactose in the bacterial enzymes about active site geometries of ST6Gal-I.

There have been a number of examples using paramagnetic constraints to examine distances between substrates and protein residues detected by NMR (1,5,19–23). The spin-labeled moiety chosen here is the commonly used and chemically stable TEMPO moiety (2,2,6,6-tetramethylpiperidine-1-oxyl). Its nitroxide group contains an unpaired electron that can significantly increase relaxation rates of protons and other magnetic nuclei. The effect is relatively long range because the electron magnetic moment is 658 times larger than the proton magnetic moment and even larger when compared to most other nuclei. Effects also are easily measured; as the spin relaxation times are shortened (in particular T_2), the protein resonances are broadened and the cross-peak intensity in most multidimensional experiments is reduced. The general procedure is to first acquire a spectrum of the protein with the spin-labeled ligand in its oxidized form, and then again after reduction of the nitroxide radical with ascorbate. The peak intensity differences in the oxidized and reduced spectra should be directly attributable to the paramagnetic relaxation enhancement and simply related to r^{-6}. These effects are observable up to 20 Å or more in most protein ^1H-^{15}N HSQC experiments, allowing distances between the nitroxide spin and the nuclei giving rise to the cross-peaks to be calculated.

Similar distance-dependent perturbations can be measured for the ligand in lieu of the protein. The protein sample does not need isotope labeling, and sites in the ligand (usually ^1H-^{13}C) can be detected at natural abundance as long as rapid exchange exists and a sufficiently high concentration of the ligand can be used. In principle, relaxation perturbation of an exchanging ligand in the presence of a second exchanging ligand carrying a nitroxide group can be measured from any type of relaxation parameter. There are certain advantages in measuring longitudinal relaxation times (T_1) rather than the transverse relaxation time (T_2) often measured in protein systems, including the fact that exchanging ligands usually suffer some resonance broadening due to intermediate exchange rates on and off the protein. These exchange broadening effects can be difficult to separate from the paramagnetic T_2 effects. As long as

the exchange is fast compared to T_1, T_1s are not affected by the exchange process. In choosing between T_1s of protons or carbons, T_1s of carbon have an advantage because T_1s of protons in large complexes suffer from spin diffusion. There may be some concern about effects of direct electron transfer for heteronuclei as has been postulated when paramagnetic metals are used (24). However, to our knowledge, these effects have not been documented for nitroxides.

Transferred nuclear Overhauser (trNOE) and saturation transfer difference (STD) experiments are common methods for determining bound conformations and binding motifs of substrates bound to proteins (25). Transferred NOE experiments are able to demonstrate a substrate's binding to a macromolecule and provide distance restraints between proximate pairs of protons in the substrate itself. When ligand molecules bind to large receptor proteins, NOEs undergo drastic changes leading to the observation of strong negative trNOEs as opposed to the normally weak and positive NOEs of small ligands. Because of the dominance of bound-ligand properties in the measured NOE, these negative NOEs can be converted to distance restraints for proximate pairs of ligand protons, as they exist in the bound state.

In STD experiments, the binding epitope of the substrates can be determined by measuring saturation transfer from protein protons to the individual protons of the substrates when the protein is saturated. In 1D ^1H experiments, a cascade of Gaussian-shaped pulses is applied to the aliphatic protons of protein. The resultant regional saturation quickly diffuses through the protein by intramolecular spin diffusion and then to the ligand protons by intermolecular spin diffusion. Due to fast exchange of free and bound ligands, the saturation carried by the ligands is transferred into solution and detected as a saturation of specific protons in the solution form of the ligand. By subtracting this spectrum from a spectrum with protein irradiated at a frequency well away from its proton resonances, an NMR spectrum is obtained in which the intensities of ligand signals are reduced in proportion to the proximity of ligand protons to protons on the protein surface.

In what follows a spin-labeled acceptor, LacNAc-TEMPO (21) (Fig. 1), has been used to observe changes in additional resonances of ^{15}N-Phe-labeled ST6Gal-I, and these have been compared to previous changes observed during the binding of a spin-labeled donor. The donor analog, CMP-4carboxyTEMPO (Fig. 1), is then used to perturb resonances of a non-spin labeled acceptor, N-acetyl-lactosamine (LacNAc). Distances between the nitroxide of the donor and the acceptor are estimated based on carbon T_1 changes for the acceptor in the presence and absence of the spin-labeled donor. A ^{13}C-labeled version of LacNAc was incorporated to improve the precision of relaxation rate measurements which were conducted via indirect detection of carbons through protons in arrayed HSQC experiments. Finally, these new data were combined with results from trNOEs and STD experiments to provide insight into how both donor and acceptor are positioned in the active site of ST6Gal-I.

MATERIALS AND METHODS

The synthesis of LacNAc-TEMPO has been published (21); the preparation of ^{15}N-Phe ST6Gal-I has likewise been reported previously (5).

Two-step enzymatic synthesis of CMP-3FNeuAc

N-Acetyl mannosamine, 3-fluoropyruvate, cytidine triphosphate, and NeuAc aldolase (EC 4.1.3.3) were purchased from Sigma-Aldrich. CMP-NeuAc synthetase (EC 2.7.7.43) was purchased from Calbiochem. CMP-3FNeuAc was synthesized using CMP-NeuAc synthetase to react cytidine triphosphate with 3-fluoroNeuAc which was prepared from the reaction of 3-fluoropyruvate and N-acetylmannosamine catalyzed by NeuAc aldolase (26). The purification of the CMP-3FNeuAc product was done by centrifugation, ethanol precipitation (9:1 v/v)

(27) and size exclusion chromatography on a Bio-gel P2 column (extra fine, water, 4°C). The overall yield was 77%.

Three-step enzymatic synthesis of ^{13}C enriched LacNAc

$[^{13}C_6]$ labeled GlcNAc and $[13C_6]$ labeled Gal were purchased from Omicron Biochemicals and the remaining reagents were purchased from Sigma-Aldrich. The following enzymes were used, galactokinase (EC 2.7.1.6) purified in the lab following the literature procedure (28), galactose-1-phosphate uridyltransferase (EC 2.7.7.12) from Sigma-Aldrich, and galactosyltransferase (EC 2.4.1.90) from Calbiochem. A solution (1.5 ml) containing 50 mM Tris-HCl, 5 mM Mg^{2+}, 20 mM ATP, and 20 mM Gal or $[^{13}C_6]$ Gal at pH 7.9 was prepared in a 5 mm NMR tube (29). Ten units of galactokinase were added and the reaction was carried out in this tube heated to 85°C for 40 minutes. After cooling the reaction mix to room temperature, one unit of Gal-1-phosphate uridyltransferase was added together with 40 mM UDP-glucose. After 40 minutes incubation at 37°C, one unit of galactosyltransferase was added with 5 mM $MnCl_2$ and 10 mM GlcNAc or $[^{13}C_6]$ GlcNAc, and the pH was adjusted to 7.5. The reaction mixture was left standing overnight at 37°C. The crude product was purified sequentially by Centricon filtration of enzymes, anion exchange chromatography on a Dowex 1X-2 column (formate form, 1×10 cm, water) to remove the phosphates, size exclusion chromatography on a Bio-gel P2 column (extra fine, 1.5×45 cm, water) to separate the salts, and finally lyophilization to collect 4.9 mg of white solid (yield 86%).

Purification of reduced CMP-4carboxyTEMPO

While in situ reduction of nitroxides with ascorbate normally works well, in some cases it is advantageous to add nitroxide in a pre-reduced form. To prepare this form the oxyl radical of CMP-4carboxyTEMPO was therefore reduced to a hydroxyl group by the addition of 3-µl phenylhydrazine to 2.7 mg of CMP-4carboxyTEMPO in 0.5 ml methanol. After 20 minutes of gentle swirling, the solvent was evaporated on a rotary evaporator. The following steps were mostly done under argon: adding 2–3 ml ether; vortexing; decanting; repeating twice until the supernatant was clear; adding methanol to dissolve the residue and evaporating the solvent; repeating the ether wash twice; and co-evaporating the residue with methanol to remove the remaining ether. The dried sample weighed 2.1 mg. ^{1}H NMR was taken to verify the reduction and complete removal of phenylhydrazine.

Paramagnetic perturbation of ^{15}N-Phe ST6Gal-I NMR spectra

NMR samples were prepared in 10% D_2O/90% Bis-Tris buffer (20 mM; containing 200 mM NaCl) at pH 6.5 with the protein concentration at 0.3 mM. The spectra were acquired on a Varian INOVA spectrometer operating at 900 MHz for protons. The spectrometer was equipped with a triple resonance cryogenic probe having z pulsed gradient coils. Spectra were run with the fast HSQC sequence supplied as part of the Varian BioPak. Data were typically acquired at 25°C with 11990 Hz as SW, 1800 Hz as SW1, 48 indirect t1 complex points, 514 direct t2 complex points, and a recycling rate of 1.5 seconds. Each spectrum represents a total acquisition time of about 3 hours using 64 scans at each t1 point.

T_1 relaxation measurements on the carbons of ^{13}C- LacNAc

NMR samples were prepared in 10% D_2O/90% Bis-Tris buffer (20 mM; containing 200 mM NaCl) at pH 6.5 with the concentration of ST6Gal-I at 0.3 mM. The spectra were acquired on a Varian INOVA spectrometer equipped with a cryogenic z gradient probe operating at 800 MHz (for protons). The spectra were run using the ^{13}C T1 relaxation mode of the gChsqc sequence in the Varian BioPak. Data were typically acquired at 25°C with 6828 Hz as SW, 7038 Hz as SW1, 64 indirect F1 complex points, 819 direct F2 complex points, and a recycling delay of 1.2 seconds. The relaxation delays after Z storage were 0.01, 0.08, 0.15, 0.25, 0.4, 0.6,

0.9, 1.3 seconds. Each set of arrayed HSQC spectra represent a total acquisition time of about 9 hours.

trNOE & STD NMR data

NMR samples were prepared in deuterated 12 mM phosphate buffer (containing 200 mM NaCl) at pH 6.5. The spectra were acquired with the Varian BioPak pulse sequences on Varian INOVA spectrometers operating at 600 or 800 MHz (^{1}H frequency). The trNOE experiments were performed using zero-quantum filtered 2D NOESY at an array of mixing times (50ms, 100ms, 200ms, 500ms, 1s, 2s) and a recycling rate of 1.2 seconds. The STD experiments followed the saturation transfer 1D protocol (DPFGSE 1D) without a T_1 filter or solvent saturation, and used a 100 Hz Gaussian saturation pulse train of 450 ms centered at 0.8 ppm in the aliphatic region. The off-frequency spectra were recorded by irradiating at +24 ppm. The recycling delay was 2.5 seconds.

Calculation of distances from relaxation enhancements

T_1 spin relaxation data were fit to single exponential curves using Rate Analysis in the NMRView software. The extracted relaxation rates ($R_1 = T_1^{-1}$) were then related to distances between the spin label and the site of measurement using the following equations:

$$R_{1C(NO)}/R_{1CH}=(r_{CH}/r_{C(NO)})^6 * (\gamma_E/\gamma_H)^2 \text{ (x 0.5 for the two protons on C6 carbons)} \tag{1}$$

$$R_{1CHobs}=(1-F_b) * R_{1CHfree}+F_b * R_{1CH} \tag{2}$$

$$R_{1C(NO)obs}=(1-F_b') * R_{1CHfree}+F_b' * (R_{1CH}+R_{1C(NO)}) \tag{3}$$

Here R_{1CH} is the relaxation rate due to protons directly bound to carbons. $R_{1C(NO)}$ is the relaxation rate contribution due to the presence of the nitroxide electron. $R_{1CHfree}$ is the relaxation rate contribution from the acceptor free in solution. R_{1CHobs} is the relaxation rate observed in the presence of non-spin label donor analogs and protein. $R_{1C(NO)obs}$ is the relaxation rate observed in the presence of CMP-4carboxyTEMPO and protein. F_b refers to the fraction of acceptor complexed with non-spin label donor analogs. F_b' refers to the fraction of acceptor complexed with CMP-4carboxyTEMPO in the protein solution. The bond length of C–H (r_{CH}) used in the calculations was 1.09 Å. The scaling factor (4.25×10^5) in equation (1) is based on an estimated correlation time for the protein of 20 ns and contains dependencies on gyromagnetic ratios of the unpaired electron γ_E (-1.76×1011) and the 1H nuclei γ_H (2.68×10^8).

RESULTS

Paramagnetic perturbation of ^{15}N-Phe ST6Gal-I by spin-labeled acceptor LacNAc-TEMPO

An HSQC spectrum of a 0.3 mM ST6Gal-I sample prepared with all phenylalanines labeled in ^{15}N is presented in Fig. 2A. It shows 16 cross-peaks (not yet assigned) corresponding to the 16 labeled amide sites in ST6Gal-I. When a 4 fold excess of CMP-4-carboxyTEMPO was added, 4 out of the 16 protein cross-peaks were perturbed by the presence of spin label (5) (boxed in Fig. 2B). Just as in previous work with the sugar donor analog, CMP-4carboxyTEMPO, we can identify cross-peaks belonging to phenylalanines near the LacNAc-binding site based on the steep $1/r^6$ distance dependence of the effects.

Before correlating cross-peaks with proximity to the binding site, appropriate controls must be run. Since ligand exchange dynamics can lead to peak broadening and intensity loss, it is important to compare spectra with a non-paramagnetic ligand having a structure as close as possible to that of the TEMPO analog. The simplest way to do this is to reduce the TEMPO analog. When the nitroxide radical was reduced by sodium ascorbate, the partial or complete reappearance of the perturbed resonances of the protein was observed (Fig. 2C). Changes in intensities between reduced and oxidized spectra were then related to distances between the nitroxide and the corresponding phenylalanines. The possibility of nonspecific binding of the TEMPO moiety to secondary sites was also excluded by a control experiment in which ST6Gal-I was titrated with 4-carboxyTEMPO; no changes were observed for protein resonances.

LacNAc-TEMPO perturbs three of the four peaks seen with CMP-4carboxyTEMPO (Fig. 2D). This confirms binding of the donor near the acceptor site. However, there are significant differences. For example, one peak almost completely disappears in the presence of LacNAc-TEMPO (circled in Fig. 2D) but it was not perturbed by CMP-4carboxyTEMPO. Nor was this peak perturbed by the non-paramagnetic donor mimic CMP-3FNeuAc which broadened three other protein peaks (5). A change in this particular peak did, however, occur when CMP-3FNeuAc was included along with LacNAc-TEMPO (Fig. 2E). Integrations of the peaks show that the intensity of the circled peak decreases an additional factor of two on comparing Figure 2D to 2E, and the peak reappeared to full intensity when the nitroxide was quenched (Fig. 2F). The additional decrease could come from enhanced binding of the LacNAc-TEMPO molecule in the presence of a donor. Cooperative binding of donor and acceptor has been suggested for a number of glycosyltransferases (16,30–32). The decrease could also come from a conformational change in the glycosyltransferase induced on donor binding. The latter types of conformational change have also been documented in other glycosyltransferases (16,33). We believe the latter to be the explanation in the current case, because, even in the absence of sugar donor, the binding site was nearly saturated at the concentrations of LacNAc-TEMPO used. Based on a fixed fraction of bound LacNAc-TEMPO, it is estimated that the donor brought the acceptor 6% closer to the ^{15}N-labeled Phe associated with the perturbed peak (see equation 1 in reference (5)). This observation is consistent with a loop movement seen in Δ24PmST1 in which a buried tryptophan flips out of the protein core to redefine the acceptor binding site once the donor binds to the protein (16).

Relaxation enhancement of the bound acceptor by a bound spin-labeled donor analog

The relative positions of bound acceptor and bound donor analog can also be probed using paramagnetic perturbations. The paramagnetic relaxation enhancement of the bound acceptor due to bound CMP-4carboxyTEMPO was measured by comparing the T_1 of acceptor carbons in control and in relaxation-enhanced samples. The enhancements were in turn converted into intermolecular distances between the substrates in the protein binding site. The relaxation-enhanced sample contained 0.3 mM ST6Gal-I, 1.0 mM ^{13}C-LacNAc and a near saturating amount of CMP-4carboxyTEMPO (3 mM; estimated at 79% saturation of protein using a dissociation constant (K_D) of 0.75 mM (5)). The two control samples contained ST6Gal-I, the acceptor, and the equivalent amount of CMP-3FNeuAc or reduced CMP-4carboxyTEMPO. T_1s of carbons were calculated from peak intensities measured in an arrayed ^{1}H-^{13}C HSQC experiment in which an extra z-storage, relaxation delay, and J evolution element had been inserted just before the INEPT transfer back to protons. Examples of the data are given in Fig. 3 (that from the which depicts exponential fits to a resonance showing a significant change in T_1 C4 carbon of the galactose residue in LacNAc) and a resonance showing minimal perturbation (that from the C6 carbon of the galactose residue in LacNAc).

The changes in longitudinal relaxation rate (R_1, T_1^{-1}) for the sugar residues are actually small, ranging from 0.05 to 0.37 s^{-1} (Fig. 4). This is partly due to the fact that paramagnetic effects

on heteronuclear longitudinal relaxation rates by nitroxide are inherently small at high magnetic fields, and partly due to the small fraction of the acceptor bound to ST6Gal-I while CMP-4carboxyTEMPO also was binding. With an estimated value of 1.6 mM for the acceptor, the K_D experimental sample of 0.3 mM ST6Gal-I, 1.0 mM acceptor, 3 mM CMP-4carboxyTEMPO would yield only 9% of the acceptor bound to CMP-4carboxyTEMPO-occupied ST6Gal-I. Clearly it would be advantageous to have acceptor analogues with higher affinities for these experiments. Nevertheless, some useful data can be extracted. The enhanced relaxation rates of the Gal carbon 4 and of the α-GlcNAc carbon 6 are obviously significant. The absolute propagated standard error on average is 0.14 s^{-1} based on the standard error of 5% from the exponential fit of the peak intensity data.

Based on the R_1 changes of Gal and GlcNAc carbons between the control and relaxation-enhanced samples (shown in Fig. 4), the distances from the donor nitroxide to the 14 acceptor carbons can be calculated. In these calculations, the internal motions of the ligand were ignored and the modulation of interactions by the tumbling correlation time for the protein was assumed to dominate relaxation in all cases. The estimates are dependent on the fraction of acceptor complexed with proteins carrying non-spin-labeled donor analogs, 9.2% for CMP-3FNeuAc and 7.3% for reduced CMP-4carboxyTEMPO, as well as the fraction of acceptor complexed with protein, carrying CMP-4carboxyTEMPO, which is calculated as 8.9%. The details of the calculation protocol are given in the methods section. The distances calculated range from 8–12 Å with 9 Å for Gal4 and 8 Å for GlcNAc6α being the shortest distances. The data are consistent with preferential exposure to nitroxide of carbons on a side of LacNAc where both Gal4 and GlcNAc6 can be exposed.

Investigation of bound conformations from trNOE and binding epitopes from STD experiments

Transferred NOEs can give useful information on the conformations of the individual bound ligands. Strong negative NOEs were observed from samples having a large excess of CMP-NeuAc and CMP-3FNeuAc in the presence of ST6Gal-I (Fig. 5). This clearly indicates the dominance of NOEs by contributions from the ligands in bound state. In Figure 5, the circled cross-peak connects the H6 proton of the cytosine ring and the H2 and H3 protons of the ribose. This NOE also was strong for CMP-3FNeuAc in the absence of the protein, but it had the opposite sign. This observation in both free and bound states supports the dominant population of the anti conformation about the cytosine-ribose glycosidic bond in which the H6 and H2/H3 protons are in close proximity. Given that the distance of cytosine H5 and cytosine H6 is 2.5 Å, the distance between cytosine H6 and ribose H3 was calculated as 2.3 Å in protein-free solution and 2.6 Å in the presence of the protein, based on their respective cross-peak intensities. These values are very close to one another and also close to distances expected based on both predictions of minimum energy structures and the conformations observed in other crystal structures. Unfortunately, there were no cross-peaks detected between the nucleotide and the donor sugar. However, the absence of cross-peaks is consistent with the long distances between nucleotide and sugar protons seen in the preferred low energy states and the states typically seen in the crystal structures. In modeling structures we will use these previously observed minimum energy conformers for the bound donor.

Transferred NOE data on the LacNAc acceptor were difficult to collect because of its weak binding to ST6Gal-I. NOEs collected at 25°C were in fact positive and difficult to distinguish from NOEs seen for LacNAc in the absence of the protein. The dominant NOEs obtained in a 1D NOE experiment, in which the anomeric proton resonance of galactose was saturated, showed the expected intra-ring connectivity to the axial H3 and H5 protons as well as a poorly resolved but distinguishable connection to the trans-glycosidic GlcNAc H4 proton. The latter observation is consistent with LacNAc favoring the low energy ψ = −123.26 (C1-O4-C4′-C5′),

φ = −64.33 (O5-C1-O4-C4′) conformation in solution. To make effects of protein binding observable, the temperature was lowered to 17°C where LacNAc in free solution showed minimal NOEs (actually a mix of positive and negative NOEs, possibly due to a mix of contributions from overall anisotropic motions and internal motions). As depicted in Fig. 6, the addition of ST6Gal-I now produces some clear negative NOEs (to Gal H3) and some negative contributions that cancel residual positive contributions at (Gal H5 and GlcNAc H4). The GlcNAc H4 contribution is consistent with maintenance of a dominant population for the minimum energy structure in the bound complex as well.

Saturation transfer difference (STD) experiments were used in an attempt to identify the binding epitopes of both the donor and the acceptor. Saturation transfer is most efficient when protein protons (connected by spin diffusion to a general protein proton pool) are in close proximity to ligand protons. When the saturated spectrum is subtracted from a reference spectrum and protein peaks are removed by subtracting an STD pair without ligand, the amplitude of the remaining peaks on the resulted difference spectrum correlate with proximity of the respective ligand protons to protons on the protein surface. STD data for both CMP-3FNeuAc and the natural donor, CMP-NeuAc, were collected. These data are presented in Supplemental Figures S1 and S2. Data collection on the natural donor was of higher quality, but its interpretation was complicated by the slow hydrolysis of the donor because the cytosine H5 and ribose proton resonances from the CMP product overlap resonances from the donor. Data for the non-hydrolysable CMP-3FNeuAc are summarized in Fig. 7. STD percentages of the ligands were calculated as the ratio of individual peak intensities in the corresponding difference spectrum and in the reference spectrum, and normalized using the STD effect of the ribose H1. The derived STD percentages of CMP-3FNeuAc show a clear preferential enhancement of the nucleotide base and adjacent parts of the ribose ring over the 3FNeuAc ring. This observation is consistent with the evidence that CMP, contained in both ligands, is an efficient competitive inhibitor of ST6Gal-I. The less efficient magnetization transfer to the sugar residue is consistent with the necessity of having space near the sugar residue to accommodate the acceptor. The binding epitope of CMP-3FNeuAc demonstrates that CMP serves as a recognition motif for ST6Gal-I that provides reasonable affinity for a number of different CMP-sugar donors (34).

An STD effect for LacNAc could only be detected when the ratio of the ligand to the protein was 35-fold or greater and the sample was cooled down to 5°C. However, the effects were small and quite uniformly distributed over all protons of the LacNAc moiety. This can be the result of binding in a mode where magnetization transfers from protein protons to ligand protons are inefficient compared to intra-ligand transfers, such as would occur when hydrogen bonding interactions dominate rather than close hydrophobic contacts.

DISCUSSION

Based on the data presented we can make several statements about the placement of donor and acceptor in the binding site of ST6Gal-I. Even in the absence of a structure for the protein, this placement can provide insights into a mechanism of action and donor/acceptor modifications that might alter this action. First, we can make some qualitative statements about binding site occupancy and relative placement of donor and acceptor from the data on perturbation of phenylalanine ^{1}H-^{15}N HSQC cross-peaks using a spin-labeled acceptor analog. The fact that the acceptor analog perturbed three of the four peaks that were previously perturbed by a spin-labeled donor analog suggests that they clearly bind in the same general region on the protein surface. However, the fact that one HSQC cross-peak, not perturbed by the spin-labeled donor analog, was perturbed by the spin-labeled acceptor analog suggests that the reducing end of LacNAc (where the spin label is bound) is well removed from contact with the donor sugar (TEMPO in the donor analog). This is as it should be in a complex poised for addition of the

donor sugar to the non-reducing end of the acceptor. These observations will become more useful as the cross-peaks in the HSQC spectra of ST6Gal-I are assigned and a structure is determined.

In principle, related paramagnetic relaxation enhancement (PRE) experiments using the spin-labeled donor analog and observing perturbations of HSQC peaks from an acceptor, allow determination of approximate distances between carbons of the acceptor and the nitroxide of the donor, and construction of a more complete model for placement of donor and acceptor in the ST6Gal-I binding site. Perturbations are, however, weak and broadly distributed except for two well separated carbons, Gal-4 and GlcNAc-6. This may suggest that a distribution of sites is occupied in which the spin label can make close approaches to either site. This would require, at a minimum, exposure of surfaces on the acceptor carrying these two sites.

Figure 8 depicts a model in which the Gal-4 and GlcNAc-6 sites are exposed. The structure of α-LacNAc shown here is the minimal energy conformer generated by Glycam Biomolecular Builder (35) (also http://glycam.ccrc.uga.edu/). The color code indicates the proximity of the LacNAc carbons to the nitroxide of CMP-4carboxyTEMPO with red for closest (Gal4, GlcNAc6), yellow for second closest (Gal3, GlcNAc 4, GlcNAc5), and green for the furthest (Gal5, Gal6, GlcNAc2, GlcNAc3). If we assume the less perturbed side on the acceptor is protected by binding to the protein and represents regions that must be conserved for acceptor recognition, the model is consistent with the pattern of the groups rejecting/accepting modifications in the specificity studies of the ST6Gal-I (36).

It would be most useful at this point to add the donor to this picture using experimental distances from the nitroxide to Gal-C4 and the nitroxide to GlcNAc-C6 along with an assumed distance for an SN2 attack of the Gal-O6 oxygen on the C2 of the NeuAc residue (~4Å). Unfortunately, it is not possible to generate models satisfying all three of these closest approach constraints without bringing some other carbons closer. This likely indicates enhanced contributions at C4 and/or C6 due to occupation of secondary binding sites. However, we have produced a best fit model as depicted in Figure 9A. The Gal O6 to the TEMPO C1 distance is 4.3Å. The Gal C4 to the TEMPO N4 distance is 7.5Å, the GlcNAc C6 to the TEMPO N4 distance is 8.0Å, and all other carbons in the LacNAc acceptor are at longer distances to the TEMPO N4. The regions depicted in gray represent protein contact regions as determined from STD experiments or protection from contact with the exchanging spin label.

Comparison to the placement of donor relative to acceptor as shown in the crystal structure of bacterial sialyltransferases is of some interest. There are actually two structures of a bacterial 2,3/6 sialyltransferase, Δ24PmST1 from *Pasteurella Multocida*, reported in the literature and deposited in the PDB that have both donor and acceptor present. Both structures, 2IHZ and 2IY8, have the inactive donor CMP-3FNeuAc bound, but 2IHZ has α-lactose and the 2IY8 has β-lactose. It has been suggested that the binding orientation of α-lactose in 2IHZ might represent the bound conformation for α-2,6-activity (16). The donor-acceptor pair from the crystal structure, 2IHZ, is shown in Figure 9B. In this structure, rotating the torsion angle around the C5–C6 bond of Gal can move the O6 of Gal from 6.6 to within 5.1 Å of the anomeric C2 of donor. Comparison of Figure 9A and 9B suggest that the two donor-acceptor pairs occupy similar spatial regions. Both show protein contacts with the nucleotide base and C2–C3 region of the ribose ring. Both show protein contacts with the C2–C3 regions of the Glc or GlcNAc residues of the respective acceptors, and both show protein contacts with the C6 region of the Gal residues of the acceptors. In the crystal structure there are more contacts with parts of the NeuAc residue of the donor that are not shown on the TEMPO adduct depicted in the figure, but there was a fairly high STD transfer (47%) to the H3 proton of the NeuAc residue in experiments conducted with CMP-3FNeuAc and ST6Gal-I. This level of agreement is reasonable given the differences in the enzyme sequences and donor-acceptor pairs used.

While the structure presented cannot yet be considered of sufficient resolution for detailed mechanistic modeling, it does provide a stepping stone for future work. In particular, the sequential assignments of ^{15}N-Phe ST6Gal-I resonances in the HSQC spectra should add key distances between specific amino acids and sugar donor to the model. NMR assignment for resonances from proteins that must be expressed in eukaryotic hosts is not a trivial task, however, appropriate methodology is being developed (37). Working with higher affinity acceptors should improve the sensitivity of paramagnetic probes, and orientation-sensitive data such as ligand RDCs (38–40) should greatly refine the orientation of ligands in the models. With such improvements, we can expect a better structural understanding of ST6Gal-I and a basis for structure-based design of specific inhibitors. The strategies we developed should also find important applications on probing the structures of other glycosyltransferases.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

ABBREVIATIONS

CMP, cytidine-5′-monophosphate; NeuAc, N-acetylneuraminic acid; CMP-3FNeuAc, cytidine 5′-monophosphate 3-(axial)-fluoro-N-aceytlneuraminic acid; CMP-4carboxyTEMPO, cytidine 5′-monophosphate 4-*O*-(4-carboxy-2,2,6,6-tetramethylpiperidinine-1-oxyl); GlcNAc, N-acetylglucosamine; Gal, galactose; LacNAc, N-acetyllactosamine; LacNAc-TEMPO, N-acetyllactosamine 4-*O*-(4-carboxy-2,2,6,6-tetramethylpiperidinine-1-oxyl).

References

1. Macnaughtan MA, Kamar M, Alvarez-Manilla G, Venot A, Glushka J, Pierce JM, Prestegard JH. NMR structural characterization of substrates bound to N-acetylglucosaminyltransferase V. J Mol Biol 2007;366:1266–1281. [PubMed: 17204285]

2. Becker BA, Larive CK. Probing the Binding of Propranolol Enantiomers to alpha(1)-Acid Glycoprotein with Ligand-Detected NMR Experiments. J Phys Chem B 2008;112:13581–13587. [PubMed: 18839985]

3. Constantine KL, Davis ME, Metzler WJ, Mueller L, Claus BL. Protein-ligand NOE matching: A high-throughput method for binding pose evaluation that does not require protein NMR resonance assignments. J Am Chem Soc 2006;128:7252–7263. [PubMed: 16734479]

4. Skinner AL, Laurence JS. High-Field Solution NMR Spectroscopy as a Tool for Assessing Protein Interactions with Small Molecule Ligands. J Pharm Sci 2008;97:4670–4695. [PubMed: 18351634]

5. Liu S, Venot A, Meng L, Tian F, Moremen KW, Boons GJ, Prestegard JH. Spin-labeled analogs of CMP-NeuAc as NMR probes of the alpha-2,6-sialyltransferase ST6Gal I. Chemistry & Biology 2007;14:409–418. [PubMed: 17462576]

6. Angulo J, Langpap B, Blume A, Biet T, Meyer B, Krishna NR, Peters H, Palcic MM, Peters T. Blood group B galactosyltransferase: Insights into substrate binding from NMR experiments. J Am Chem Soc 2006;128:13529–13538. [PubMed: 17031966]

7. Post CB. Exchange-transferred NOE spectroscopy and bound ligand structure determination. Current Opinion in Structural Biology 2003;13:581–588. [PubMed: 14568612]

8. Mayer M, Meyer B. Characterization of ligand binding by saturation transfer difference NMR spectroscopy. Angewandte Chemie-International Edition 1999;38:1784–1788.

9. Dall'Olio F, Chiricolo M. Sialyltransferases in cancer. Glycoconjugate J 2001;18:841–850.

10. Crocker PR, Paulson JC, Varki A. Siglecs and their roles in the immune system. Nature Reviews Immunology 2007;7:255–266.

11. Weinstein J, Lee EU, Mcentee K, Lai PH, Paulson JC. Primary Structure of Beta-Galactoside Alpha-2,6-Sialyltransferase - Conversion of Membrane-Bound Enzyme to Soluble Forms by Cleavage of the Nh2-Terminal Signal Anchor. J Biol Chem 1987;262:17735–17743. [PubMed: 3121604]

12. Weinstein J, Desouzaesilva U, Paulson JC. Purification of a Gal-Beta-1-]4glcnac Alpha-2-]6 Sialyltransferase and a Gal-Beta-1-]3(4)Glcnac Alpha-2-]3 Sialyltransferase to Homogeneity from Rat-Liver. J Biol Chem 1982;257:3835–3844.

13. Weinstein J, Desouzaesilva U, Paulson JC. Sialylation of Glycoprotein Oligosaccharides N-Linked to Asparagine - Enzymatic Characterization of a Gal-Beta-1-]3(4)Glcnac Alpha-2-]3 Sialyltransferase and a Gal-Beta-1-]4glcnac Alpha-2-]6 Sialyltransferase from Rat-Liver. J Biol Chem 1982;257:3845–3853.

14. Chiu CPC, Watts AG, Lairson LL, Gilbert M, Lim D, Wakarchuk WW, Withers SG, Strynadka NCJ. Structural analysis of the sialyltransferase CstII from Campylobacter jejuni in complex with a substrate analog. Nature Structural & Molecular Biology 2004;11:163–170.

15. Chiu CPC, Lairson LL, Gilbert M, Wakarchuk WW, Withers SG, Strynadka NCJ. Structural analysis of the alpha-2,3-sialyltransferase cst-i from Campylobacter jejuni in apo and substrate-analogue bound forms. Biochemistry 2007;46:7196–7204. [PubMed: 17518445]

16. Ni LS, Chokhawala HA, Cao HZ, Henning R, Ng L, Huang SS, Yu H, Chen X, Fisher AJ. Crystal structures of Pasteurella multocida sialyltransferase complexes with acceptor and donor analogues reveal substrate binding sites and catalytic mechanism. Biochemistry 2007;46:6288–6298. [PubMed: 17487984]

17. Ni LS, Sung MC, Yu H, Chokhawala H, Chen X, Fisher AJ. Cytidine 5′-monophosphate (CMP)-induced structural changes in a multifunctional sialyltransferase from Pasteurella multocida. Biochemistry 2006;45:2139–2148. [PubMed: 16475803]

18. Kakuta Y, Okino N, Kajiwara H, Ichikawa M, Takakura Y, Ito M, Yamamoto T. Crystal structure of Vibrionaceae Photobacterium sp JT-ISH-224 alpha 2,6-sialyltransferase in a ternary complex with donor product CMP and acceptor substrate lactose: catalytic mechanism and substrate recognition. Glycobiology 2008;18:66–73. [PubMed: 17962295]

19. Battiste JL, Wagner G. Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear overhauser effect data. Biochemistry 2000;39:5355–5365. [PubMed: 10820006]

20. Ramos A, Varani G. A new method to detect long-range protein-RNA contacts: NMR detection of electron-proton relaxation induced by nitroxide spin-labeled RNA. J Am Chem Soc 1998;120:10992–10993.

21. Jain NU, Venot A, Umemoto K, Leffler H, Prestegard JH. Distance mapping of protein-binding sites using spin-labeled oligosaccharide ligands. Protein Sci 2001;10:2393–2400. [PubMed: 11604544]

22. Johnson PE, Brun E, MacKenzie LF, Withers SG, McIntosh LP. The cellulose-binding domains from Cellulomonas fimi beta-1,4-glucanase CenC bind nitroxide spin-labeled cellooligosaccharides in multiple orientations. J Mol Biol 1999;287:609–625. [PubMed: 10092463]

23. Kuliopulos A, Westbrook EM, Talalay P, Mildvan AS. Positioning of a Spin-Labeled Substrate-Analog into the Structure of Delta-5–3-Ketosteroid Isomerase by Combined Kinetic, Magnetic-Resonance, and X-Ray-Diffraction Methods. Biochemistry 1987;26:3927–3937. [PubMed: 2888482]

24. Hansen DF, Led JJ. Mapping the electronic structure of the blue copper site in plastocyanin by NMR relaxation. J Am Chem Soc 2004;126:1247–1252. [PubMed: 14746497]

25. Meyer B, Peters T. NMR Spectroscopy techniques for screening and identifying ligand binding to protein receptors. Angewandte Chemie-International Edition 2003;42:864–890.

26. Watts AG, Withers SG. The synthesis of some mechanistic probes for sialic acid processing enzymes and the labeling of a sialidase from Trypanosoma rangeli. Canadian Journal of Chemistry-Revue Canadienne De Chimie 2004;82:1581–1588.

27. Knorst M, Fessner WD. CMP-sialate synthetase from Neisseria meningitidis - Overexpression and application to the synthesis of oligosaccharides containing modified sialic acids. Adv Synth Catal 2001;343:698–710.

28. Verhees CH, Koot DGM, Ettema TJG, Dijkema C, De Vos WM, Van Der Oost J. Biochemical adaptations of two sugar kinases from the hyperthermophilic archaeon Pyrococcus furiosus. Biochem J 2002;366:121–127. [PubMed: 11978175]

29. Gilhespymuskett AM, Partridge J, Jefferis R, Homans SW. A Novel C-13 Isotopic Labeling Strategy for Probing the Structure and Dynamics of Glycan Chains in-Situ on Glycoproteins. Glycobiology 1994;4:485–489. [PubMed: 7827410]

30. Boix E, Zhang YN, Swaminathan GJ, Brew K, Acharya KR. Structural basis of ordered binding of donor and acceptor substrates to the retaining glycosyltransferase, alpha-1,3-galactosyltransferase. J Biol Chem 2002;277:28310–28318. [PubMed: 12011052]

31. Ramakrishnan B, Boeggeman E, Ramasamy V, Qasba PK. Structure and catalytic cycle of beta-1,4-galactosyltransferase. Current Opinion in Structural Biology 2004;14:593–600. [PubMed: 15465321]

32. Unligil UM, Rini JM. Glycosyltransferase structure and mechanism. Current Opinion in Structural Biology 2000;10:510–517. [PubMed: 11042447]

33. Qasba PK, Ramakrishnan B, Boeggeman E. Substrate-induced conformational changes in glycosyltransferases. Trends Biochem Sci 2005;30:53–62. [PubMed: 15653326]

34. Blixt O, Han SF, Liao L, Zeng Y, Hoffmann J, Futakawa S, Paulson JC. Sialoside analogue arrays for rapid identification of high affinity siglec ligands. Journal of the American Chemical Society 2008;130:6680-+. [PubMed: 18452295]

35. Kirschner KN, Yongye AB, Tschampel SM, Gonzalez-Outeirino J, Daniels CR, Foley BL, Woods RJ. GLYCAM06: A generalizable Biomolecular force field. Carbohydrates. J Comput Chem 2008;29:622–655. [PubMed: 17849372]

36. Wlasichuk KB, Kashem MA, Nikrad PV, Bird P, Jiang C, Venot AP. Determination of the Specificities of Rat-Liver Gal(Beta-1–4)Glcnac Alpha-2,6-Sialyltransferase and Gal(Beta-1–3/4) Glcnac Alpha-2,3-Sialyltransferase Using Synthetic Modified Acceptors. J Biol Chem 1993;268:13971–13977. [PubMed: 8314763]

37. Nkari WK, Prestegard JH. NMR Resonance Assignments of Sparsely Labeled Proteins: Amide Proton Exchange Correlations in Native and Denatured States. Journal of the American Chemical Society 2009;131:5344–5349. [PubMed: 19317468]

38. Seidel RD, Zhuang TD, Prestegard JH. Bound-state residual dipolar couplings for rapidly exchanging ligands of His-tagged proteins. J Am Chem Soc 2007;129:4834–4839. [PubMed: 17385862]

39. Zhuang TD, Leffler H, Prestegard JH. Enhancement of bound-state residual dipolar couplings: Conformational analysis of lactose bound to Galectin-3. Protein Sci 2006;15:1780–1790. [PubMed: 16751604]

40. Zhuang TD, Lee HS, Imperiali B, Prestegard JH. Structure determination of a Galectin-3-carbohydrate complex using paramagnetism-based NMR constraints. Protein Sci 2008;17:1220–1231. [PubMed: 18413860]

Fig. 1.
Structures of the spin-labeled substrates of ST6Gal-I. CMP-4carboxyTEMPO is a donor mimic while LacNAc-TEMPO is an acceptor.
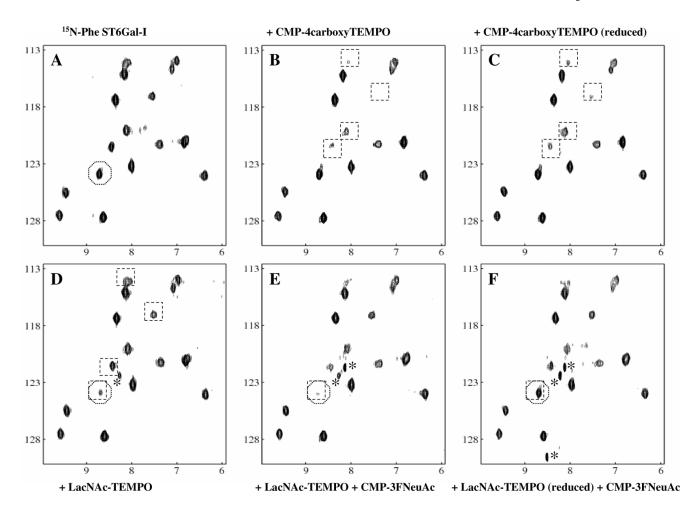
Fig. 2.
^{1}H-^{15}N HSQC spectra of ^{15}N-Phe ST6Gal-I in the presence of the indicated substrates.
(A) ^{15}N-Phe ST6Gal-I in the absence of substrates. (B and C) ^{15}N-Phe ST6Gal-I with a 4-fold molar excess of CMP-4carboxyTEMPO (B) before and (C) after reduction using ascorbate. (D, E and F) ^{15}N-Phe ST6Gal-I with a 20-fold excess of LacNAc-TEMPO (D), and added 15-fold excess of CMP-3FNeuAc (E), and added ascorbate (F). The boxed peaks were broadened due the presence of spin-labeled substrates. The circled peak was most strongly perturbed and was used to estimate structural changes of ST6Gal-I upon the addition of CMP-3FNeuAc. The circled peak did not change upon the addition of solely CMP-3FNeuAc to ST6Gal-I (Fig. 2B in reference [2]). Additional peaks that appeared because of natural abundance ^{15}N-H in the NH-NeuAc moiety of CMP-3FNeuAc or the GlcNAc moiety of LacNAc-TEMPO are denoted with an asterisk.
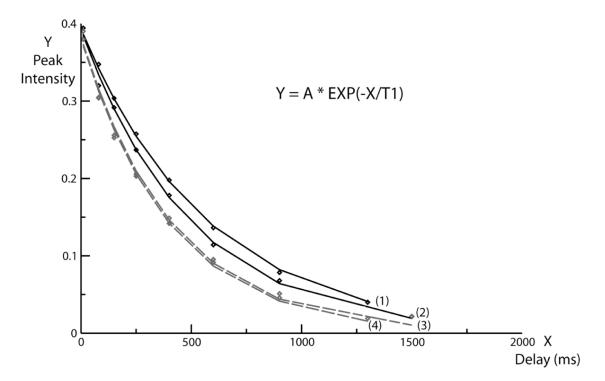
Fig. 3.
T_1 measurement examples for carbon Gal4 (solid black) and carbon Gal6 (dashed gray) with and without spin-label perturbation by CMP-4carboxyTEMPO binding to the protein. (1) data for Gal4 without a spin label. The T_1 was calculated as 574 ms, (2) data for Gal4 with a spin label. The T_1 was calculated as 496 ms, (3) data fo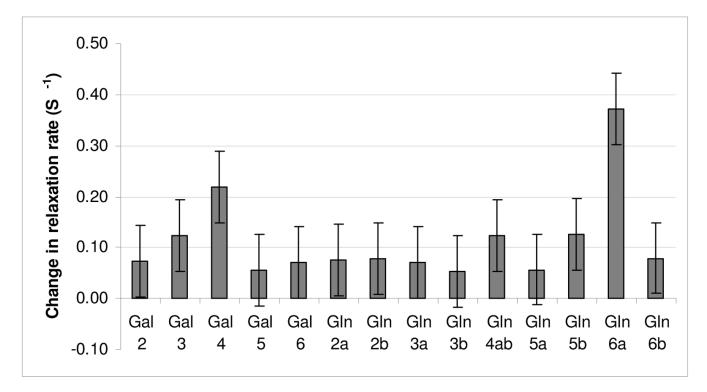r Gal6 with a spin label. The T_1 was calculated as 416 ms, (4) data for Gal6 without a spin label. The T_1 was calculated as 406 ms.

Fig. 4.
The differences in R_1 of LacNAc carbons in the control and experimental samples. All the changes are positive and they indicate paramagnetic perturbation from bound CMP-4carboxyTEMPO. The *N*-acetyl methyl carbons were not measured in the ^{1}H-^{13}C HSQC due to the lack of ^{13}C-enrichment in the methyl group. The data on Gal1 and GlcNAc1α and GlcNAc1β are not shown because their peak intensities were attenuated by water suppression during the acquisition of spectra.
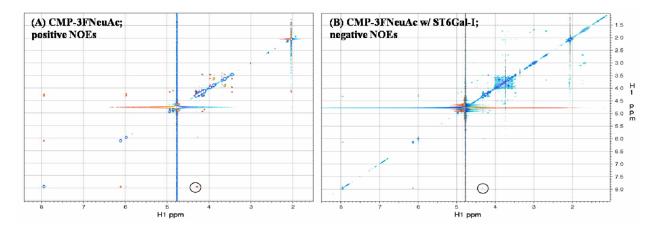
Fig. 5.
^{1}H-^{1}H 2D NOESY spectra of CMP-3FNeuAc (1.5 mM) without (A) and with (B) ST6Gal-I (0.3 mM) at a mixing time of 500 ms and at 5°C. Red indicates a positive NOE and blue indicates a negative NOE. No cross-peaks between the cytosine/ribose ring and the NeuAc ring were observed suggesting long distances. The circled cross-peaks indicate the NOEs connecting cytosine H6 and ribose H3 protons.
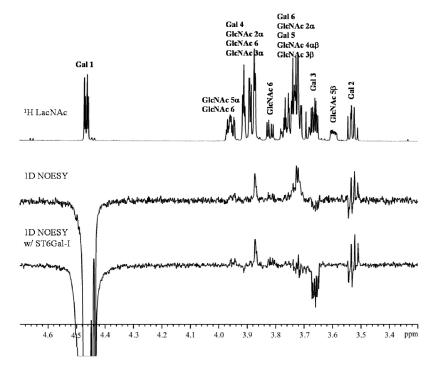
Fig. 6.
[1]H of LacNAc and 1D-NOESY spectra of LacNAc (4.5 mM) without and with ST6Gal-I (0.3 mM) acquired at 17°C.

Fig. 7.
The derived STD percentages for the protons of CMP-3FNeuAc (1.5 mM) bound to ST6Gal-I (0.3 mM) at 25°C. The double-headed arrows denote average STD effects for the protons that are unresolved in the STD difference spectrum.
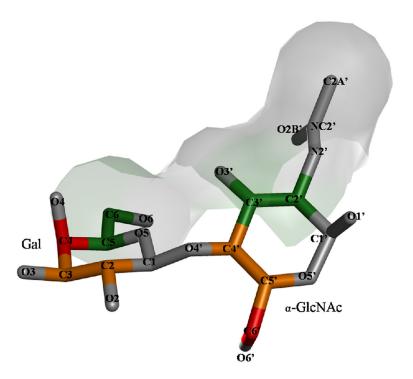
Fig. 8.
Position of α-LacNAc in the active site of ST6Gal-I. The color code indicates the measured proximity from the α-LacNAc carbons to the nitroxide of CMP-4carboxyTEMPO: red for closest (Gal4, GlcNAc6), yellow for second closest (Gal3, GlcNAc 4, GlcNAc5), and green for the furthest (Gal5, Gal6, GlcNAc2, GlcNAc3). The shaded area is proposed to be protected by the protein surface.
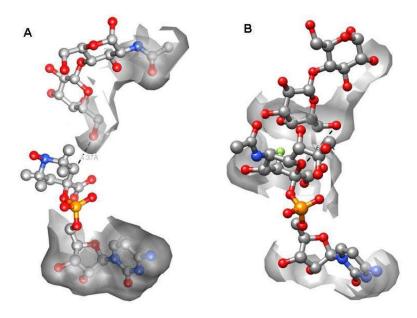
Fig. 9.
Comparison of relative positions and protein surface contacts of donor and acceptor. (A) CMP-4carboxyTEMPO and α-LacNAc in the ST6Gal-I binding site. (B) CMP-3FNeuAc and α-Lactose in the crystal structure of the 2,3/6 sialyltransferase from Δ24PmST1. The dashed lines depict attack on the anomeric carbons of the donor by the O6 oxygen of the acceptor. The gray regions depict protein contact regions; identified by STD or nitroxide exposure in ST6Gal-I and by 3.2Å proximity to protein atoms in Δ24PmST1.