BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

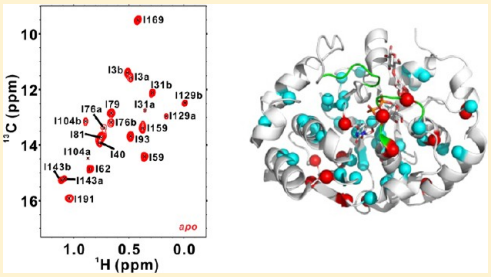
Nuclear Magnetic Resonance Spectral Assignments of α-1,4-Galactosyltransferase LgtC from *Neisseria meningitidis*: Substrate Binding and Multiple Conformational States

Patrick H. W. Chan,[†,‡] Sophie Weissbach,[†] Mark Okon,[†,§] Stephen G. Withers,[†,‡,§] and Lawrence P. McIntosh*[,†,‡,§,∥]

[†]Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, BC V6T 1Z3, Canada
[‡]Centre for High-throughput Biology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada
[§]Department of Chemistry, University of British Columbia, Vancouver, BC V6T 1Z1, Canada
[∥]Michael Smith Laboratories, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

Supporting Information

ABSTRACT: Lipopolysaccharide α-1,4-galactosyltransferase C (LgtC) from *Neisseria meningitidis* is responsible for a key step in lipooligosaccharide biosynthesis involving the transfer of α-galactose from the sugar donor UDP-galactose to a terminal acceptor lactose. Crystal structures of the complexes of LgtC with Mn^{2+} and the sugar donor analogue UDP-2-deoxy-2-fluorogalactose in the absence and presence of the sugar acceptor analogue 4′-deoxylactose provided key insights into the galactosyl-transfer mechanism. Combined with kinetic analyses, the enzymatic mechanism of LgtC appears to involve a "front-side attack" S_Ni-like mechanism with a short-lived oxocarbenium−phosphate ion pair intermediate. As a prerequisite for investigating the required roles of structural dynamics in this catalytic mechanism by nuclear magnetic resonance techniques, the transverse relaxation-optimized amide ^{15}N heteronuclear single-quantum correlation and methyl ^{13}C heteronuclear multiple-quantum correlation spectra of LgtC in its apo, substrate analogue, and product complexes were partially assigned. This was accomplished using a suite of complementary spectroscopic approaches, combined with selective isotopic labeling and mutagenesis of all the isoleucine residues in the protein. Only ∼70% of the amide signals could be detected, whereas more than the expected number of methyl signals were observed, indicating that LgtC adopts multiple interconverting conformational states. Chemical shift perturbation mapping provided insights into substrate and product binding, including the demonstration that the sugar donor analogue (UDP-2FGal) associates with LgtC only in the presence of a metal ion (Mg^{2+}). These spectral assignments provide the foundation for detailed studies of the conformational dynamics of LgtC.



Neisseria meningitidis is a Gram-negative bacterium that causes meningitis, an inflammation of the protective membranes covering the central nervous system.[1] The pathogenesis of *N. meningitidis*, with 10% fatality levels, is dependent upon the bacterium mimicking human lacto-*N*-neotetraose via its cell wall lipooligosaccharides (LOSs).[2] The LOSs are synthesized in part by glycosyltransferases (GTs) encoded by three genetic loci (*lgt-1*, *-2*, and *-3*), in which seven ORFs (*lgtcA*, *lgtB*, *lgtC*, *lgtD*, *lgtE*, *lgtH*, and *lgtZ*) are in a single *lgt-1* locus.[2] By studying the structures and catalytic mechanisms of these GTs, we can gain general insights into therapeutic approaches against *N. meningitidis*.

Lipopolysaccharide α-1,4-galactosyltransferase (LgtC) is responsible for the transfer of α-galactose from sugar donor UDP-Gal to the LOS terminal sugar acceptor lactose. LgtC belongs to GT family 8 and follows an ordered bi-bi reaction with binding of the donor preceding the acceptor.[3] The X-ray crystal structures of LgtC with Mn^{2+} and the UDP-2-deoxy-2-fluorogalactose (UDP-2FGal) sugar donor analogue in the absence (binary complex) and presence (ternary complex) of the acceptor sugar analogue 4′-deoxylactose have been determined.[4] The two structures are essentially identical and reveal that the monomeric enzyme consists of a large N-terminal mixed α/β domain, containing the active site, and a smaller C-terminal helical domain, which mediates membrane attachment. The full sugar acceptor binding site is proposed to be formed only upon the binding of a sugar donor in the active site and its concomitant burial by the ordering of two potentially flexible loop regions. Although these loops are thought to at least transiently adopt an opened conformational state to allow donor binding and product release,[4,5] there is no structural information for the substrate-free enzyme to support or refute this hypothesis.

On the basis of extensive kinetic analyses and computational calculations, LgtC appears to exploit an S_Ni-like mechanism to catalyze the transfer of galactose to the LOS acceptor with

Received: July 30, 2012
Revised: September 17, 2012
Published: September 19, 2012

   dx.doi.org/10.1021/bi3010279 | *Biochemistry* 2012, 51, 8278−8292

retention of its anomeric stereochemistry.[3,6−9] This "front-side attack" mechanism would require at least localized conformational and electrostatic changes to stabilize a short-lived oxocarbenium−phosphate ion pair intermediate. Hence, the dynamic properties of LgtC and its catalytic mechanism could be highly correlated, at the level of both substrate binding and glycosyl transfer. To provide experimental evidence of this hypothesis, we have used NMR spectroscopy to investigate LgtC along its reaction pathway from the free enzyme to binary and ternary complexes with substrate analogues and finally to the UDP product complex.

NMR spectroscopy is arguably the most informative experimental method for characterizing protein dynamics.[10] However, almost all known structural models of GTs have been obtained by X-ray crystallography. The only two NMR spectroscopically derived structures of partial GTs reported to date are the glycosyl donor binding domain of the bipartite glycosyltransferase (Alg13) and the transmembrane subunit of the oligosaccharyltransferase (OST) complex.[11,12] In both cases, the catalytic domain was absent. There are several reasons why it is difficult to investigate GTs by NMR spectroscopy. These enzymes are typically "large" (i.e., >30 kDa and frequently oligomeric) and thus yield complex spectra with broad, overlapping signals. Furthermore, GTs are often membrane-bound and thus have limited solubility, even when their membrane-associating regions are deleted. GTs also appear to exhibit a range of motions, especially in the active site, that can potentially lead to conformational exchange broadening of NMR signals.[8] Indeed, the crystal structures of these enzymes are usually determined with substrate or product analogues bound. In the apo forms of GTs, the electron density of active site loop regions is often missing, indicating significant flexibility. This is also a plausible reason why the apo forms of GTs are often difficult to crystallize. Fortunately, many new NMR spectroscopic methods and instrument technologies have been developed for studying large proteins, such as selective deuteration and [^{13}C]methyl labeling combined with transverse relaxation-optimized spectroscopy (TROSY) experiments.[13] These approaches complement X-ray crystallography and provide a window into understanding the roles of dynamics in protein function and enzymatic catalysis.

The requisite first step toward the structure and dynamics of LgtC by NMR spectroscopy is to assign the signals from its apo and substrate analogue, product-bound complexes. In this work, we report the partial spectral assignments of the amide ^{15}N-TROSY-HSQC and methyl-TROSY spectra of LgtC using a battery of NMR spectroscopic approaches, combined with mutagenesis of all the isoleucine residues in the protein. An analysis of these assignments provides insights into substrate and product complex formation, including the direct demonstration of UDP-2FGal binding only in the presence of Mg^{2+} and evidence of multiple conformations of LgtC.

MATERIALS AND METHODS

Cloning and Site-Directed Mutagenesis of LgtC. The previously described gene encoding LgtC, with the C128S and C174S mutations and a deletion of the C-terminal 25-residue membrane association sequence, was used as the starting point for this project.[4] Using directed evolution, the additional T273A mutation and two other silent mutations (GTC → GTT for Val133 and CAG → CAA for Gln189) were found to help increase the level of expression of LgtC without affecting its enzymatic activity (R. Kittl, personal communication). The reason for this behavior is unclear. Versions of this mutant were tested with combinations of N- and C-terminal His_6 tags and thrombin and TEV protease cleavage sites.[14] Maximal expression, purification, and proteolytic processing were obtained using the truncated LgtC (C128S/C174S/T273A) followed by a C-terminal TEV protease site and a His_6 tag. The final 292-residue construct, including the C-terminal hexapeptide Glu-Asn-Leu-Tyr-Phe-Gln remaining after tag cleavage, is henceforth denoted as LgtC. Additional mutations were introduced sequentially using the QuikChange site-directed mutagenesis procedure (Stratagene).

Protein Expression and Purification. The unlabeled His_6-tagged LgtC proteins were expressed in *Escherichia coli* BL21(λDE3) cells. The cells were grown in 2×YT broth medium at 37 °C to an OD_{600} of 0.8 and then induced with IPTG at a final concentration of 0.5 mM. After further growth at 16 °C for 16 h, the cells were harvested by centrifugation and lysed by sonication in the presence of 50 μg/mL lysozyme (Sigma). The cell debris was removed by centrifugation at 15000 rpm in a Sorvall SS32 rotor, and LgtC was isolated from the supernatant using a HisTrap HP column (GE Healthcare). The His_6 tag of the purified recombinant protein was then cleaved by His_6-tagged TEV protease (8 nM) during a 16 h dialysis against 50 mM Tris, 0.5 mM EDTA, and 1 mM TCEP (pH 8.0) at room temperature. The plasmid encoding the TEV protease construct was provided by the Structural Genomics Consortium, and the enzyme was expressed and purified according to standard protocols. After TEV protease digestion, the HisTrap HP column was used to remove the His_6 tag and uncleaved LgtC, as well as the tagged protease.

^{15}N-labeled proteins were expressed in M9 minimal medium containing 1 g/L $^{15}NH_4Cl$ and purified as described above. Proteins selectively labeled with [^{15}N]alanine, [^{15}N]leucine, [^{15}N]tyrosine, [^{15}N]valine, [^{15}N]glutamate, or [^{15}N]aspartate were produced using auxotrophic strain BL21(λDE3) *ilvE*, *tyrB*, *aspC*, *avtA*, and *trpB*, grown in a medium containing a mixture of unlabeled amino acids along with one of the following labeled amino acids: 100 mg/L [^{15}N]-L-alanine, 100 mg/L [^{15}N]-L-leucine, 50 mg/L [^{15}N]-L-tyrosine, 50 mg/L [^{15}N]-L-valine, 500 mg/L [^{15}N]-L-glutamate, or 200 mg/L [^{15}N]-L-aspartate (Sigma-Aldrich).[15,16]

Proteins selectively labeled with [$^{13}CH_3$]Ile$^{\delta1}$, [$^{13}CH_3$,$^{12}CD_3$]-Leu, and [$^{13}CH_3$,$^{12}CD_3$]Val in an otherwise deuterated background were expressed essentially according to published protocols.[17] The salts, antibiotics, and IPTG were dissolved in 99% D_2O and lyophilized to remove the exchangeable protium atoms. The cells were first grown in 2×YT/H_2O broth at 37 °C to an OD_{600} of 0.8 and then pelleted and resuspended in standard M9/H_2O medium to an OD_{600} of 0.2. Next, the cells were grown at 37 °C to an OD_{600} of 0.7 and then pelleted and resuspended in standard M9/D_2O medium containing 1 g/L $^{15}ND_4Cl$ and 3 g/L [D_7]-D-glucose in 99% D_2O (Cambridge Isotope Laboratories) to an OD_{600} of 0.1. The D_2O culture was grown until OD_{600} reached 0.5, diluted with the latter medium to an OD_{600} of 0.25, grown to an OD_{600} of 0.5, and again diluted to an OD_{600} of 0.1. After further growth to an OD_{600} of 0.25, 70 mg/L 2-keto-3-d_2-4-[^{13}C]butyrate and 120 mg/L of 2-keto-3-methyl-d_3-3-d_1-4-[^{13}C]butyrate (that is, α-ketoisovalerate deuterated at the β-position with one of the two methyl groups being $^{13}CH_3$ and the other being $^{12}CD_3$) were added (Cambridge Isotope Laboratories).[18−20] Precursor 2-keto-3-d_2-4-[^{13}C]butyrate was prepared from the 2-keto-3-H_2-4-[^{13}C]butyrate (Cambridge Isotope Laboratories) by incubation in 50 mM Na_2DPO_4 (pH* 10)

dx.doi.org/10.1021/bi3010279 | *Biochemistry* 2012, 51, 8278−8292

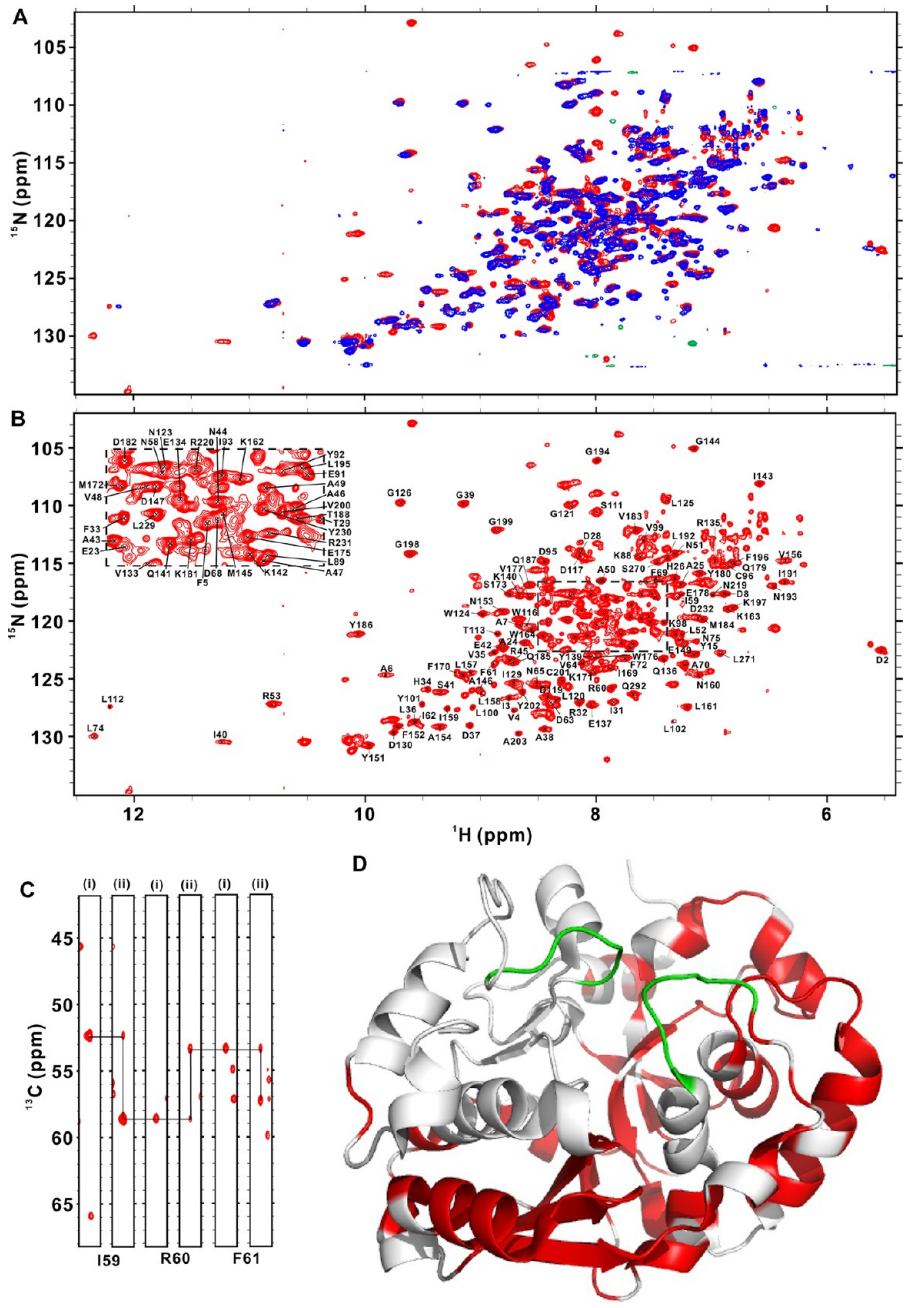


Figure 1. (A) Overlaid ^{15}N-TROSY-HSQC spectra of LgtC in the absence (red) and presence (blue) of 1 mM $MgCl_2$ and 1.5 mM UDP-2FGal show extensive perturbations caused by substrate analogue binding. Aliased peaks are colored green. (B) Partially assigned ^{15}N-TROSY-HSQC spectra of uniformly deuterated ^{13}C- and ^{15}N-labeled apo LgtC. The crowded central region is redrawn in the expanded, dashed box. (C) Strips from ^{15}N planes of three-dimensional TROSY-HN(CO)CA (i)/HNCA (ii) spectra show the backbone assignments of I59, R60, and F61 obtained by linking signals from $^{13}C^{\alpha}$ nuclei of residue $i-1$ (strip i) and both residues $i-1$ and i (strip ii). Similar patterns were observed for the $^{13}C'$ signal from complementary HN(CA)CO/HNCO spectra and the $^{13}C^{\beta}$ signals from HN(COCA)CB/HN(CA)CB spectra (not shown). (D) The assigned amides are mapped in red on the crystal structure of the LgtC ternary complex, and the active site loops are colored green. The active site residues could not be assigned because of their absence from the ^{15}N-TROSY-HSQC spectra.

in 99% D_2O at 45 °C for 16 h to exchange the β-protium atom with deuterium. D_2O-exchanged IPTG (final concentration of 0.5 mM) was added when the culture OD_{600} reached 0.8. The cells were then grown at 16 °C for 8 h postinduction before being harvested, and LgtC was purified as described above.

Protein selectively labeled with [$^{13}CH_3$]Ile$^{\delta1}$, [$^{13}CH_3$,$^{12}CD_3$]-Leu, and [$^{13}CH_3$,$^{12}CD_3$]Val in an otherwise deuterated and ^{13}C-labeled background was expressed using the same method as described above except 3 g/L D-[D_7]-[$^{13}C_6$]glucose and 70 mg/L methyl labeling precursors 2-keto-3-d_2-1,2,3,4-[^{13}C]butyrate and 120 mg/L 2-keto-3-methyl-d_3-3-d_1-1,2,3,4-[^{13}C]butyrate (Cambridge Isotope Laboratories) were used.[18−20]

Unless stated otherwise, all purified proteins were concentrated to ∼500 μM and exchanged into 20 mM Tris and 5 mM TCEP (pH 8.5) or 20 mM d_{11}-Tris and 5 mM TCEP (pH* 8.5) in D_2O, using an Amicon Ultra-15 centrifugal filter device. High concentrations of $MgCl_2$ (final concentration of 10 mM), UDP (final concentration of 1 mM), UDP-2FGal (final concentration

dx.doi.org/10.1021/bi3010279 | Biochemistry 2012, 51, 8278−8292

of 1 mM), and lactose (final concentration of 300 mM) relative to their K_m values were added to the protein to form binary and ternary complexes.[4,21] Protein concentrations were determined by UV absorbance, using the predicted ε_{280} of 66350 M^{-1} cm^{-1}.[22]

Activity Assays. The activities of LgtC and its mutants (1 μM) in a buffer of 50 mM HEPES and 10 mM $MnCl_2$ (pH 7.0) were determined qualitatively by a TLC assay with fluorescent detection using UDP-Gal (1 mM) and bodipy-lactose (0.5 mM) as substrates.[23] The TLC solvent consists of ethyl acetate, methanol, H_2O, and acetic acid in a 7:2:1:0.1 ratio. Steady-state kinetic parameters for LgtC were determined quantitatively by a continuous coupled enzyme assay in which the time-dependent formation of NAD^+ was monitored by absorption spectroscopy at 340 nm.[3,7,24] The assay utilized 0.1 μg/mL LgtC in a buffer consisting of 20 mM HEPES, 50 mM KCl, 15 mM $MnCl_2$, 0.5 mM NADH, 0.1% bovine serum albumin (BSA), 0.7 phosphoenolpyruvate (PEP), and 5 mM DTT (pH 7.5) with 2.7 units of lactate dehydrogenase (LDH) and 2 units of pyruvate kinase (PK) (Sigma-Alrich). Depending on which parameter was being measured, saturated amounts of either UDP-Gal (3 mM) or lactose (160 mM) were also included, and the concentration of the other substrate was varied.

NMR Spectroscopy. NMR spectra were recorded at 25 °C on Varian Inova 600 MHz and Bruker Avance III 600 and 850 MHz spectrometers equipped with cryogenic probes. Spectra were processed using NMRPipe[25] and analyzed with SPARKY-3.[26] One-bond sensitivity-enhanced ^{15}N-HSQC[27] and ^{15}N-TROSY-HSQC[28] spectra were recorded for ^{15}N-labeled LgtC and its variants. The sequence-specific backbone assignments of ^{2}H-, ^{13}C-, and ^{15}N-labeled LgtC [500 μM protein, 20 mM Tris, and 5 mM TCEP (pH 8.5) in 10% D_2O] were achieved using three-dimensional $^{1}H^{N}$-detected TROSY-HNCA, TROSY-HN(CO)CA, TROSY-HN(CA)CB, TROSY-HN(COCA)CB, TROSY-HNCO, and TROSY-HN(CA)CO spectra.[29]

Methyl-TROSY experiments[18,19,30,31] were recorded with [$^{13}CH_3$]methyl-labeled deuterated LgtC [500 μM protein in 99% D_2O with 20 mM d_{11}-Tris and 5 mM TCEP (pH* 8.5)]. Partial assignments of methyl-protonated [Ile(δ_1-$^{13}CH_3$ only), Leu($^{13}CH_3$,$^{12}CD_3$), and Val($^{13}CH_3$,$^{12}CD_3$)] in otherwise U-^{2}H,^{13}C,^{15}N-labeled LgtC were obtained using Ile,Leu-(HM)CM(CGCBCA)NH, Val-(HM)CM(CBCA)NH, Ile,Leu-HMCM(CGCBCA)CO, Val-HMCM(CBCA)CO, and HMCM(CG)CBCA experiments.[30] In addition, a complete set of LgtC mutants with systematic isoleucine to alanine or valine substitutions were generated and selectively labeled with [$^{13}CH_3$]methyl with deuteration for analysis by methyl-TROSY spectroscopy.

UDP-2FGal Synthesis. Initial samples of UDP-2FGal were obtained via chemical synthesis.[3] Subsequently, this donor analogue was generated enzymatically from chemically synthesized 2-deoxy-2-fluorogalactose by sequential use of galactokinase and galactose-1-phosphate uridyltransferase (GalK/GalT),[14] as summarized in Figure S1 of the Supporting Information.

Structural Analyses and Molecular Graphics. LgtC structural figures were rendered with PyMol. The coordinates for the complexes of LgtC with Mn^{2+}·UDP-2FGal in the absence [Protein Data Bank (PDB) entry 1G9R] and presence of 4′-deoxylactose (PDB entry 1GA8) were used, with substrates shown only as appropriate.

RESULTS

Optimizations of Experimental Conditions. Although an initial ^{15}N-HSQC spectrum of LgtC was promising, only ~120 of 300 expected $^{1}H^{N}$–^{15}N signals were detected (not shown). This behavior is both consistent with the size of the LgtC construct (32 kDa, 292 residues) and suggestive of aggregation or conformational exchange broadening. Therefore, we conducted an extensive screen of conditions to increase the solubility and stability of LgtC for long-term analyses by NMR spectroscopy. As summarized in Table S1 of the Supporting Information, LgtC was least prone to aggregation under basic conditions at pH 8.0–9.0. An elevated ionic strength and additives, such as UDP, UDP-2FGal, and lactose, did not help significantly. Although Mn^{2+} is required for the activity of LgtC and has been suggested to stabilize its substrate-bound structure,[4] Mg^{2+} can replace this paramagnetic metal ion with an only 50% reduction in activity.[21] Unfortunately, Mg^{2+} promotes slow (~1–2 days) precipitation of the enzyme, and we were unable to find conditions to prevent this behavior. In the end, to also maximize cryoprobe sensitivity, a simple buffer of 20 mM Tris (pH 8.5) with 5 mM TCEP reducing agent was selected as optimal for all subsequent studies. Spectra were recorded at 25 °C as a compromise between long-term stability, which was best for LgtC at lower temperatures, and signal line width, which generally becomes sharper at higher temperatures.

We also examined conditions for exchanging the backbone amide deuterium atoms of LgtC produced in D_2O medium to protium atoms for $^{1}H^{N}$ detection. Despite exhaustive trials with several different methods, including thermal unfolding and refolding, on-column refolding of His_6 tag-immobilized LgtC, and dilution of GdnHCl-denatured LgtC into 96 different conditions (Table S2 and Figure S2 of the Supporting Information, modified from ref 32), active LgtC could not be recovered in any practical yield. Fortunately, because LgtC is stable and soluble under alkaline conditions, we found that the amides of the folded protein could be reprotonated simply by incubation in H_2O buffer at pH 8.5 and 25 °C for 1–2 days (Figure S3 of the Supporting Information).

^{15}N-TROSY-HSQC Spectra and Assignments. The ^{15}N-TROSY-HSQC spectrum of uniformly deuterated ^{15}N-labeled apo LgtC is shown in Figure 1B. Although improved greatly relative to the conventional ^{15}N-HSQC spectrum of the nondeuterated protein recorded at pH 7.5, the spectrum exhibited only ~210 of ~300 expected peaks. Upon formation of a binary complex with Mg^{2+} and UDP-2FGal, many amides showed chemical shift perturbations, indicative of a change in the structure and dynamics of LgtC (Figure 1A). Importantly, over the course of a titration experiment, only signals from the free or bound LgtC were detected. Thus, the binding of this donor analogue occurs in the slow-exchange regime on the chemical shift time scale (i.e., $k_{ex} < \Delta\omega$, where $\Delta\omega$ is the chemical shift difference between the free and bound states).

TROSY-based ^{1}H, ^{13}C, ^{15}N correlation experiments (Figure 1C) were undertaken to help assign as well as possible the amide ^{15}N and $^{1}H^{N}$ signals of LgtC.[29,33−35] Along with selective ^{15}N amino acid labeling (Figure S4 of the Supporting Information) and site-directed mutagenesis (Figure S5 of the Supporting Information), 146 of ~210 observable peaks in the ^{15}N-TROSY-HSQC spectrum of apo LgtC were assigned confidently (Table S3 of the Supporting Information). Unfortunately, this corresponds to only ~50% of the total number of residues in LgtC. Furthermore, when they are mapped onto the structure of

dx.doi.org/10.1021/bi3010279 | Biochemistry 2012, 51, 8278–8292

the protein (Figure 1D), it is apparent that the assigned amides are distal from the active site, whereas those within the active site region of LgtC remained essentially unassigned. It is unlikely that this is simply a result of spectral overlap, incomplete reprotonation, or rapid amide hydrogen exchange at pH 8.5, as none of these factors should be unusually pronounced for this particular region of the protein. Instead, the lack of assignments likely reflects line broadening caused by millisecond to microsecond time scale conformational dynamics of the active site of LgtC. Indeed, we speculate that such backbone dynamics are linked to catalysis and are currently investigating this hypothesis using NMR relaxation experiments.

Parenthetically, of the assigned amides, the ^{15}N signals of Ile40, Arg53, Leu74, and Leu112 are unusually downfield-shifted, whereas that of Asp2 is unusually upfield-shifted (Figure 1B). On the basis of the X-ray crystal structure of substrate analogue-bound LgtC, the amide protons of Ile40 and Arg53 are hydrogen bonded with the carboxylate oxygens of the side chains of Asp37 and Glu23, respectively. This could lead to their downfield signals. In contrast, the amide protons of Asp2 and Leu112 are close to the indole ring of Trp116 and imidazole ring of His26, respectively, and thus are likely perturbed by ring current effects. Interestingly, the amide proton of Leu74, in one of the active site loops, appears to be solvent-exposed and is neither involved in an intramolecular hydrogen bond nor adjacent to an aromatic ring. Thus, the origin of its unusually downfield-shifted signal is unclear and suggests that the conformation of Leu74 might differ in apo LgtC versus that determined for the LgtC complexes by X-ray crystallography.

Methyl-TROSY Spectra Reveal Multiple Conformations. Because the amide signals of active site residues could not be identified in ^{15}N-TROSY-HSQC spectra, we turned our attention to the methyl groups of LgtC. Figure 2A shows the methyl-TROSY spectrum of LgtC selectively labeled with [^{13}CH$_3$]Ile$^{\delta 1}$, [^{13}CH$_3$,^{12}CD$_3$]Leu, and [^{13}CH$_3$,^{12}CD$_3$]Val in an otherwise deuterated background. The signals are generally well dispersed, especially in the upfield ^{13}C region corresponding to the isoleucine residues. We expected to observe 15 peaks from the 15 isoleucine δ_1-methyls and 86 peaks from the 24 leucine and 19 valine methyls (Figure 2B). Somewhat surprisingly, approximately 21 and 95 peaks could be counted in the isoleucine and leucine/valine regions, respectively, of the methyl-TROSY spectrum of the apoenzyme. Most of these peaks were of comparable intensity, yet a subset were weaker. Furthermore, as demonstrated below, the relative intensities of these peaks change with mutation and with substrate−product binding and thus arise from multiple conformations of LgtC that are in slow exchange on the chemical shift time scale.

Methyl-TROSY Assignments by Scalar Correlations. Using a suite of methyl-directed multidimensional ^{1}H, ^{15}N, and ^{13}C scalar coupling experiments,[30] the signals from nine isoleucine, five leucine, and seven valine residues in apo LgtC were assigned (Figure 2A and Figure S6 of the Supporting Information). Most of these are distal from the active site of LgtC, which of course reflects the fact that main chain ^{1}H^{N}, ^{15}N, and ^{13}CO assignments are a necessary prerequisite for this approach. One exception is Val133, which is near the active site and interacts with lactose. Unfortunately, similar spectra could not be recorded for substrate-bound forms of LgtC as the protein aggregated in the presence of Mg^{2+} within the time period required to record the necessary data (~2 days per experiment). Furthermore, because UDP-2FGal·Mg^{2+} binds in the slow-exchange regime, it was not possible to use titration experiments

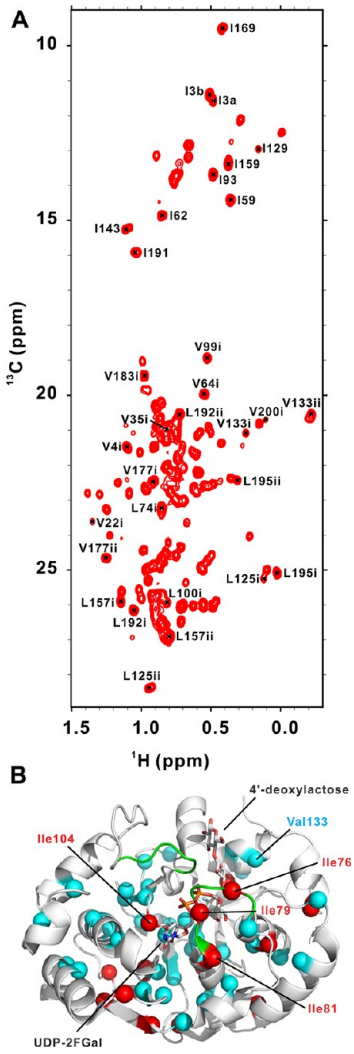


Figure 2. (A) Methyl-TROSY spectrum of apo LgtC. Partial assignments were obtained using magnetization-transfer experiments as summarized in Figure S6 of the Supporting Information. (B) Cartoon of the LgtC ternary complex with the isoleucine (red) and leucine/valine (cyan) residues identified by spheres. Selected residues are labeled; the proposed active site flexible loops are colored green, and UDP-2FGal and 4′-deoxylactose are shown as sticks (carbon, gray; oxygen, red; nitrogen, blue; phosphorus, orange).

to track chemical shift changes upon formation of the complex. To investigate the substrate binding and structural dynamics of LgtC, the spectral assignments of the remaining methyl groups had to be obtained using a site-directed mutagenesis approach.

Methyl-TROSY Assignments via Alanine Mutagenesis. In an initial attempt to assign the methyl-TROSY signals of selected active site residues, Ile76, Ile79, Ile81, Ile104, Val106, and Val133 were mutated to alanine. These "IA" and "VA" mutants were folded as evidenced by at least 5% activity in kinetic measurements (Table 1 and Figure S7 of the Supporting Information), and by ^{15}N-TROSY-HSQC (Figure S5 of the Supporting Information) and methyl-TROSY spectra (Figure S8 of the Supporting Information). Hence, their structures should be similar to that of the wild-type enzyme. Not surprisingly, mutation of Ile79, which directly contacts the bound sugar donor, had the greatest detrimental effect on catalysis. Ideally, when this spectrum is compared to the spectrum of wild-type LgtC, one peak should be absent because of an isoleucine

dx.doi.org/10.1021/bi3010279 | Biochemistry 2012, 51, 8278−8292

Table 1. Steady-State Kinetic Parameters for the IA and VA Mutants of LgtC

	UDP-Gal[a]			lactose[b]		
	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (μM^{-1} s^{-1})	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})
WT LgtC	29	16	0.6	101	23	0.2
LgtC-I76A	35	5	0.1	59	6	0.1
LgtC-I79A	19	1	0.05	275	2	0.01
LgtC-I104A	18	3	0.2	19	3	0.2
LgtC-V106A	32	26	0.8	24	26	1.1

[a]Data for sugar donor UDP-Gal in the presence of 160 mM acceptor lactose at pH 7.5 and 25 °C. [b]Data for sugar acceptor lactose in the presence of 10 mM donor UDP-Gal at pH 7.5 and 25 °C.

mutation and two because of a valine mutation. However, multiple spectral changes were observed upon mutation with both the apoenzymes and, to a lesser extent, their UDP-2FGal·Mg^{2+} binary complexes (Figures S8 and S9 of the Supporting Information). This made it difficult to confidently identify which signal was indeed absent and which were shifted due to conformational perturbations resulting from the mutation. In the end, only the two methyl signals from Val133 in apo LgtC, identified previously from scalar correlation spectra, could be confirmed. Because of improved spectral dispersion with UDP-2FGal·Mg^{2+} bound, the signals from Ile76, Ile81, Ile104, and Val133 in the binary complex could be assigned on the basis of alanine mutagenesis (Figure S9 of the Supporting Information).

Methyl-TROSY Assignments of Apo LgtC via Isoleucine to Valine Mutagenesis. Recently, Amero et al.[36] successfully assigned signals from the Ala and Ile δ_1-methyl groups of the 468 kDa multimeric aminopeptidase PhTET2 by generating a complete set of single mutants with every isoleucine and alanine in the protein changed to leucine and valine, respectively. By considering the patterns of methyl-TROSY spectral perturbations for the entire set, these researchers were able to confidently distinguish absent versus shifted peaks. Inspired by this approach, we generated a set of LgtC mutants in which each isoleucine has been individually substituted with a valine. Note that valine was chosen, rather than alanine as we used initially or leucine as used by others,[36] because this is also a hydrophobic β-branched residue that differs from isoleucine by the loss of only one CH_2 group. In principle, this should introduce fewer new unfavorable contacts than the number that would occur with a leucine substitution, while retaining most van der Waals contacts that might otherwise be lost with an alanine substitution. Thus, consistent with the BLOSUM matrices for amino acid substitutions,[37] the isoleucine to valine mutations are most likely to have minimal effects on the structure and NMR spectra of a protein.

For unclear reasons, two of the IV mutants, LgtC-I3V and LgtC-I159V, could not be produced with our *E. coli* expression system. Fortunately, the δ_1-methyl signals from these isoleucines had already been assigned via scalar correlations. The activities of the 13 remaining IV mutants were tested qualitatively by TLC assays (Figure S7 of the Supporting Information). Nine were wild-type-like, whereas the I76V and I79V mutations within an active site loop and I31V and I40V mutations that are distal from the active site had reduced activity.

Via comparison of individual methyl-TROSY spectra of the apo IV mutants with that of apo wild-type LgtC, the signals from all isoleucine δ_1-methyl groups were assigned (Figure 3). The assignments of Ile59, Ile62, Ile93, Ile129, Ile143, Ile169, and Ile191 previously obtained via scalar correlation experiments were also confirmed. In a few cases, such as LgtC-I169V, the only significant spectral change was the loss of one peak, thus providing an immediate identification of the Ile169 signal. However, in most cases, additional spectral perturbations were observed, and thus, comparisons of spectra from several mutants were required to deduce a consistent set of assignments. Fortunately, the perturbations resulting from the valine substitutions in LgtC-I76V, LgtC-I79V, and LgtC-I104V (Figure 3) were indeed smaller than those from the corresponding alanine mutations (Figure S8 of the Supporting Information), thus facilitating this approach. Some, but not all, of these perturbations can be rationalized by the proximity of interacting residues in the crystal structure of the enzyme. For example, substitution of Ile81 influences the chemical shifts of Ile76 and Ile79 in the adjacent active site loop (Figure 3O). Similarly, the δ_1-methyls of Ile62 and Ile93 are in van der Waals contact (Figure 3P), and indeed, the chemical shift of one changes upon the mutation of the other, albeit by different amounts. However, Ile159 is also near Ile93 and within the same β-strand as Ile62, yet neither pair appears to be coupled. More perplexing are the nonreciprocal changes in the signals of Ile31 and Ile104 caused by the mutation of the distal Ile129 (Figure 3J). This reflects the difficulty in rationalizing chemical shift perturbations on the basis of simple expectations from a static crystal structure. Furthermore, the available structures are of the LgtC binary and ternary complexes, which may differ from that of apo LgtC.

In the completely assigned methyl-TROSY spectrum of apo LgtC, Ile3, Ile31, Ile76, Ile104, Ile129, and Ile143 each yielded two peaks, with the weaker denoted as "a" and the more intense as "b" (Figure 3N). The ratios of the "a" to "b" peak intensities varied between residues, which could result from differential relaxation or from sensitivity to multiple conformational states of LgtC with differing populations. Furthermore, the relative intensities of the "a" versus "b" peaks from each of these isoleucines generally increased with the various IV mutants relative to the wild-type enzyme. This observation, which suggests that the mutants favor the conformations leading to the "a" peaks, will be discussed further below.

Methyl-TROSY Assignments of the UDP-2FGal·Mg^{2+} Binary Complex via Isoleucine to Valine Mutagenesis. The signals from the isoleucine δ_1-methyl groups in the LgtC binary complex were also assigned by comparing the methyl-TROSY spectra of the 13 IV mutants saturated with 10 mM Mg^{2+} and 1 mM UDP-2FGal (Figure 4). As with the apoenzyme samples, the IV mutations generally caused spectral perturbations, some of which can be rationalized on the basis of the known structure of the complex. Fortunately, the spectra of the binary complexes of the IV (and initial "IA") mutants showed better peak dispersion than those of the corresponding apo samples, making it easier to discriminate between absent and shifted signals. This was particularly true for the isoleucines closest to the active site because these residues are most influenced by substrate analogue binding.

A comparison of the spectra of LgtC in its apo versus binary complex forms is provided in Figures 5 and 6. When UDP-2FGal·Mg^{2+} binds, the δ_1-methyl signals of Ile76, Ile79, Ile81, Ile104, and Ile191 were the most substantially perturbed. As expected, these isoleucines generally clustered near the donor binding site and thus may experience chemical shift changes because of conformational changes because of the sugar donor binding and/or because of electrostatic and ring current effects

dx.doi.org/10.1021/bi3010279 | *Biochemistry* 2012, 51, 8278−8292

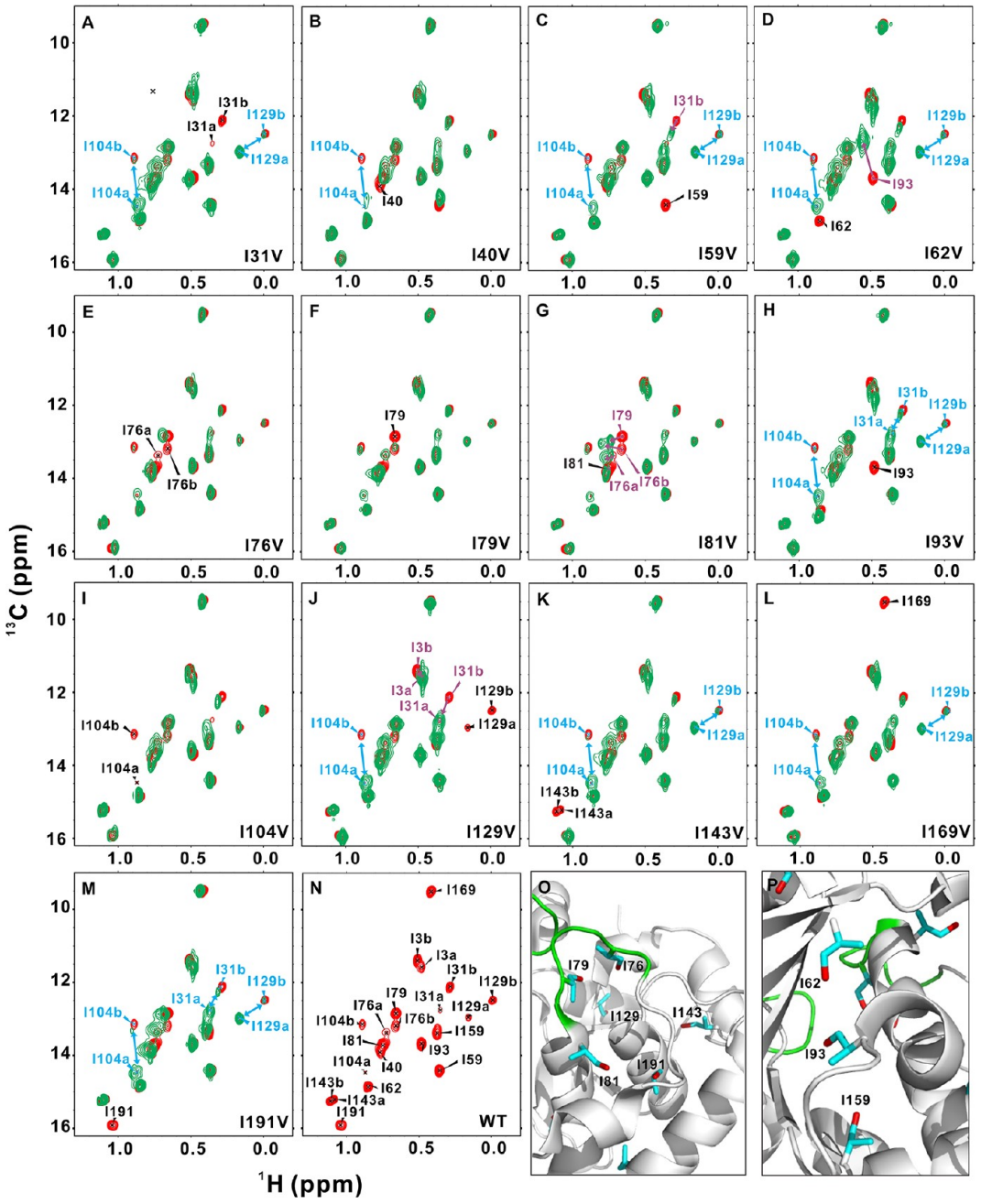


Figure 3. Assignment of the isoleucine δ_1-methyl signals of apo LgtC. Methyl-TROSY spectra of uniformly deuterated and selectively [^{1}H,^{13}C]methyl-labeled apo (A) LgtC-I31V, (B) LgtC-I40V, (C) LgtC-I59V, (D) LgtC-I62V, (E) LgtC-I76V, (F) LgtC-I79V, (G) LgtC-I81V, (H) LgtC-I93V, (I) LgtC-I104V, (J) LgtC-I129V, (K) LgtC-I143V, (L) LgtC-I169V, and (M) LgtC-I191V (green) overlaid with that of wild-type LgtC (red). Assigned peaks from the mutated residues are labeled in black. The peaks that are most significantly shifted in the mutant spectra are labeled in purple, whereas those with changes in the relative intensities of their "a" and "b" peaks are labeled in cyan. (N) Assigned methyl-TROSY spectrum of apo wild-type LgtC. (O and P) Expanded regions of LgtC showing isoleucines (cyan; C$^{\delta 1}$-methyl, red), as discussed in the text.

from the metal ion and sugar donor. In the cases of Ile81 and Ile191, which are more distal from the active site, the chemical shift perturbations upon substrate binding might result from structural changes of adjacent aromatic residues and hence altered ring current effects. It is also noteworthy that the introduction of mutations I40V and I79V both altered the chemical shift of Ile104 and caused visible line broadening (Figure 4B,F). The side chains of Ile104 and Ile79 are approximately 10 Å apart and are located on the opposite sides of the bound UDP-2FGal (Figure 4O). Mutation of Ile79 to a valine may alter the interaction of the donor analogue with Ile104, both structurally and dynamically, thereby yielding the observed spectral changes. In contrast, Ile40 is approximately 8 Å from Ile104 yet more distant from the active site of LgtC (Figure 4O). Mutation of this residue might alter packing within the hydrophobic core of the enzyme and thus indirectly perturb the interaction of Ile104 with the bound UDP-2FGal. In the spectrum of LgtC-I104V, the signals of Ile40 and Ile79 are not significantly changed (Figure 4I). However, chemical shifts depend upon the environment of each residue within the protein, and there is no reason to expect reciprocal spectral perturbations because of the mutations of these three isoleucine residues.

dx.doi.org/10.1021/bi3010279 | *Biochemistry* 2012, 51, 8278−8292

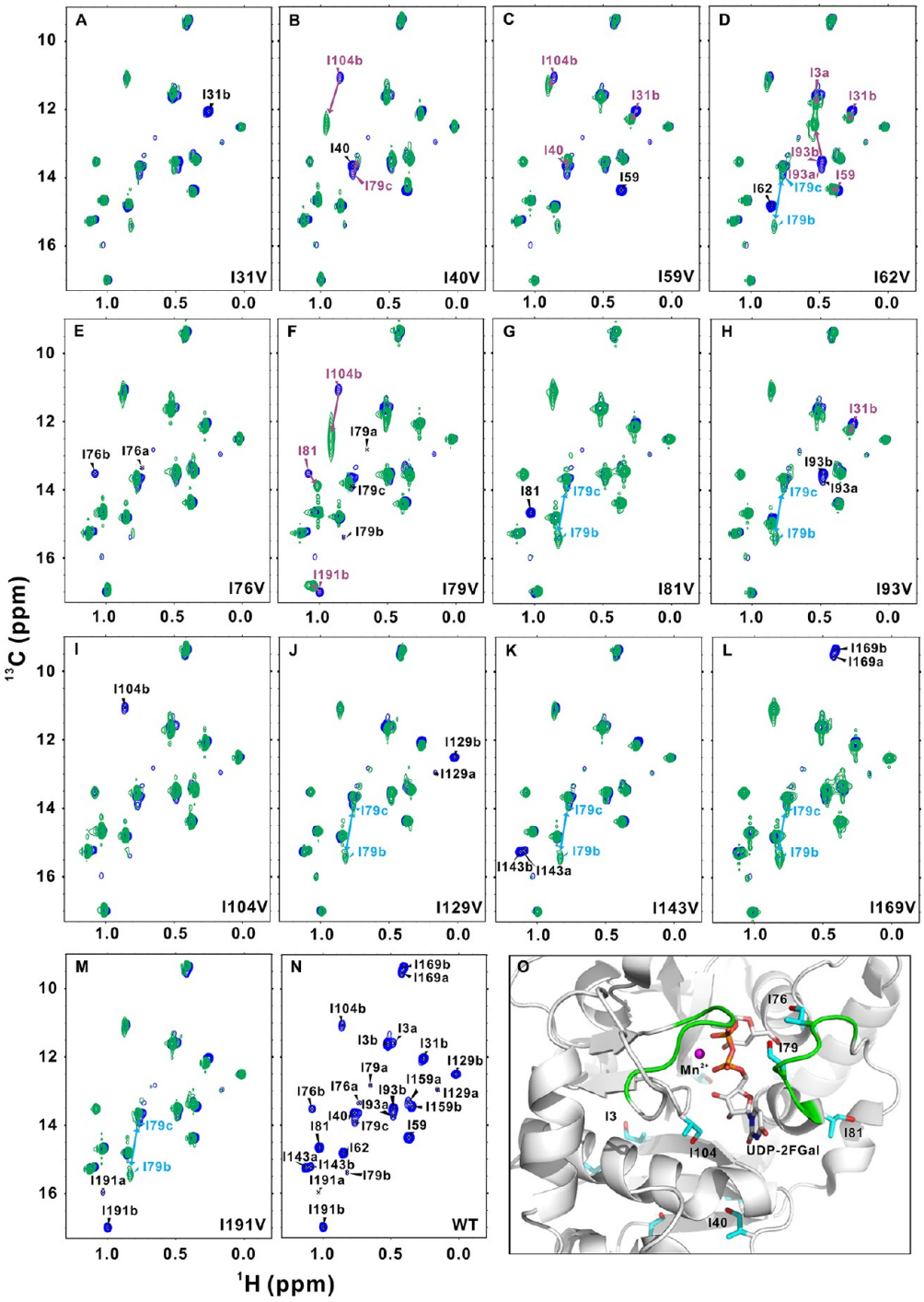


Figure 4. Assignment of the isoleucine δ_1-methyl signals in the UDP-2FGal·Mg^{2+} binary complex of LgtC. Methyl-TROSY spectra of the uniformly deuterated and selectively [^{1}H,^{13}C]methyl-labeled LgtC IV mutants (A) LgtC-I31V, (B) LgtC-I40V, (C) LgtC-I59V, (D) LgtC-I62V, (E) LgtC-I76V, (F) LgtC-I79V, (G) LgtC-I81V, (H) LgtC-I93V, (I) LgtC-I104V, (J) LgtC-I129V, (K) LgtC-I143V, (L) LgtC-I169V, and (M) LgtC-I191V (green) overlaid with that of wild-type LgtC (blue). All samples were saturated with 10 mM Mg^{2+} and 1 mM UDP-2FGal. Assigned peaks from the mutated residues are labeled in black. The peaks that are most significantly shifted in the mutant spectra are labeled in purple, whereas those with changes in the relative intensities of their "a" and "b" peaks are labeled in cyan. The "c" peak of Ile79 represents a possible third state of LgtC. (N) Overall assigned methyl-TROSY spectrum of the isoleucine residues of wild-type LgtC binary complex. (O) Crystal structure of the LgtC binary complex showing the active site isoleucine residues (cyan; $C^{\delta 1}$-methyl, red) and UDP-2FGal (carbon, gray; oxygen, red; nitrogen, blue; phosphorus, orange; Mn^{2+}, magenta).

When UDP-2FGal·Mg^{2+} binds, the magnitudes of the δ_1-methyl signals from the isoleucine "a" peaks decreased in relative intensity, whereas those from the "b" peaks both increased in intensity and shifted in spectral position (Figures 5 and 6). This suggested that the sugar donor analogue bound only to the form of LgtC giving rise to the "b" peaks, and that the "a" peaks remained detectable because of incomplete saturation with the donor analogue or perhaps a small population of inactive enzyme with an apo-like spectrum. However, the methyl signals of Ile93, Ile159, and Ile169 also split slightly, indicating possible conformational heterogeneity in the binary complex. Furthermore, Ile79 in the active site loop yielded three peaks with

dx.doi.org/10.1021/bi3010279 | *Biochemistry* 2012, 51, 8278−8292

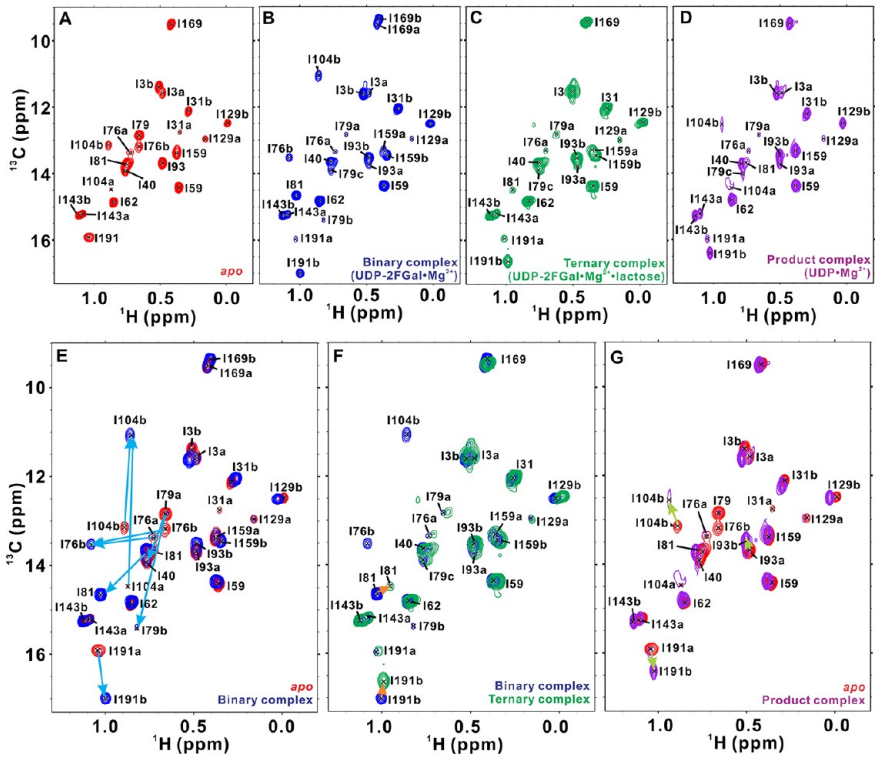


Figure 5. Assigned methyl-TROSY spectra of LgtC along its reaction pathway: (A) apo form, (B) binary complex with 10 mM Mg^{2+} and 1 mM UDP-2FGal, (C) ternary complex with 10 mM Mg^{2+}, 1 mM UDP-2FGal, and 300 mM lactose, and (D) product complex with 20 mM Mg^{2+} and 1 mM UDP. Also shown are the spectra of (E) the UDP-2FGal·Mg^{2+} binary complex overlaid on that of the apoenzyme, (F) the UDP-2FGal·Mg^{2+}·lactose ternary complex overlaid on that of the binary complex, and (G) the UDP·Mg^{2+} product complex overlaid on that of the apoenzyme. Arrows highlight spectral changes.

relative intensities that changed upon mutation (Figures 4 and 5). Thus, LgtC exhibits multistate conformational behavior.

Methyl-TROSY Assignments of the UDP-2FGal·Mg^{2+}·Lactose Ternary Complex. The addition of lactose to the preformed complex of LgtC with UDP-2FGal·Mg^{2+} led to only modest methyl-TROSY spectral changes. This allowed us to readily extend the assignments of the isoleucine δ_1-methyl signals from the binary to the resulting ternary complex by monitoring these changes as a function of added lactose (Figure 5). Most significantly, the signals from Ile76 and Ile104 were no longer observed, possibly because of conformational exchange broadening. Otherwise, only minor chemical shift perturbations were observed for Ile81 and Ile191 (Figure 6). The smaller spectral effects of lactose binding to form the ternary complex relative to UDP-2FGal·Mg^{2+} binding to form the binary complex are consistent with the nearly identical crystal structures of the two complexes, as well as the smaller interface of LgtC with the sugar acceptor, and its neutral, non-aromatic character.

Methyl-TROSY Assignments of the UDP·Mg^{2+} Product Complex. The addition of UDP·Mg^{2+} to apo LgtC also resulted in modest spectral changes. Thus, it was also straightforward to assign the isoleucine δ_1-methyl signals in the product complex by direct spectral comparisons (Figure 5). The smaller chemical shift perturbations resulting from UDP·Mg^{2+} binding relative to the larger effects of UDP-2FGal·Mg^{2+} suggest that the product complex is more similar structurally to apo LgtC than to the binary substrate complex (Figure 6).

The Sugar Donor Requires Mg^{2+} for Binding to LgtC. On the basis of crystallographic studies and molecular dynamics simulations,[4,5] it was suggested that LgtC binds Mn^{2+} prior to the sugar donor UDP-Gal. However, there was no clear experimental test of this binding order. Therefore, NMR spectroscopy was used to monitor the titration of LgtC with Mg^{2+} and the donor analogue UDP-2FGal (Figure 7). The addition of either 10 mM Mg^{2+} or 1 mM UDP-2FGal independently did not perturb the methyl-TROSY spectrum of methyl-labeled LgtC or the ^{15}N-TROSY-HSQC spectrum of [^{15}N]tyrosine-labeled LgtC. In contrast, when both Mg^{2+} and UDP-2FGal were present, signals from several methyls and tyrosine amides were significantly shifted. Therefore, within these concentration ranges, it is clear that UDP-2FGal binds LgtC only when Mg^{2+} (and, by inference, Mn^{2+}) is present. It is also likely that Mg^{2+} binds LgtC cooperatively with the sugar donor, rather than independently. However, the latter conclusion is less certain as it is based on the assumption the binding of Mg^{2+} would have at least altered the chemical shifts of Leu102, Ile104, Val156, and Tyr245, which are reporter groups in the proximity of the crystallographically defined catalytic metal binding site of LgtC. Unfortunately, the methyl signals of Leu102 and Val156 were not assigned, and the amide signal of Tyr245 was not detected. Thus, the insensitivity of the Ile104 methyl signal to the presence of Mg^{2+} provides the main evidence of the simultaneous binding of UDP-2FGal·Mg^{2+} to LgtC.

Multiple Conformations Probed by Mn^{2+} Binding. The binding of Mn^{2+} to the multiple conformational states of LgtC was investigated by monitoring the loss of isoleucine δ_1-methyl signal intensity in methyl-TROSY spectra because of paramagnetic relaxation. To avoid complications due to the spectral changes accompanying substrate binding, a preformed LgtC binary complex saturated with Mg^{2+} and UDP-2FGal was titrated with Mn^{2+}. Because the K_m value of Mn^{2+} (27 μM) is 13-fold

dx.doi.org/10.1021/bi3010279 | *Biochemistry* 2012, 51, 8278−8292

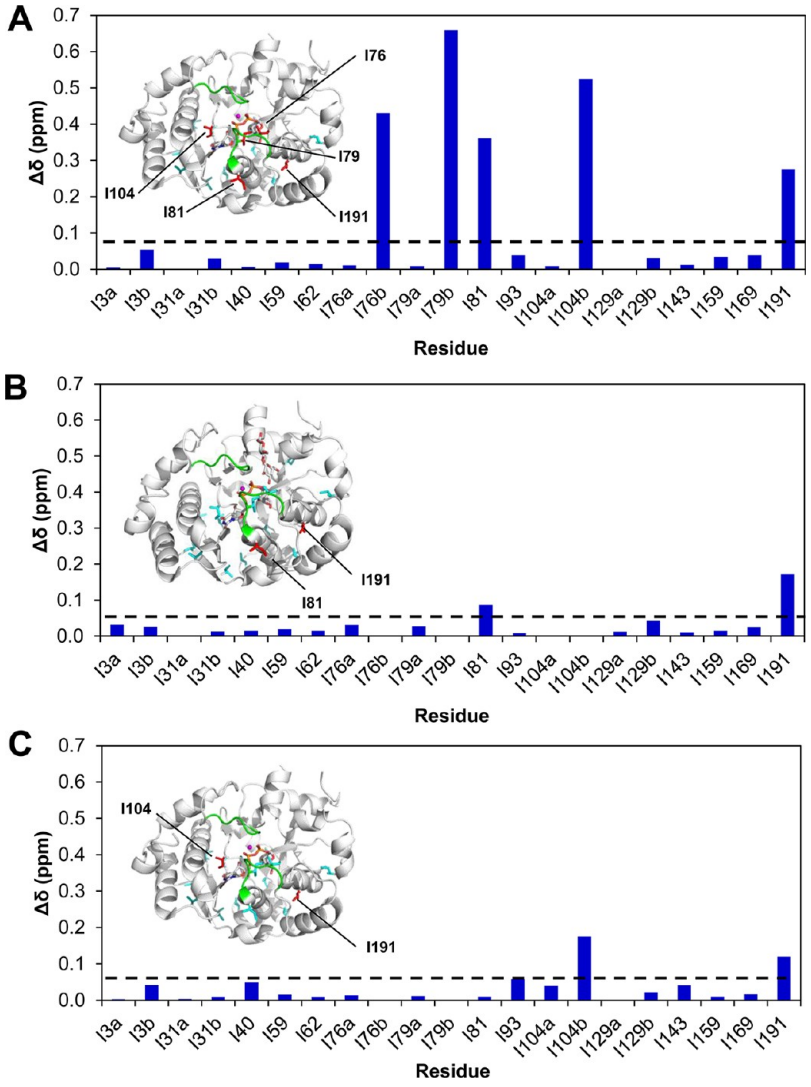


Figure 6. Isoleucine δ_1-methyl chemical shift perturbations (CSPs) along the LgtC reaction pathway: (A) UDP-2FGal·Mg^{2+} binary complex vs apo LgtC, (B) UDP-2FGal·Mg^{2+}·lactose ternary complex vs the binary complex, and (C) UDP·Mg^{2+} product complex vs apo LgtC, generated using the assigned spectra of Figure 5. Blank values correspond to residues not detected under one or more conditions. Isoleucines with a CSP of >0.08 ppm are highlighted in red on the ribbon diagrams and the remainder with smaller perturbations or unobserved signals in cyan (UDP-2FGal and 4′-deoxylactose shown with carbons colored gray, oxygens red, nitrogens blue, phosphoruses orange, Mn^{2+} ions magenta, and active site loops green).

lower than that of Mg^{2+} (370 μM),[21] we assumed that Mn^{2+} will displace Mg^{2+} and alter the intensities, but not chemical shifts, of nearby isoleucine methyl groups. As summarized in Figure 8, this indeed occurred. Qualitatively, the signals from residues closest to the metal binding sites, including Ile76, Ile79, and Ile104, disappeared at lower Mn^{2+} concentrations than those from more distal residues. However, a clear correlation of intensity loss with distance was not observed, possibly because of nonspecific effects from the relatively high concentrations of Mn^{2+} in the bulk solvent and differential nonparamagnetic relaxation rates of the methyl groups. Most strikingly, in the presence of 400 μM Mn^{2+}, only signals from "a" peaks were observed in methyl-TROSY spectra, albeit at a reduced intensity. This result also indicates that UDP-2FGal·Mn^{2+} binds only to the form of the protein giving rise to the "b" peaks, and that the residual "a" peaks originate from a population of protein not saturated with the sugar donor analogue.

DISCUSSION

Spectral Assignments. Using a combination of mutagenesis, isotope labeling strategies, and TROSY-based pulse sequences, we have made substantial progress toward assigning the NMR spectra of LgtC along its reaction pathway. Initially, we focused on its amide ^{15}N-TROSY-HSQC spectrum as this can provide a probe of each non-proline residue in the backbone of a protein. Although the assignments were facilitated by deuteration and amino acid selective ^{15}N labeling, we were able to detect only ~70% and assign only ~50% of the expected $^{1}H^{N}$–^{15}N signals from apo LgtC with TROSY-based ^{1}H, ^{13}C, ^{15}N correlation experiments. Unfortunately, these assigned signals corresponded to amides that are remote from the active site. A similar result occurred in a related study of *Campylobacter jejuni* sialyltransferase CstII[38] and is suggestive of active site backbone dynamics in glycosyltransferases on a millisecond to microsecond time scale that led to conformational exchange broadening. However, such motions were not substantially dampened in either protein upon addition of a sugar donor

dx.doi.org/10.1021/bi3010279 | *Biochemistry* 2012, 51, 8278–8292

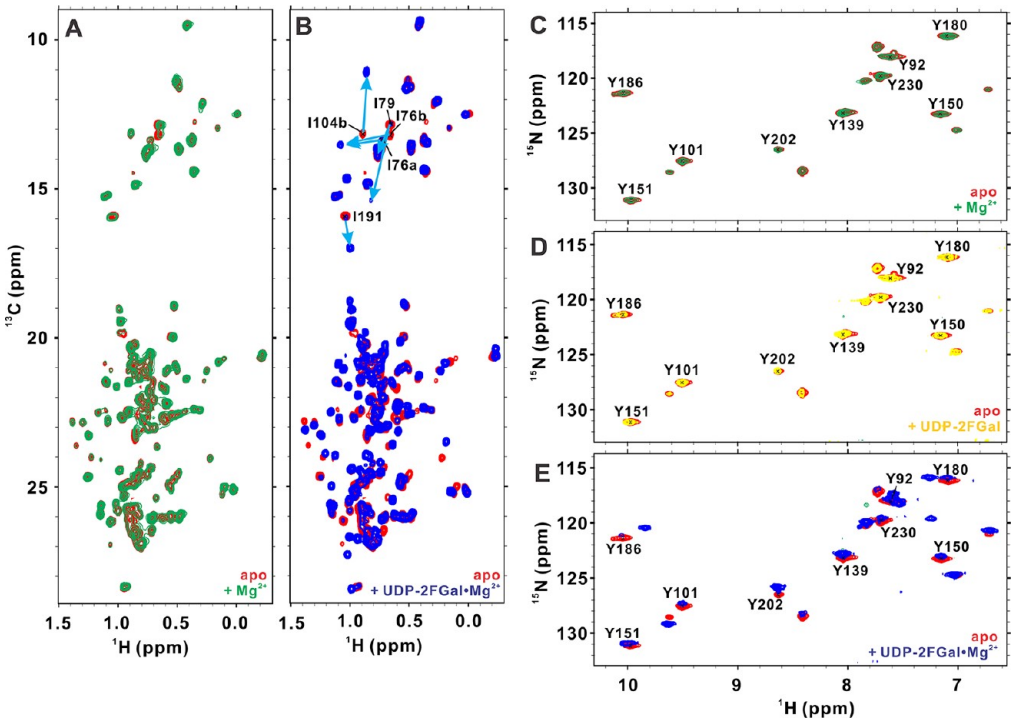


Figure 7. UDP-2FGal binding by LgtC requires Mg^{2+}. (A, C, and D) Overlaid methyl-TROSY and ^{15}N-TROSY-HSQC spectra of LgtC in the absence (red) and presence of 10 mM Mg^{2+} (green) or 1 mM UDP-2FGal (yellow) show no significant differences. Thus, UDP-2FGal alone does not bind LgtC. Assuming that metal binding should cause some detectable spectral perturbations, the same conclusion likely holds for Mg^{2+}. (B and E) The clear differences between the overlaid methyl-TROSY and ^{15}N-TROSY-HSQC spectra of LgtC in the absence (red) and presence (blue) of 10 mM Mg^{2+} and 1 mM UDP-2FGal confirm that the sugar donor binds as a metal complex.

analogue, which caused extensive chemical shift changes but not the appearance of the numerous missing amide signals.

In contrast to amide-directed approaches, the methyl-TROSY spectra of selectively $^{13}CH_3$-labeled LgtC in an otherwise deuterated background yielded excellent quality spectra. Because of the limited main chain assignments, combined with the aggregation of LgtC in the presence of Mg^{2+}, only a subset of these methyl signals could be assigned using multidimensional ^1H, ^{13}C, ^{15}N correlation experiments. Accordingly, a mutational approach was required. However, initial attempts to assign selected active site residues by alanine substitutions were hampered by chemical shift perturbations, which precluded the unambiguous identification of the one absent peak in the methyl-TROSY spectrum of each mutant. Thus, a more laborious approach of mutating each isoleucine to a valine was undertaken. Via comparison of the rapidly measured spectra from the entire set of mutants, it was possible to fully assign all of the isoleucine δ_1-methyl signals of LgtC in its apo form, as well as with substrate analogues and product UDP bound (Figure 5). As expected, the complicating spectral perturbations resulting from the valine mutations were generally smaller than those observed with the initial alanine substitutions. In some cases, the spectral changes could be rationalized, whereas in others, an obvious explanation was not readily apparent on the basis of simple expectations from the available static crystal structures. This suggests the latter may not completely reflect the solution conformation of LgtC at all locations and that mutations may have subtle, long-range effects on its structure and dynamics.

Kinetic Effects Caused by Single-Site Mutations. The effects of mutations on the activity of LgtC were examined either through quantitative kinetic measurements or via a qualitative fluorescence-based TLC assay. On the basis of a detailed kinetic analysis conducted with the initial set of alanine mutants (Table 1), substitution of Ile79 had the greatest detrimental effect on catalysis, reducing k_{cat}/K_m for UDP-Gal and lactose to ~5% of this wild-type level. This is likely due to the direct contact of its side chain with UDP-Gal. More modest effects were observed upon mutation of Ile76 and Ile104, both of which also contact the substrate. In contrast, mutation of Val106 to alanine increased k_{cat}/K_m toward lactose by ~5-fold for reasons that are not immediately obvious as its side chain is >7 Å from any substrate atom. Furthermore, from qualitative assays, the substitution of Ile40 and, more so, Ile31 impaired LgtC activity (Figure S7 of the Supporting Information). Both are quite remote from the active site of LgtC and appear to exert long-range effects on the structure or dynamics of the enzyme. However, such effects were not manifest via any unusual methyl-TROSY spectral perturbations, except for a pronounced change in the chemical shift of the active site Ile104 in the UDP-2FGal·Mg^{2+} complex of the LgtC-I40V mutant.

Donor Substrate Binding. Using both methyl-TROSY and ^{15}N-TROSY-HSQC measurements, we demonstrated unambiguously that UDP-2FGal binds LgtC only in the presence of Mg^{2+}. This is consistent with both the mandatory requirement of the enzymatic activity of LgtC for a divalent metal ion and the crystallographic structures of its complexes in which a Mn^{2+} bridges two phosphate oxygens of the UDP moiety with the side chains of Asp103, Asp105, and His244.[4] These two aspartates form the conserved DxD motif that is found in a wide range of glycosyltransferases.[4] In the absence of a metal ion, an unfavorable electrostatic interaction between this motif and the negatively charged sugar donor would certainly arise. The lack of spectral perturbations upon addition of Mg^{2+} to apo LgtC also suggests that the sugar donor and the metal ion bind the enzyme

dx.doi.org/10.1021/bi3010279 | Biochemistry 2012, 51, 8278−8292

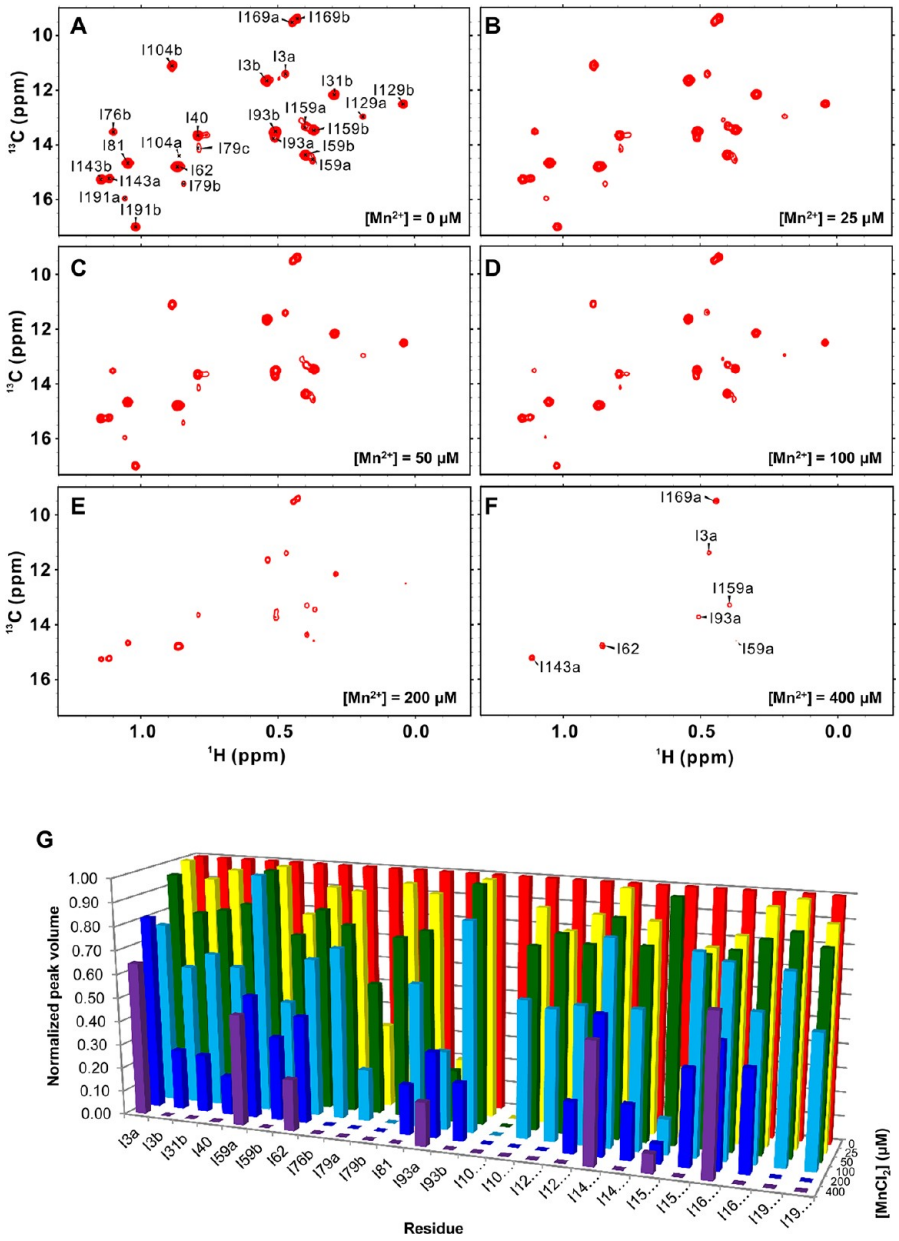


Figure 8. Analysis of Mn^{2+} binding by LgtC. (A) Methyl-TROSY spectrum of the LgtC binary complex (200 μM protein) in the presence of 10 mM $MgCl_2$ and 1 mM UDP-2FGal. The complex was titrated with (B) 25, (C) 50, (D) 100, (E) 200, and (F) 400 μM $MnCl_2$. (G) Histogram showing the relative intensity change of each isoleucine δ_1-methyl group with respect to increasing concentrations of Mn^{2+} due to paramagnetic relaxation.

cooperatively as a UDP-2FGal·Mg^{2+} complex. Indeed, it is well-established that Mg^{2+} interacts directly with nucleotides,[39,40] and previous kinetic measurements showed that LgtC has similar K_m values for UDP-Gal and Mn^{2+}.[4,21] However, it is possible that the chemical shifts of the methyl and tyrosine amide reporter groups are insensitive to the presence of bound Mg^{2+}. If so, then the metal ion could bind LgtC either independently before or simultaneously with UDP-Gal. Regardless of pathway, by thermodynamic linkage, the affinities of both the metal and sugar donor for the enzyme will be enhanced in the presence of each other.

NMR-monitored titration experiments also demonstrated that UDP-2FGal·Mg^{2+} binds LgtC in the slow-exchange limit such that distinct signals from the free and bound protein were observed. This indicates that the exchange rate ($k_{ex} = k_{on}$[UDP-2FGal·Mg^{2+}] + k_{off}) is smaller than the chemical shift difference for both amides and methyl groups in the apo versus binary complex states of LgtC. Extrapolating from previously reported kinetic data [metal ions Mg^{2+} (K_m = 370 μM with UDP-Gal) and Mn^{2+} (K_m = 27 μM with UDP-Gal); sugar donors UDP-Gal (K_m = 30 μM with Mn^{2+}) and UDP-2FGal (K_i = 2 μM with Mn^{2+})],[4,21] we think it is likely that UDP-2FGal·Mg^{2+} binds LgtC with a K_d value on the order of ~25 μM. Given that this modest affinity usually results in binding in the fast- to intermediate-exchange limit,[10] it appears that both the association (k_{on}) and dissociation (k_{off}) rate constants for the sugar donor analogue with LgtC are slower than those typically found for simple protein−ligand complexes with comparable K_d values. On the basis of the X-ray crystal structure of the LgtC binary and ternary complexes, a plausible explanation for this result is that the two loops flanking the active site loops must undergo a conforma-

dx.doi.org/10.1021/bi3010279 | Biochemistry 2012, 51, 8278−8292

tional change to allow access of the UDP-2FGal·Mg^{2+} to an otherwise buried position within the enzyme's active site.

Acceptor Lactose Binding. Relatively small spectral changes occurred when lactose was added to the binary complex of LgtC with UDP-2FGal·Mg^{2+}, indicating that acceptor binding does not substantially perturb the structure of the enzyme. This is consistent with the highly similar crystal structures of its binary and ternary substrate complexes. Surprisingly, however, in methyl-TROSY-monitored titrations, the "b" δ_1-methyl signals of the three isoleucines (Ile76, Ile79, and Ile104) that contact the bound UDP-2FGal all disappeared upon the addition of lactose. Such exchange broadening (i.e., interconversion between states with different chemical shifts at a rate $k_{ex} \sim \Delta\omega$) is unexpected given the very low affinity of lactose for the LgtC binary complex (K_m = 13 mM)[4] and hence predicted fast-exchange ($k_{ex} > \Delta\omega$) behavior. Possible explanations for this result are that the exchange rate for lactose is slowed to the millisecond to microsecond time scale typically associated with exchange broadening because of a requisite structural or dynamic change for binding or that, once bound, lactose induces conformational motions within the ternary complex on this time scale.

Product UDP Binding. As typically observed with glycosyltransferases, LgtC is subject to product inhibition by UDP (K_i = 80 μM).[41] When apo LgtC was titrated with UDP·Mg^{2+}, only minor isoleucine δ_1-methyl chemical shift perturbations were observed in the slow-exchange regime, with the largest effect seen for Ile104 and Ile191. It is noteworthy that these chemical shift perturbations were substantially smaller than those accompanying UDP-2FGal·Mg^{2+} binding, despite the fact that several isoleucines are expected to contact the UDP moiety in both cases. Furthermore, the signals of Ile76 and Ile79, which are within one of the active site loops, disappeared upon UDP·Mg^{2+} binding. This is suggestive of conformational exchange broadening. To exclude alternative possibilities, such as aggregation, UDP was converted "in situ" to UTP by the addition of pyruvate kinase and phosphoenolpyruvate (Figure S10 of the Supporting Information). Indeed, after this treatment, the methyl-TROSY spectrum of the LgtC sample returned to that observed with the apoenzyme as UTP no longer binds. Although a crystal structure of LgtC with UDP·Mg^{2+} has not been determined, these results indicate that conformational and dynamic differences exist between its substrate and product complexes.

Multiple Conformational States of LgtC. More peaks than expected were observed in the methyl-TROSY spectra of apo LgtC and its substrate and product complexes. This suggested two general hypotheses. First, the sample used in the NMR experiments could be heterogeneous, perhaps containing chemically modified forms of LgtC (e.g., oxidized or proteolyzed) or other unrelated *E. coli* contaminants. However, the fact that mutations and substrate or product binding alter the relative intensities of the "a" and "b" peaks argues against any predominant irreversible chemical modification. Nevertheless, although the LgtC samples appeared to be >90% pure via sodium dodecyl sulfate−polyacrylamide gel electrophoresis, a signal of ~10% relative intensity was observed in a MALDI-TOF mass spectrum with a mass ~130 Da greater than that measured for apo LgtC (Figure S11 of the Supporting Information). We speculated initially that this could be due to a galactose or glucose covalently linked with Asp190 in the active site during protein expression because a similar phenomenon had previously been observed with the Q189E mutant of LgtC.[7] However, despite many efforts to cleave the proposed saccharide from the active site of LgtC by the addition of UDP, Mn^{2+}, and lactose to complete the glycosyl-transfer reaction, the higher-molecular mass species remained in the sample. This minor species is thus assumed to be an unknown modification of the enzyme.

An alternative and more consistent hypothesis is that apo LgtC exists in at least two conformational states that interconvert slowly enough to yield distinct NMR signals, denoted as "a" and "b" peaks in a methyl-TROSY spectrum. Indeed, preliminary measurements have detected magnetization transfer between the corresponding "a" and "b" peaks of several methyl groups in LgtC with rate constants on the order of ~0.2 s^{-1}.[14] These states are in equilibrium when no substrate is present, and their relative populations vary with mutation. When the enzyme was titrated with UDP-2FGal·Mg^{2+} or UDP·Mg^{2+}, the intensities of the "a" peaks diminished, whereas the intensities of the "b" peaks both grew in intensity and, in several cases, changed in chemical shift. Furthermore, the "b", but not "a", peaks disappeared upon replacement of Mg^{2+} with paramagnetic Mn^{2+}. Collectively, these data suggest that LgtC exists in at least two interconverting conformations, of which only the form giving rise to "b" peaks is competent for substrate and product binding.

Although the differences between the putative active and inactive conformations of LgtC are currently unknown, one plausible albeit purely speculative explanation is a slow cis−trans isomerization of one of the two proline residues in an active site loop. It is likely that this loop transiently opens to allow binding and release of UDP-2FGal or UDP from an otherwise buried position in the active site of LgtC. However, the splitting of the signals from Ile159 and Ile169 and the appearance of peak "c" for Ile79 in the UDP-2FGal·Mg^{2+} binary complex, as well as the fact that different residues in apo LgtC yield different "a" versus "b" methyl-TROSY peak intensity ratios, indicate that such a simple two-state model is inadequate. Furthermore, although reduced in intensity, "a" peaks are still observed when LgtC is titrated with what should be saturating amounts of substrate analogues or UDP product, as well as excess paramagnetic Mn^{2+}. This suggests that a fraction of the "a"-state population cannot interconvert to the "b" state and bind these species, possibly because of the unknown modification detected by mass spectrometry.

Collectively, these studies reveal that LgtC exhibits complex conformational and motional properties that extend beyond what has been observed in the static crystal structures of its binary and ternary substrate complexes. With spectral assignments in hand, we are currently using a range of NMR techniques to examine the behavior of LgtC along its reaction pathway and thereby gain further insights into the structural and dynamic mechanisms underlying glycosyl transfer.

■ ASSOCIATED CONTENT

Supporting Information

Four tables of screening conditions and spectral assignments and 11 figures of activity assays, NMR spectra, and a mass spectrum. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Department of Biochemistry and Molecular Biology, Life Sciences Centre, University of British Columbia, Vancouver, BC, Canada V6T 1Z3. E-mail: mcintosh@chem.ubc.ca. Phone: (604) 822-3341. Fax: (604) 822-5227.

Funding

This research was supported by grants from the Natural Sciences and Engineering Research Council of Canada (L.P.M. and S.G.W.) and the Canadian Institutes for Health Research (S.G.W.). S.G.W. is the recipient of a Canada Research Chair in Chemical Biology. Instrument support was provided by the Canadian Institutes for Health Research, the Canada Foundation for Innovation, the British Columbia Knowledge Development Fund, the University of British Columbia Blusson Fund, and the Michael Smith Foundation for Health Research.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Roman Kittl for providing the LgtC (C128S, C174S, and T273A) construct.

■ ABBREVIATIONS

GT, glycosyltransferase; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; HSQC, heteronuclear single-quantum correlation; IPTG, isopropyl β-D-1-thiogalactopyranoside; k_{cat}, catalytic rate constant; k_{cat}/K_m, second-order rate constant; K_d, equilibrium dissociation constant; K_i, inhibitor equilibrium dissociation constant; K_m, Michaelis constant; LgtC, lipopolysaccharide α-1,4-galactosyltransferase C; LOS, lipooligosaccharide; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; NMR, nuclear magnetic resonance; pH*, pH meter reading without correction for isotope effects; TLC, thin layer chromatography; TROSY, transverse relaxation-optimized spectroscopy; UDP, uridine 5′-diphosphate; UDP-2FGal, uridine-5′-diphosphate 2-deoxy-2-fluorogalactose; UDP-Gal, uridine-5′-diphosphate galactose.

■ REFERENCES

(1) Tzeng, Y. L., and Stephens, D. S. (2000) Epidemiology and pathogenesis of *Neisseria meningitidis*. *Microbes Infect. 2*, 687−700.

(2) Zhu, P., Klutch, M. J., Bash, M. C., Tsang, R. S., Ng, L. K., and Tsai, C. M. (2002) Genetic diversity of three lgt loci for biosynthesis of lipooligosaccharide (LOS) in *Neisseria* species. *Microbiology 148*, 1833−1844.

(3) Ly, H. D., Lougheed, B., Wakarchuk, W. W., and Withers, S. G. (2002) Mechanistic studies of a retaining α-galactosyltransferase from *Neisseria meningitidis*. *Biochemistry 41*, 5075−5085.

(4) Persson, K., Ly, H. D., Dieckelmann, M., Wakarchuk, W. W., Withers, S. G., and Strynadka, N. C. (2001) Crystal structure of the retaining galactosyltransferase LgtC from *Neisseria meningitidis* in complex with donor and acceptor sugar analogs. *Nat. Struct. Biol. 8*, 166−175.

(5) Snajdrova, L., Kulhanek, P., Imberty, A., and Koca, J. (2004) Molecular dynamics simulations of glycosyltransferase LgtC. *Carbohydr. Res. 339*, 995−1006.

(6) Tvaroska, I. (2004) Molecular modeling insights into the catalytic mechanism of the retaining galactosyltransferase LgtC. *Carbohydr. Res. 339*, 1007−1014.

(7) Lairson, L. L., Chiu, C. P., Ly, H. D., He, S., Wakarchuk, W. W., Strynadka, N. C., and Withers, S. G. (2004) Intermediate trapping on a mutant retaining α-galactosyltransferase identifies an unexpected aspartate residue. *J. Biol. Chem. 279*, 28339−28344.

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

(9) Gomez, H., Polyak, I., Thiel, W., Lluch, J. M., and Masgrau, L. (2012) Retaining glycosyltransferase mechanism studied by QM/MM method: LgtC transfers α-galactose via an oxocarbenium ion-like transition state. *J. Am. Chem. Soc. 134*, 4743−4752.

(10) Kleckner, I. R., and Foster, M. P. (2011) An introduction to NMR-based approaches for measuring protein dynamics. *Biochim. Biophys. Acta 1814*, 942−968.

(11) Wang, X., Weldeghiorghis, T., Zhang, G., Imperiali, B., and Prestegard, J. H. (2008) Solution structure of Alg13: The sugar donor subunit of a yeast N-acetylglucosamine transferase. *Structure 16*, 965−975.

(12) Gayen, S., and Kang, C. (2011) Solution structure of a human minimembrane protein Ost4, a subunit of the oligosaccharyltransferase complex. *Biochem. Biophys. Res. Commun. 409*, 572−576.

(13) Kay, L. E. (2011) Solution NMR spectroscopy of supra-molecular systems, why bother? A methyl-TROSY view. *J. Magn. Reson. 210*, 159−170.

(14) Chan, H. W. (2012) Investigating the enzymatic mechanisms of the inverting and retaining glycosyltransferases by NMR spectroscopy. Ph.D. Thesis, Department of Biochemistry and Molecular Biology, The University of British Columbia, Vancouver.

(15) McIntosh, L. P., Wand, A. J., Lowry, D. F., Redfield, A. G., and Dahlquist, F. W. (1990) Assignment of the backbone ^{1}H and ^{15}N NMR resonances of bacteriophage T4 lysozyme. *Biochemistry 29*, 6341−6362.

(16) Waugh, D. S. (1996) Genetic tools for selective labeling of proteins with α-^{15}N-amino acids. *J. Biomol. NMR 8*, 184−192.

(17) Tugarinov, V., Kanelis, V., and Kay, L. E. (2006) Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nat. Protoc. 1*, 749−754.

(18) Sprangers, R., and Kay, L. E. (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. *Nature 445*, 618−622.

(19) Sprangers, R., and Kay, L. E. (2007) Probing supramolecular structure from measurement of methyl ^{1}H-^{13}C residual dipolar couplings. *J. Am. Chem. Soc. 129*, 12668−12669.

(20) Tugarinov, V., and Kay, L. E. (2004) An isotope labeling strategy for methyl-TROSY spectroscopy. *J. Biomol. NMR 28*, 165−172.

(21) Lougheed, B. (1998) *Neisseria meningitidis* lipopolysaccharide galactosyl transferase: Mechanistc investigations and applications for oligosaccharide synthesi. Ph.D. Thesis, Department of Chemistry, The University of British Columbia, Vancouver.

(22) Wilkins, M. R., Gasteiger, E., Bairoch, A., Sanchez, J. C., Williams, K. L., Appel, R. D., and Hochstrasser, D. F. (1999) Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol. 112*, 531−552.

(23) Lairson, L. L., Wakarchuk, W. W., and Withers, S. G. (2007) Alternative donor substrates for inverting and retaining glycosyltransferases. *Chem. Commun.*, 365−367.

(24) Gosselin, S., Alhussaini, M., Streiff, M. B., Takabayashi, K., and Palcic, M. M. (1994) A continuous spectrophotometric assay for glycosyltransferases. *Anal. Biochem. 220*, 92−97.

(25) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277−293.

(26) Goddard, T. D., and Kneeler, D. G. (1999) *Sparky 3*, University of California, San Francisco.

(27) Kay, L. E., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity. *J. Am. Chem. Soc. 114*, 10663−10665.

(28) Pervushin, K., Riek, R., Wider, G., and Wuthrich, K. (1997) Attenuated T_2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U.S.A. 94*, 12366−12371.

(29) Yang, D., and Kay, L. E. (1999) Improved $^{1}H^{N}$-detected triple resonance TROSY-based experiments. *J. Biomol. NMR 13*, 3−10.

(30) Tugarinov, V., and Kay, L. E. (2003) Ile, Leu, and Val methyl assignments of the 723-residue malate synthase G using a new labeling strategy and novel NMR methods. *J. Am. Chem. Soc. 125*, 13868−13878.

(31) Sprangers, R., Li, X., Mao, X., Rubinstein, J. L., Schimmer, A. D., and Kay, L. E. (2008) TROSY-based NMR evidence for a novel class of 20S proteasome inhibitors. *Biochemistry 47*, 6727−6734.

(32) Vincentelli, R., Canaan, S., Campanacci, V., Valencia, C., Maurin, D., Frassinetti, F., Scappucini-Calvo, L., Bourne, Y., Cambillau, C., and

Bignon, C. (2004) High-throughput automated refolding screening of inclusion bodies. *Protein Sci. 13*, 2782−2792.
(33) Gardner, K. H., and Kay, L. E. (1998) The use of ^{2}H, ^{13}C, ^{15}N multidimensional NMR to study the structure and dynamics of proteins. *Annu. Rev. Biophys. Biomol. Struct. 27*, 357−406.
(34) Loria, J. P., Rance, M., and Palmer, A. G., III (1999) Transverse-relaxation-optimized (TROSY) gradient-enhanced triple-resonance NMR spectroscopy. *J. Magn. Reson. 141*, 180−184.
(35) Tossavainen, H., and Permi, P. (2004) Optimized pathway selection in intraresidual triple-resonance experiments. *J. Magn. Reson. 170*, 244−251.
(36) Amero, C., Asuncion Dura, M., Noirclerc-Savoye, M., Perollier, A., Gallet, B., Plevin, M. J., Vernet, T., Franzetti, B., and Boisbouvier, J. (2011) A systematic mutagenesis-driven strategy for site-resolved NMR studies of supramolecular assemblies. *J. Biomol. NMR 50*, 229−236.
(37) Henikoff, S., and Henikoff, J. G. (1992) Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. U.S.A. 89*, 10915−10919.
(38) Chan, P. H., Lairson, L. L., Lee, H. J., Wakarchuk, W. W., Strynadka, N. C., Withers, S. G., and McIntosh, L. P. (2009) NMR spectroscopic characterization of the sialyltransferase CstII from *Campylobacter jejuni*: Histidine 188 is the general base. *Biochemistry 48*, 11220−11230.
(39) Tran-Dinh, S., and Neumann, J. M. (1977) A ^{31}P-NMR study of the interaction of Mg^{2+} ions with nucleoside diphosphates. *Nucleic Acids Res. 4*, 397−403.
(40) Vogel, H. J., and Bridger, W. A. (1982) Phosphorus-31 nuclear magnetic resonance studies of the methylene and fluoro analogues of adenine nucleotides. Effects of pH and magnesium ion binding. *Biochemistry 21*, 394−401.
(41) Lougheed, B., Ly, H. D., Wakarchuk, W. W., and Withers, S. G. (1999) Glycosyl fluorides can function as substrates for nucleotide phosphosugar-dependent glycosyltransferases. *J. Biol. Chem. 274*, 37717−37722.