Specificity of a UDP-GalNAc Pyranose-Furanose Mutase: A Potential Therapeutic Target for Campylobacter jejuni Infections

ARTICLE *in* CHEMBIOCHEM · JANUARY 2014

Impact Factor: 3.09 · DOI: 10.1002/cbic.201300653 · Source: PubMed

DOI: 10.1002/cbic.201300653

Specificity of a UDP-GalNAc Pyranose–Furanose Mutase: A Potential Therapeutic Target for *Campylobacter jejuni* Infections

Myles B. Poulin,[a] Yun Shi,[b] Carla Protsko,[c] Sean A. Dalrymple,[c] David A. R. Sanders,[c] B. Mario Pinto,[b] and Todd L. Lowary*[a]

Pyranose–furanose mutases are essential enzymes in the life cycle of a number of microorganisms, but are absent in mammalian systems, and hence represent novel targets for drug development. To date, all such mutases show preferential recognition of a single substrate (e.g., UDP-Gal). We report here the detailed structural characterization of the first bifunctional pyranose–furanose mutase, which recognizes both UDP-Gal and UDP-GalNAc. The enzyme under investigation (cjUNGM) is involved in the biosynthesis of capsular polysaccharides (CPSs) in *Campylobacter jejuni* 11168. These CPSs are known virulence factors that are required for adhesion and invasion of human epithelial cells. Using a combination of UV/visible spectroscopy, X-ray crystallography, saturation transfer difference NMR spectroscopy, molecular dynamics and CORCEMA-ST calculations, we have characterized the binding of the enzyme to both UDP-Gal*p* and UDP-Gal*p*NAc, and compared these interactions with those of a homologous monofunctional mutase enzyme from *E. coli* (ecUGM). These studies reveal that two arginines in cjUNGM, Arg59 and Arg168, play critical roles in the catalytic mechanism of the enzyme and in controlling its specificity to ultimately lead to a Gal*f*NAc-containing CPS. In ecUGM, these arginines are replaced with histidine and lysine, respectively, and this results in an enzyme that is selective for UDP-Gal. We propose that these changes in amino acids allow *C. jejuni* 11168 to produce suitable quantities of the sugar nucleotide substrate required for the assembly of a CPS containing Gal*f*NAc, which is essential for viability.

Introduction

Campylobacter jejuni are commensal bacteria in many animals, but infections by these organisms are the most common cause of diarrheal disease worldwide in humans.[1] Infection by *C. jejuni* is also regarded as a predetermining factor for the development of the neurological disorder Guillain–Barré syndrome (GBS).[2,3] *C. jejuni* produces a high-molecular-weight capsular polysaccharide (CPS), which is a virulence factor of the organism.[4] The CPS is required for chicken colonization,[5] the adhesion and invasion of human epithelial cells, serum resistance, and the maintenance of cell-surface charge by the organism.[6] *C. jejuni* CPSs are composed of structurally diverse carbohydrates, including unusual heptoses and sugars in the furanose ring form.[4,7,8] Given the importance of surface carbohydrates for these bacteria, there has been an interest in understanding biosynthesis of these structurally diverse glycans[4] and the enzymes involved in furanose biosynthesis have received particular attention.[9,10]

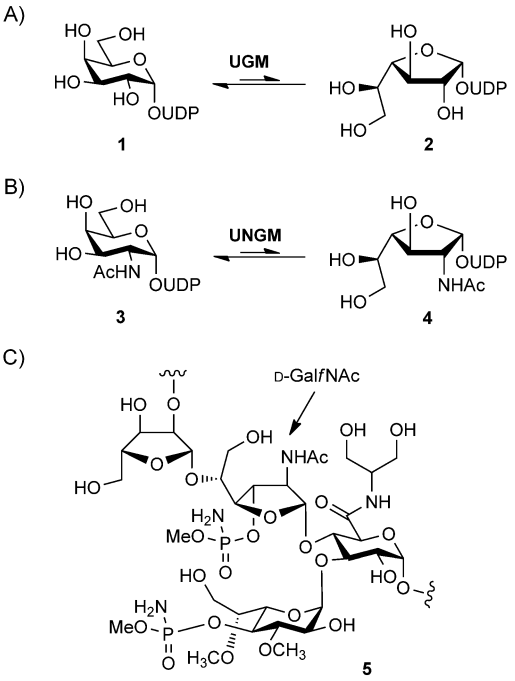
UDP-galactopyranose mutase (UGM) enzymes, encoded by *glf* genes,[11] are responsible for the biosynthesis of the galactofuranose (Gal*f*) sugars found in many microorganisms.[12] These enzymes catalyze the isomerization of UDP-galactopyranose (UDP-Gal*p*, **1**, Scheme 1 A) to UDP-Gal*f* (**2**), the precursor to the Gal*f* residues found in glycoconjugates. In previous studies, we showed that the CPS biosynthetic gene cluster in *C. jejuni* 11168 possesses a homologue of the *glf* gene (*cj1439c*), which encodes for a UDP-*N*-acetylgalactopyranose mutase (cjUNGM, Scheme 1 B).[9] The enzyme is involved in the biosynthesis of the *N*-acetylgalactofuranose (Gal*f*NAc) residue found in the CPS tetrasaccharide repeating unit in this strain (**5**, Scheme 1 C).[13] cjUNGM is a bifunctional enzyme that can catalyze the isomerization of both UDP-Gal*p* (**1**) and UDP-Gal*p*NAc (**3**) to the corresponding furanose sugar nucleotides (**2** and **4**, respectively). Although no Gal*f* residues have been found in glycoconjugates of *C. jejuni* 11168, cjUNGM is able to complement UGM activity in an *E. coli* ΔUGM knockout strain to restore lipopolysaccharide (LPS) biosynthesis.[9] Thus, cjUNGM can function in vivo as a UGM.

In this paper we characterize the structure and specificity of cjUNGM and compare its substrate-recognition patterns to those of a UGM from *E. coli* (ecUGM). These studies have allowed us to understand the specificity of this interesting protein, the first pyranose–furanose mutase demonstrated to effi-

[a] *Dr. M. B. Poulin, Prof. T. L. Lowary*
Alberta Glycomics Centre and Department of Chemistry
The University of Alberta, Gunning–Lemieux Chemistry Centre
11227 Saskatchewan Drive, Edmonton, AB T6G 2G2 (Canada)
E-mail: tlowary@ualberta.ca

[b] *Y. Shi, Prof. B. M. Pinto*
Department of Chemistry, Simon Fraser University
8888 University Drive, Burnaby, BC V5A 1S6 (Canada)

[c] *C. Protsko, Dr. S. A. Dalrymple, Prof. D. A. R. Sanders*
Department of Chemistry, University of Saskatchewan
110 Science Place, Saskatoon, SK S7N 5C9 (Canada)



Scheme 1. A) UGM-catalyzed interconversion of UDP-Gal*p* and UDP-Gal*f*. B) UNGM-catalyzed interconversion of UDP-Gal*p*NAc and UDP-Gal*f*NAc. C) Structure of the *C. jejuni* serotype HS:2 CPS highlighting the Gal*f*NAc residue.

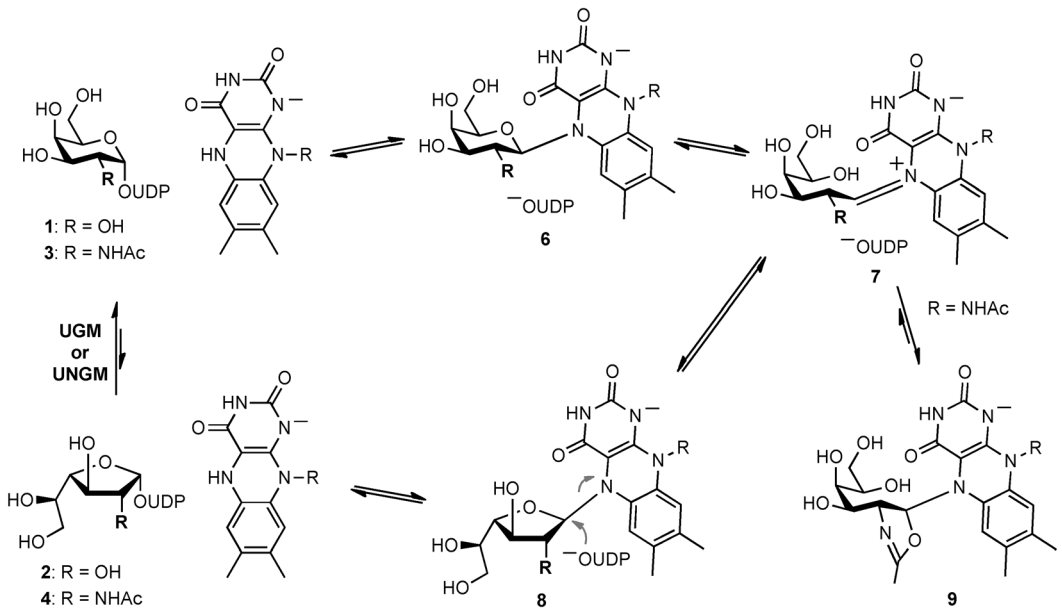
ciently turn over two different, albeit structurally similar, substrates. We conclude that the salient differences between the two enzymes are subtle changes in substrate recognition mediated by two active-site arginine residues in cjUNGM, which are replaced with lysine and histidine in ecUGM. These investigations provide the first insights into the dual specificity of this mutase, a new drug target for treating *C. jejuni* infections.

Results and Discussion

There is no covalent intermediate between ecUGM and UDP-GalNAc

Both ecUGM and cjUNGM are flavoenzymes. The N-5 of reduced $FADH^-$ reacts directly with the UDP-Gal(NAc) substrate (Scheme 2) to displace UDP and give **6**. Next, the pyran ring opens to form a covalent iminium ion intermediate, **7**, which proceeds through intermediate **8** on the way to the product.[14,15] Despite the high sequence homology between cjUNGM and ecUGM, only the former turns over UDP-GalNAc as a substrate. In our previous study[9] we proposed that arginine residues in cjUNGM (R59 and R168) play a role in the recognition and turnover of UDP-GalNAc, and we hypothesized that an interaction between R59 and the acetamido group of UDP-GalNAc is required to prevent the formation of unproductive oxazoline side products (**9**, Scheme 2). It is possible that the ecUGM, which contains instead a histidine at this position, could become trapped as the oxazoline (**9**) after formation of the iminium ion (**7**) from UDP-GalNAc. Alternatively, the ecUGM may not bind to UDP-GalNAc at all. To distinguish between these possibilities, we sought to directly observe the formation of any of these species.

Flavin derivatives often exhibit characteristic UV/visible absorbances that can be used to identify the species that are produced during an enzymatic reaction. For example, it was previously shown that the covalent iminium intermediate of the UGM reaction (**7**, R = OH) gives rise to an increased absorb-



Scheme 2. Proposed mechanism for cjUNGM based on that of UGM. The reduced form of the flavin cofactor is required for activity. In an earlier investigation, it was hypothesized that UGMs lacking R59 (i.e., as in ecUGM) form an unproductive oxazoline intermediate (**9**) with UDP-Gal*p*NAc (**3**), thus preventing turnover of this substrate.[9]

ance at 500 nm with an isosbestic point around 475 nm.[14] We hypothesized that a similar covalent intermediate would exist en route to the proposed unproductive oxazoline intermediate **9** with ecUGM and UDP-Gal*p*NAc (**3**). Therefore, we sought to detect such an iminium ion intermediate from the absorbance spectrum of the ecUGM in the presence of UDP-Gal*p*NAc. First, we examined the absorbance of ecUGM with increasing amounts of UDP-Gal*p* (**1**). Similar to published results with *Klebsiella pneumoniae* UGM,[14,16] an increase in the absorbance at 500 nm with increasing concentration of UDP-Gal*p* (Figure 1 A) was observed. However, when UDP-Gal*p*NAc (**3**) was added, there was no change in the absorbance of the sample. As seen in Figure 1 B, the constant FAD absorbance with increasing concentrations of UDP-Gal*p*NAc suggests that no iminium ion species is formed between the ecUGM FAD cofactor and this substrate. Consequently the formation of an unproductive oxazoline intermediate, as previously proposed is not supported. These absorbance measurements provide evidence that ecUGM does not employ UDP-Gal*p*NAc as a substrate; however, the origin for the specificity of this ecUGM vis-à-vis cjUNGM remained unclear.[9] Therefore, we tried other approaches such as X-ray crystallography, saturated transfer difference NMR spectroscopy, molecular dynamics, and CORCEMA-ST calculations.

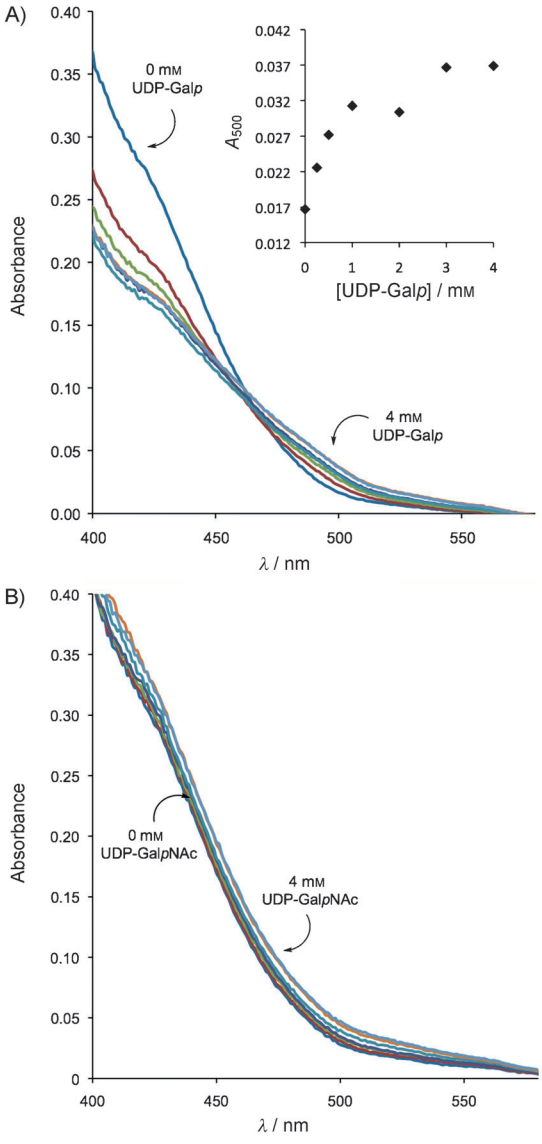


Figure 1. UV/visible spectrum for titrations of ecUGM with A) UDP-Gal*p* (**1**) and B) UDP-Gal*p*NAc (**2**). Concentrated UDP-Gal*p* or -Gal*p*NAc was titrated into reduced ecUGM. The formation of a covalent adduct between the FAD cofactor and the substrate can be seen by the increase in A_{500} with increasing [UDP-Gal*p*] (inset), reaching a maximum at ~3 mM. Titrating with UDP-Gal*p*NAc revealed no ΔA_{500}.

Crystal structure of cjUNGM

The crystal structure of cjUNGM is very similar to that of the other members of UGM family (Figure 2). Superposition of cjUNGM on ecUGM (chain B) reveals that the two dimers superimpose with a root mean square deviation (RMSD) of 0.9 Å for all Cα's, a similar superposition with the monomer of kpUGM revealed an RMSD of 1.3 Å. The substrate cleft is located between domains 1 (FAD-binding) and 2 (5-helix bundle), and, in the cjUNGM structure, the mobile loop noted in other UGM structures is in the "open" conformation. This loop plays an important role in the binding of UDP-Gal, closing around the substrate in enzyme–ligand complexes.[16–20] All of the conserved residues noted as important for enzyme activity[19,20] are located in the same positions in cjUNGM as in the other structures, with the noted exception of the residues in the mobile

Figure 2. Superposition of a monomer of UNGM (red) with monomer B of ecUGM (blue; PDB ID: 1I8T). The mobile loop that moves position when substrate binds is colored yellow in the UNGM monomer and green in the ecUGM monomer. FAD from UNGM is shown in a ball-and-stick representation.

loop. Given that we have been unsuccessful in obtaining a cocrystal of the enzyme bound to a substrate, we used STD-NMR spectroscopy in conjunction with molecular dynamics and CORCEMA-ST calculations[21–26] to provide a model of the complex structures.

Differentiating the binding epitopes for ecUGM and cjUNGM with UDP-Gal*p* and UDP-Gal*p*NAc by STD-NMR spectroscopy

Binding epitopes of UDP-Galp and UDP-GalpNAc: STD-NMR experiments were performed to explore the binding of both ligands to cjUNGM or ecUGM in its reduced state. The largest STD effects were observed for H1R/H5U protons in all four cases; this is consistent with previous studies of the kpUGM–UDP-Gal*p* complex.[24] Thus, the STD intensity of this peak was set to 100% as a reference, and relative STD intensities for other protons were normalized to this peak (Figure 3). STD effects were generally larger from the uridine moiety than those from the galactopyranose or *N*-acetylgalactosamine, and only H1G and H4G showed comparable STD intensities with uridine protons. In addition, UDP-Gal*p* exhibited similar STD enhancement profiles in both cjUNGM and ecUGM binding sites, with the exception of H2G; this suggests similar binding modes of UDP-Gal*p* in the active sites of both enzymes. For UDP-Gal*p*NAc, a few discrepancies such as H4G and H6G were observed for its enhancement profiles, thus implying somewhat different binding patterns for the Gal*p*NAc portion. Yet, STD intensities alone cannot directly afford substrate binding modes, and therefore we resorted to MD simulations in conjunction with CORCEMA-ST calculations to develop a semiquantitative model of substrate recognition.[21–26]

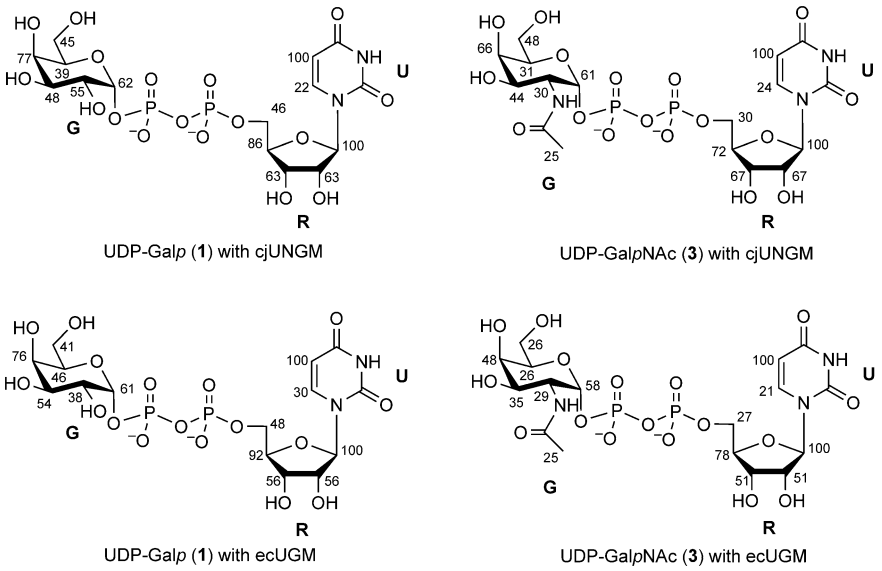


Figure 3. Epitope mapping of UDP-Gal*p* (**1**) and UDP-Gal*p*NAc (**3**) in the active site of cjUNGM or ecUGM (U: uracil, R: ribose, G: galactopyranose or 2-acetamido-2-deoxy-galactopyranose). All numbers shown are percentage STD values of non-exchangeable protons after normalization against that of H1R/H5U.

Molecular dynamics studies of binding modes

The active-site structures of bacterial UGMs and cjUNGM are highly conserved,[9,16–18] and the binding conformations of UDP-Gal*p* with *Deinococcus radiodurans* UGM (drUGM) and kpUGM are almost identical.[16,19] Hence, the ligand binding modes for ecUGM and cjUNGM should be similar to those of other UGMs, and the bound structures of both UDP-Gal*p* and UDP-Gal*p*NAc to kpUGM, the latter of which was generated by molecular docking, were used as reasonable starting points for MD studies. Simulations of 50 ns were performed to allow sufficient time to probe ligand binding to both cjUNGM and ecUGM. It has been suggested that the movement of the mobile loop (residues 162–175 in ecUGM and 161–174 in cjUNGM) towards a closed conformation is essential for UGM substrate binding and catalytic activity.[23,27] Indeed, this loop closure was observed for all four U(N)GM–ligand complexes during the course of MD simulations. The last 20 ns trajectory of each system, all of which have the mobile loop closed and thus represent the "bound" structure, was chosen for further analysis. Unless otherwise noted, residue numbers refer to cjUNGM. Clustering was carried out based on the RMSD of active-site and ligand residues, and the central structure of the largest cluster was used for static structure comparisons (Figure 4).

Similar uridine binding patterns were observed across the four complexes. A strong hydrogen bond with a donor (uracil N3)–acceptor (Leu146 O) distance of 3.0 Å or less persisted during the MD simulation. The π-stacking interaction between uracil and the Tyr150 side chain was also maintained. These agreed with observations from related crystal structures.[16,19] However, hydrogen bonds between ribose O2 and Gln154 OE1 as well as Gln154 NE2 and ribose O3 (except for the cjUNGM–UDP-Gal*p* complex) were also present throughout the MD trajectories. This is different from related crystal structures, in which NE1 of Trp184 (drUGM) or Trp160 (kpUGM) forms a hydrogen bond with ribose O2 or O3, and the adjacent Gln has its side chain pointing away from the ligand.[16,19] Yet, the interaction modes revealed by MD did provide one more stable hydrogen bond, and the Trp NE1 could still participate in water-mediated hydrogen bonding with the ligand.

The diphosphate binding modes, dominated by ionic interactions with two flanking guanidinium groups, were different between cjUNGM and ecUGM, although they had similar ionic interactions, with the functionally critical Arg277 in the active

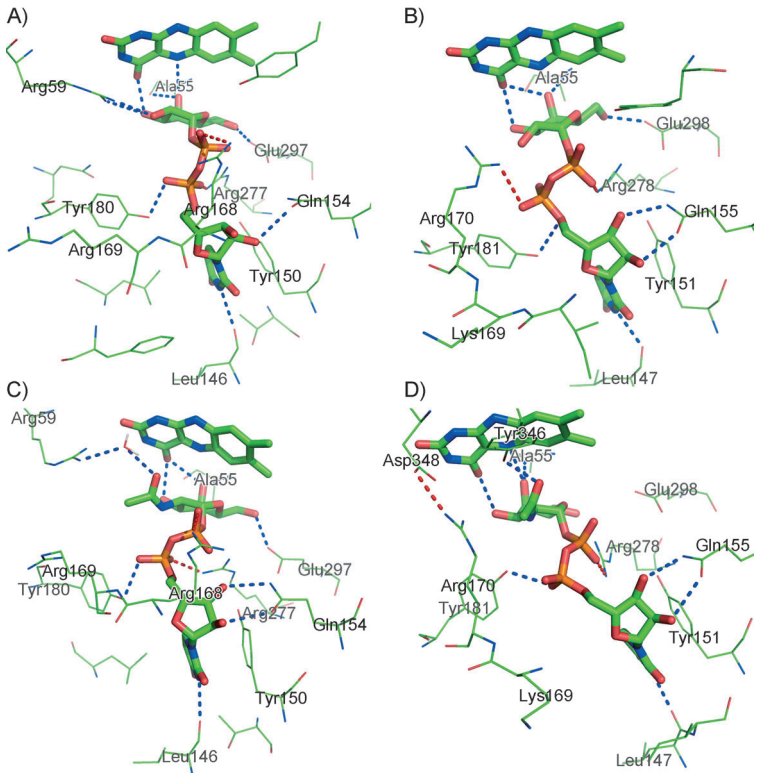


Figure 4. Binding-site interactions of A) cjUNGM–UDP-Gal*p*, B) ecUGM–UDP-Gal*p*, C) cjUNGM–UDP-Gal*p*NAc and D) ecUGM–UDP-Gal*p*NAc. Ligands and flavin isoalloxazine rings are shown as sticks, and selected active-site residues are shown as thin lines (green: carbon, blue: nitrogen, red: oxygen, orange: phosphorus). Hydrogens and water molecules are omitted. Blue dots represent hydrogen bonds; red dots indicate ionic interactions.

site.[20] The other interacting guanidinium group came from mobile loop Arg170 in the ecUGM–UDP-Gal*p* complex, which corresponds to the Arg174 in kpUGM that was shown to be essential for UGM activity.[20] In cjUNGM–UDP-Gal*p* and cjUNGM–UDP-Gal*p*NAc, however, it was Arg168 that displayed strong ionic interactions with the diphosphate. Arg168 in cjUNGM corresponds to Lys169 in ecUGM and Lys173 in kpUGM, neither of which interacts with the substrate as they point away from the active site. Interestingly, Arg168 in cjUNGM is important for its bifunctionality as it favors the conversion of UDP-Gal*p*NAc, which cannot be converted by ecUGM, to its furanose counterpart.[9] The Coulombic interaction energy between the diphosphate and the guanidine moiety from the corresponding mobile loop arginine was calculated based on the force field used in MD simulations (Table 1). Because this interaction is key for catalysis, the much smaller interaction energy in ecUGM–UDP-Gal*p*NAc, 105 kJ mol^{-1} smaller than that in ecUGM–UDP-Gal*p*, accounts for the inability of ecUGM to catalyze the conversion of UDP-Gal*p*NAc. Of note, this interaction energy in cjUNGM–UDP-Gal*p*NAc is 12.5 kJ mol^{-1} larger than that in cjUNGM–UDP-Gal*p*, thus suggesting its role in favoring UDP-Gal*p*NAc over UDP-Gal*p*. This is also consistent with a previous study of UNGM mutants that shows the role of Arg168 in substrate binding (K_m) rather than catalytic turnover (k_{cat}).[9] Hydrogen bonds donated from Tyr180 (as well as Tyr311 in the case of ecUGM–UDP-Gal*p*) further stabilized the diphosphate group, whereas hydrogen bonding from Tyr310 (except in ecUGM–UDP-Gal*p*) and Tyr345 were mostly absent during the course of MD simulations. However, hydrogen bonding interactions from these active-site tyrosines are much less important, and their absence only results in a small decrease in the catalytic efficiency of kpUGM.[20]

Binding of the Gal*p* or Gal*p*NAc moiety was characterized by a series of hydrogen bonding interactions. The flavin N5 donated a hydrogen bond to Gal*p* or Gal*p*NAc O4, which in turn donated a hydrogen bond to either Ala54 O or flavin O4 during the course of MD simulations. The O3 of Gal*p* or Gal*p*NAc was also involved in strong hydrogen bonding with flavin O4, analogous to that observed in MD simulation of the kpUGM–UDP-Gal*f* complex.[23] With the exception of ecUGM–UDP-Gal*p*NAc, the O6 of Gal*p* or Gal*p*NAc donated a hydrogen bond to Glu297 OE1. This represents the largest discrepancy with the kpUGM crystal structure, in which Asn80 ND2 donates a hydrogen bond to Gal*p* O6, as the C6 hydroxyl points in a completely different direction within the binding site (Figure 5). Nevertheless, this deviation resulted in a strong hydrogen bond (shorter donor–acceptor distance and Glu297 OE1 has a partial negative charge), and concomitant repositioning of the sugar ring did involve O3 in hydrogen bonding with flavin. Furthermore, Glu297 does have a role in substrate binding and catalytic turnover.[23] The distance between the flavin N5 and sugar C1 also explains the lack of catalytic activity in the ecUGM–UDP-Gal*p*NAc complex, as its average value is approximately 1 Å greater than those in the other three complexes (Table 2).

Arg59 of cjUNGM, corresponding to a histidine in ecUGM and kpUGM, has been proposed to interact directly with the NHAc group on Gal*p*NAc because it favors the conversion of

Table 1. Coulombic interaction energies between the ligand diphosphate and the mobile-loop guanidinium moiety.

	cjUNGM: UDP-Gal*p*	ecUGM: UDP-Gal*p*	cjUNGM: UDP-Gal*p*NAc	ecUGM: UDP-Gal*p*NAc
E [kJ mol^{-1}][a]	−138.9	−132.4	−151.4	−27.7
ΔE [kJ mol^{-1}][b]	−111.2	−104.7	−123.7	0

[a] Values are averaged over the final 20 ns of each trajectory. [b] After subtraction of E in ecUGM–UDP-Gal*p*NAc, −27.7 kJ mol^{-1}.

Table 2. Averaged distances between ligand anomeric carbon and flavin N5.

	cjUNGM–UDP-Gal*p*	ecUGM–UDP-Gal*p*	cjUNGM–UDP-Gal*p*NAc	ecUGM–UDP-Gal*p*NAc
distance [Å]	4.05	4.01	4.26	5.12

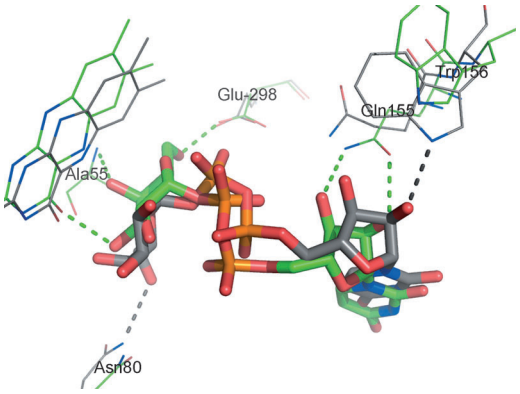


Figure 5. Major differences between UDP-Gal*p* binding modes from MD simulations and the crystal structure. Green dots represent hydrogen bonds observed in the MD simulation of ecUGM–UDP-Gal*p*, and gray dots indicate hydrogen bonds in the crystal structure of kpUGM–UDP-Gal*p* (PDB ID: 3INT:B). UDP-Gal*p* is shown as sticks; selected active-site residues are shown as thin lines (blue: nitrogen, red: oxygen, orange: phosphorus, green: carbon in ecUGM–UDP-Gal*p*, gray: carbon in kpUGM–UDP-Gal*p*). Residues are labeled for ecUGM.

UDP-Gal*p*NAc over that of UDP-Gal*p*.[9] However, the MD simulations do not support this hypothesis. Direct hydrogen bonds were observed between Arg59 NH2 and Gal*p* O2 as well as Arg59 NH1 and Gal*p* O3 during the MD simulation of cjUNGM–UDP-Gal*p*, whereas only water-mediated hydrogen bonds were observed for cjUNGM–UDP-Gal*p*NAc. Thus, we speculate that Arg59 might not contribute significantly to the K_m of UDP-Gal*p*NAc but instead might facilitate the catalytic reaction. Indeed, kinetic parameters of the R59H mutant and wild-type cjUNGM suggest a small role of Arg59 in substrate binding and a more important role in catalytic turnover that favors UDP-Gal*p*NAc over UDP-Gal*p*.[9]

A clear test of the critical role of Arg59 and Arg168 in controlling the specificity of cjUNGM would be to mutate the corresponding residues in ecUGM. Unfortunately, all attempts to prepare mutant clones of the *E. coli* enzyme failed to express.

Experimental validation of binding models

To validate the binding models, 100 snapshots were generated from the final 20 ns of MD simulations so as to perform CORCEMA-ST calculations.[21,22] Calculated STD intensities of the 100 frames were averaged for each non-exchangeable ligand proton and normalized against that of H5U/H1R. Both ecUGM–UDP-Gal*p* and cjUNGM–UDP-Gal*p* showed excellent correlations between experimental and calculated STD values (Figure 6), given the approximations in the experimental and theoretical treatments, thus indicating that the conformational ensemble generated from the MD simulation is representative of the corresponding bound structures in solution. For UDP-

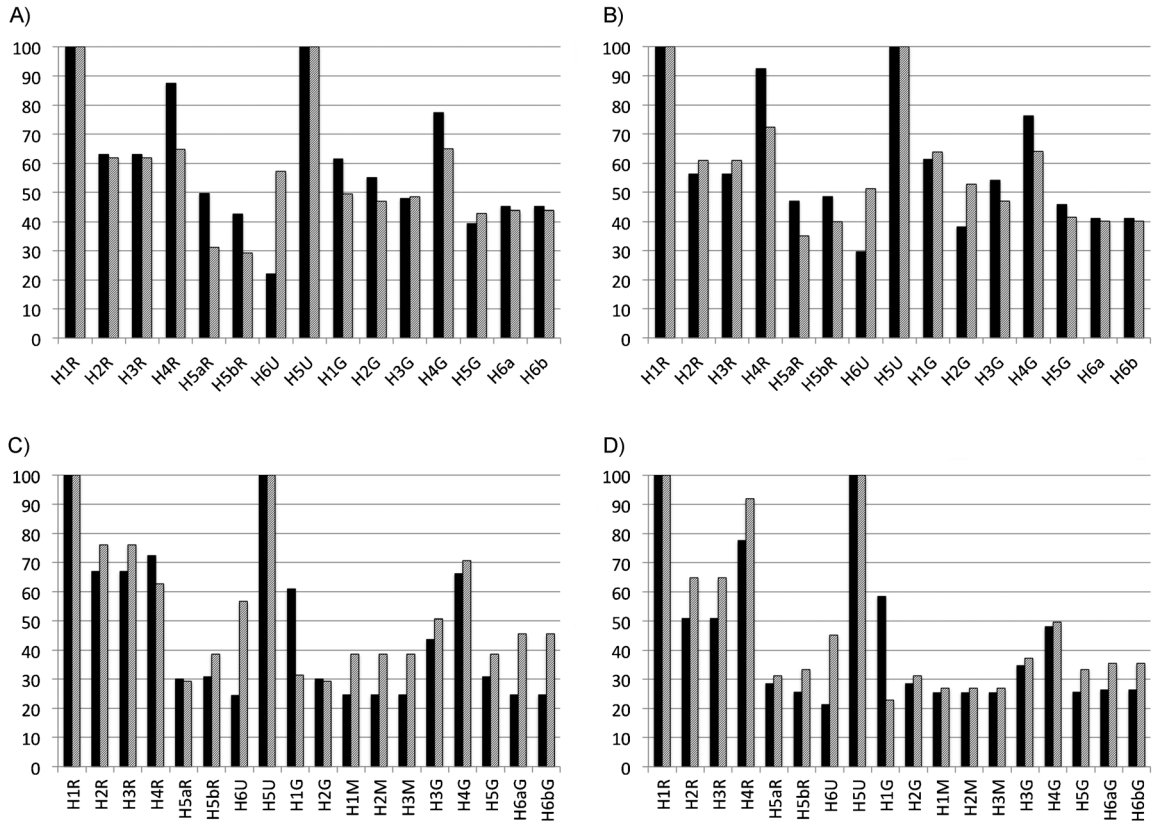


Figure 6. Comparisons of experimental (■) and calculated (■) STD values from the CORCEMA-ST protocol for A) cjUNGM–UDP-Gal*p*, B) ecUGM–UDP-Gal*p*, C) cjUNGM–UDP-Gal*p*NAc, and D) ecUGM–UDP-Gal*p*NAc. Both experimental and calculated STD values were normalized against the corresponding STD values of H1R/H5U. The NOE *R* factors, excluding contributions from H1R/H5U, are 0.25, 0.20, 0.35, and 0.33, respectively.

Gal*p*NAc, the correlations were acceptable, thus suggesting reasonable MD models for cjUNGM–UDP-Gal*p*NAc and ecUGM–UDP-Gal*p*NAc. Moreover, similar STD enhancement profiles for ecUGM–UDP-Gal*p* and cjUNGM–UDP-Gal*p* corroborated the corresponding MD models that demonstrated analogous binding modes of UDP-Gal*p* (Figure 6A and B), whereas discrepancies in STD values between cjUNGM–UDP-Gal*p*NAc and ecUGM–UDP-Gal*p*NAc in the Gal*p*NAc portion were illustrated by different binding conformations (Figure 6C and D). Therefore, the binding models produced by MD simulations can be rationalized for all four U(N)GM–ligand complexes.

Conclusions

We have employed a series of experimental and theoretical techniques to investigate the origin of substrate specificity in cjUNGM, a bifunctional pyranose–furanose mutase that converts both UDP-Gal*p* and UDP-Gal*p*NAc to their corresponding furanose derivatives. The binding modes of cjUNGM with UDP-Gal*p* and UDP-Gal*p*NAc were generated by MD simulations, which were validated by agreement between theoretical STD effects from CORCEMA-ST calculations and experimental STD enhancements. The interactions of the substrates with ecUGM, which is highly homologous to cjUNGM, but which is monofunctional for UDP-Gal, were revealed by the same approach. Our results demonstrated the important role of cjUNGM Arg168 in substrate recognition, as it formed strong ionic interactions with both substrates and could stabilize the released UDP during catalysis. Such ionic interactions were more pronounced with UDP-Gal*p*NAc than UDP-Gal*p*, thus suggesting that Arg168 favors the binding and turnover of UDP-Gal*p*NAc. In contrast, no residue from the mobile loop of ecUGM engages in ionic interactions with the diphosphate moiety of UDP-Gal*p*NAc, and this leads to the absence of catalytic activity. Indeed, UV/visible spectroscopy demonstrated that the binding of UDP-Gal*p*NAc with ecUGM is not followed by the formation of an iminium ion intermediate, a result that is corroborated by the relatively large distance between the pyranose C1 and flavin N5 atoms during the course of MD simulations.

Arg59 (His in ecUGM) is the other active-site residue that differs between cjUNGM and ecUGM. Arg59 showed different hydrogen-bonding interactions with UDP-Gal*p* and UDP-Gal*p*NAc, thus confirming its role in the substrate binding. Previous studies revealed that mutating Arg59 to His reduces the ability of cjUNGM to process UDP-Gal*f*NAc more dramatically than does mutation of Arg168. However, the work reported here does not provide a clear explanation for the effect of Arg59 on catalysis, and further study is required. We can conclude, however, that our previous proposal that Arg59 engages in direct hydrogen bonding with the acetamido residue of the GalNAc is not supported.

It is interesting to note that *C. jejuni* 11168 has not been reported to produce any Gal*f*-containing glycoconjugates. Therefore the role of a bifunctional enzyme that is capable of producing a substrate (UDP-Gal*f*) that is not further incorporated into glycoconjugates remains unclear. Certainly UDP-Gal*p*, an essential biochemical intermediate, is produced by the organism. If one assumes that cjUNGM arose though evolution of a protein similar or identical to ecUGM, then it is possible to envision that the organism has been required to modify the enzyme to produce suitable quantities of the sugar nucleotide substrate required for the (yet unidentified) glycosyltransferase involved in CPS biosynthesis. One can speculate that further evolution of the protein might lead to an enzyme that is more selective for UDP-GalNAc, although there might not be sufficient evolutionary pressure to effect such changes.

Experimental Section

Enzyme preparation: The *E. coli* UGM and *C. jejuni* UNGM enzymes were prepared as reported in our earlier investigation,[9] with slight modifications as follows. For UV/Vis spectroscopy studies, *E. coli* UGM was prepared from 3 L cultures and diluted to a final concentration of 162 μM. For crystallography, *C. jejuni* UNGM was prepared from 4 L cultures and purified by Ni-NTA agarose chromatography eluting with imidazole (250 mM) in Tris·HCl (40 mM, pH 7.4) with NaCl (150 mM). After dialysis in buffer (4 L, 40 mM Tris·HCl, pH 7.4, with 150 mM NaCl), the protein was concentrated to ~7 mg mL^{-1} by centrifugal filtration in a 30000 molecular weight cut off centrifugal filter device (Amicon Ultra-15 Ultracel-30). Protein samples used for NMR analysis were lyophilized from D_2O in 1 mg aliquots directly before use, and dissolved directly in D_2O for use. Lyophilized proteins maintained >80% activity as determined by using an HPLC assay described previously[9] (data not shown).

UV/Vis spectroscopy: FAD UV/Vis spectroscopy was performed on a Hewlett Packard 8453 UV/Vis spectrophotometer. Measurements were made in identical 50 μL cuvettes, with or without *E. coli* UGM (150 μM), in a 110 μL final volume of potassium phosphate buffer (100 mM, pH 7.4) containing freshly prepared sodium dithionite (20 mM). UV/Vis spectra were recorded in wave-scan mode as concentrated UDP-Gal*p* or UDP-Gal*p*NAc solutions were titrated into each cuvette, as described previously.[14] Attempts to carry out this experiment with *C. jejuni* UNGM were unsuccessful. At the concentrations required to get a detectable signal, protein precipitation occurred.

X-ray crystallography

Crystallization: Purified *C. jejuni* UNGM was crystallized by using the microbatch method. Initial screen were performed by using commercially available screening kits (Qiagen and Hampton Research). The initial screens were set up with 0.6 μL protein and 0.6 μL screening solution, and the drops were subsequently covered with paraffin oil. Crystals were found under several conditions, and optimization trials resulted in good-quality crystals being grown in ammonium acetate (0.2 M), trisodium citrate (0.1 M, pH 5.6) poly(ethylene glycol) 4000 (25%, *w/v*).

Diffraction and data collection: Prior to data collection, crystals were soaked for 5 min in a cryoprotectant mother liquor containing 30% glycerol, mounted on a cryoloop and flash-cooled in liquid nitrogen. All crystals were diffracted by using synchrotron radiation at a wavelength of 0.9790 Å at the Canadian Macromolecular Crystallography Facility (CMCF-1) beamline (ID08ID-1) at Canadian Light Source (CLS, Saskatoon, SK, Canada). A total of 180 images were collected on a MAR300 CCD X-ray detector; each was exposed for 1 s with 1.0° oscillation at a crystal-to-detector distance of 250 mm. The diffraction images were processed and scaled by using the XDS software package.[28] Autoindexing of the diffraction data showed that the crystal belonged to the space

group $P2_12_12_1$, with unit-cell parameters $a=48.4$, $b=116.0$, $c=165.4$ Å. The asymmetric unit appeared to contain one dimer, thus giving a Matthews coefficient[29] of 2.67 $Å^3\,Da^{-1}$ and a solvent content of 54%. The final data statistics are listed in Table 3.

Table 3. Data collection and refinement statistics. Values in parentheses are for the highest-resolution shell.

beamline	08ID-1, CLS
wavelength [Å]	0.9790
T [K]	100
space group	$P2_12_12_1$
unit-cell parameters [Å]	$a=48.42$, $b=116.15$, $c=165.24$
unit-cell parameters [°]	$\alpha=\beta=\gamma=90$
resolution [Å]	50.00–2.00 (2.05–2.00)
observed reflections	454 777
unique reflections	63 966
completeness [%]	100.0 (100.0)
multiplicity	7.11 (6.73)
mean $I/\sigma(I)$	12.6 (2.5)
R_{merge} [a]	0.140 (0.830)
R_{meas} [b]	0.151 (0.899)
$CC_{1/2}$ [c]	99.7 (71.6)
monomers per ASU	2
resolution range [Å]	47.51–2.00
R_{work}/R_{free} [%]	17.1/21.2
number of protein residues	1–365, 1–363
ligands	1 SO_4/1 GOL/1 FDA, 1 SO_4/1 GOL/2 NH_4/1 FAD
number of solvent molecules	777
RMSD bond length [Å]/angles [°]	0.002/0.670
Ramachandran favored [%]	96.4
Ramachandran allowed [%]	3.6

[a] $R_{merge}=\Sigma_{hkl}\Sigma_i|I_i(hkl)-\langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the measured intensity and $\langle I(hkl)\rangle$ is the average intensity over symmetry-equivalent reflections. [b] $R_{meas}=\Sigma_{hkl}(\sqrt{(n_h/(n_h-1))})\Sigma_i|I_i(hkl)-\langle I(hkl)\rangle|/\Sigma_{hkl}\Sigma_i I_i(hkl)$. [c] $CC^*=\sqrt{[(2CC_{1/2})/(1+CC_{1/2})]}$, where CC* is an estimate of CC_{true} based on a finite sample size.

Structure solution and refinement: A molecular replacement solution for UNGM was found by using the atomic coordinates of ecUGM[17] with the program MrBUMP implemented through the CCP4 program suite.[30,31] Initial refinement was performed by using Refmac5,[32] and model rebuilding was carried out in COOT.[33] Addition of the FAD cofactors and water molecules, and the final rounds of refinement were performed in PHENIX.[34] Noncrystallographic symmetry (NCS) restraints were applied during the initial refinement step, but were removed during the final rounds of refinement. The final model was validated using MOLPROBITY.[35] Final refinement statistics are shown in Table 3. The final structure coordinates were deposited in the PDB (PDB ID: 4MO2).

STD-NMR spectroscopy: NMR samples containing ecUGM or cjUNGM (1.0 mg) in potassium phosphate D_2O buffer (50 mM, pH 7.4) were prepared, and freshly prepared sodium dithionite (20 mM) was added to a total volume of 0.7 mL under N_2 to reduce the cofactor FAD.[17] UDP-Gal*p* or UDP-Gal*p*NAc was then added to give a final ligand concentration of 1.8 mM at a ligand/protein ratio of 50:1. Ligand resonances were assigned by using $^1H,^1H$ COSY and $^1H,^{13}C$ HMQC NMR spectroscopy. Enzyme-only samples without the ligand were also prepared to correct for protein resonances in the STD-NMR spectra. STD-NMR spectra were recorded on a Varian/Agilent VNMRS four-channel, dual receiver 700 MHz spectrometer equipped with an inverse detection, cryo-cooled $^1H\{^{15}N/^{13}C\}$ triple-resonance, Z-gradient probe at 300 K. Protein resonances at 0.0 ppm (50 ppm for reference spectra) were selectively irradiated for a total saturation time of 2.04 s by using a series of Gaussian-shaped pulses.[36]

Molecular dynamics simulations: The crystal structure of *K. pneumoniae* UGM (kpUGM) in complex with UDP-Gal*p* (PDB ID: 3INT, chain B)[16] was used as the reference bound structure of UDP-Gal*p* because kpUGM is highly homologous with both cjUNGM and ecUGM.[9] For the reference bound structure of UDP-Gal*p*NAc, docking with AutoDock Vina[37] was first performed. Water molecules and the ligand (UDP-Gal*p*) were removed from the aforementioned kpUGM structure prior to the docking of UDP-Gal*p*NAc, and the docking pose with the best overlap with UDP-Gal*p* was chosen as the reference bound structure for UDP-Gal*p*NAc. These reference bound structures were subsequently fitted to analogous positions in both monomers A and B of the cjUNGM and ecUGM (PDB ID: 1I8T) crystal structures.[17] The four resulting enzyme–ligand complexes in their dimeric forms were then processed by the LEaP module in AmberTools1.5[38] for the addition of missing atoms and parameterization of force fields. Residues were parameterized by either the Amber ff99SB[39] or GLYCAM06[40] force field wherever possible. Missing parameters for the cofactor $FADH^-$ and the diphosphate moiety in ligands were manually added; R.E.D. Server[41] was used to derive atomic charges according to the RESP model,[42] taking into account multiple conformations and multiple orientations as per AMBER convention.[43]

MD simulations of the four enzyme–ligand systems, cjUNGM–UDP-Gal*p*, ecUGM–UDP-Gal*p*, cjUNGM–UDP-Gal*p*NAc, and ecUGM–UDP-Gal*p*NAc, were conducted with GROMACS 4.5.4.[44] Each system was solvated in a dodecahedron periodic box with a TIP3P water model,[45] and NaCl (0.15 M) was added to balance ionic charge in the system. Energy minimizations were carried out with steepest-descent integrator and conjugate gradient algorithm sequentially to achieve a maximum force of $<500\,kJ\,mol^{-1}\,nm^{-1}$ on any atom. A twin-range cutoff scheme was used to evaluate short-range, nonbonded interactions, with van der Waals interactions truncated at 1.4 nm and electrostatic interactions truncated at 0.9 nm. Long-range electrostatic interactions were treated by the particle mesh Ewald (PME) method.[46,47] The temperature was maintained at 300 K by using a velocity rescaling thermostat[48] with a coupling constant of 0.1 ps, and the pressure was maintained at 1.0 atm by using a Berendsen barostat,[49] with a coupling constant of 1 ps. Simulations were performed with a time step of 2 fs, and all bonds involving hydrogen atoms were constrained by a parallel linear constraint solver (P-LINCS).[50] Each system was equilibrated under a constant volume (NVT) ensemble for 200 ps and a constant pressure (NPT) ensemble for 300 ps. A harmonic position restraint with a force constant of $1000\,kJ\,mol^{-1}\,nm^{-2}$ was applied to all the heavy atoms of the enzyme and the UDP portion of the cofactor; position restraints were applied to maintain key interactions between the Gal*p* or Gal*p*NAc moiety and the enzyme. Following equilibration, production MD simulations were conducted for 50 ns for each system, without any constraints. For consistency, only monomer B of each dimeric system was considered for analysis and CORCEMA-ST calculations.

CORCEMA-ST calculations: The theory of CORCEMA-ST and the details of executing the CORCEMA-ST protocol have been previously described.[21–26] STD intensities were calculated on 100 snapshots from the last 20 ns of MD trajectories (every 200 ps), and truncated structures with only the residues within 1.3 nm of the ligand were considered. Based on experimental conditions, the concentration

of ligand was 1.8 mM, and the ligand/enzyme ratio was kept at 50:1. The spectrometer frequency was set at 700 MHz. k_{on} was set in the range of 10^5–10^6 s^{-1} M^{-1}, and the equilibrium constant, K_{eq}, was set in the range of 10^4–10^5 M^{-1}, both of which were subject to optimization. The effect of finite delays[21,22,25] between scans was taken into account, and t_d was set to 2.1 s to account for an experimental acquisition time of 2 s and a delay of 0.1 s. The order parameter for intramethyl proton–proton interactions was set to 0.25,[21] and that for interactions from methyl to other protons was set to 0.85.[22] The rotational correlation times for bound and free ligand were estimated to be 40 and 0.2 ns, respectively, from the corresponding solvent-accessible surface areas.[51] Because protein signals at 0 ppm were selectively irradiated for the STD experiment, all methyl protons in the enzyme and its cofactor were assumed to be saturated. STD intensities of non-exchangeable protons on ligands were calculated only at a saturation time of 2 s. To evaluate the fit between experimental and calculated STD enhancements, the NOE *R*-factor[52] was calculated for normalized STD values.

Acknowledgements

This work has been supported by the Alberta Glycomics Centre, CIHR Regional Partnership Program (Saskatchewan) and the Natural Sciences and Engineering Research Council of Canada. S.A.D. was supported by a SHRF PDF fellowship. We thank Dr. Karin van Straaten for assistance with data processing. MD and CORCEMA-ST computations were performed on the Western Canada Research Grid, and we are grateful to Dr. N. Rama Krishna for providing the CORCEMA-ST program. M.B.P. is the recipient of a Studentship award from Alberta Innovates–Technology Futures and an Alexander Graham Bell Canada Graduate Scholarship from the Natural Sciences and Engineering Research Council of Canada.

Keywords: carbohydrates · CORCEMA-ST calculations · molecular dynamics · NMR spectroscopy · pyranose–furanose mutase · X-ray crystallography

[1] M. Zilbauer, N. Dorrell, B. W. Wren, M. Bajaj-Elliott, *Trans. R. Soc. Trop. Med. Hyg.* **2008**, *102*, 123 – 129.
[2] J. Kaldor, B. R. Speed, *Br. Med. J.* **1984**, *288*, 1867 – 1870.
[3] K. O. Poropatich, C. L. Fischer Walker, R. E. Black, *J. Health Popul. Nutr.* **2010**, *28*, 545 – 552.
[4] P. Guerry, F. Poly, M. Riddle, A. C. Maue, Y.-H. Chen, M. A. Monteiro, *Front. Cell. Infect. Microbiol.* **2012**, *2*, 7.
[5] M. A. Jones, K. L. Marston, C. A. Woodall, D. J. Maskell, D. Linton, A. V. Karlyshev, N. Dorrell, B. W. Wren, P. A. Barrow, *Infect. Immun.* **2004**, *72*, 3769 – 3776.
[6] D. J. Bacon, C. M. Szymanski, D. H. Burr, R. P. Silver, R. A. Alm, P. Guerry, *Mol. Microbiol.* **2001**, *40*, 769 – 777.
[7] A. Karlyshev, O. Champion, C. Churcher, J. Brisson, H. Jarrell, M. Gilbert, D. Brochu, F. St Michael, J. Li, W. Wakarchuk, I. Goodhead, M. Sanders, K. Stevens, B. White, J. Parkhill, B. Wren, C. Szymanski, *Mol. Microbiol.* **2005**, *55*, 90 – 103.
[8] A. V. Karlyshev, J. M. Ketley, B. W. Wren, *FEMS Microbiol. Rev.* **2005**, *29*, 377 – 390.
[9] M. B. Poulin, H. Nothaft, I. Hug, M. F. Feldman, C. M. Szymanski, T. L. Lowary, *J. Biol. Chem.* **2010**, *285*, 493 – 501.
[10] M. J. E. Sternberg, A. Tamaddoni-Nezhad, V. I. Lesk, E. Kay, P. G. Hitchen, A. Cootes, L. B. van Alphen, M. P. Lamoureux, H. C. Jarrell, C. J. Rawlings, E. C. Soo, C. M. Szymanski, A. Dell, B. W. Wren, S. H. Muggleton, *J. Mol. Biol.* **2013**, *425*, 186 – 197.
[11] P. Nassau, S. Martin, R. Brown, A. Weston, D. Monsey, M. McNeil, K. Duncan, *J. Bacteriol.* **1996**, *178*, 1047 – 1052.
[12] M. R. Richards, T. L. Lowary, *ChemBioChem* **2009**, *10*, 1920 – 1938.
[13] F. St Michael, C. M. Szymanski, J. Li, K. H. Chan, N. H. Khieu, S. Larocque, W. W. Wakarchuk, J. R. Brisson, M. A. Monteiro, *Eur. J. Biochem.* **2002**, *269*, 5119 – 5136.
[14] M. Soltero-Higgin, E. E. Carlson, T. D. Gruber, L. L. Kiessling, *Nat. Struct. Mol. Biol.* **2004**, *11*, 539 – 543.
[15] H. G. Sun, M. W. Ruszczycky, W. C. Chang, C. J. Thibodeaux, H. W. Liu, *J. Biol. Chem.* **2012**, *287*, 4602 – 4608.
[16] T. D. Gruber, W. M. Westler, L. L. Kiessling, K. T. Forest, *Biochemistry* **2009**, *48*, 9171 – 9173.
[17] D. A. R. Sanders, A. G. Staines, S. A. McMahon, M. R. McNeil, C. Whitfield, J. H. Naismith, *Nat. Struct. Biol.* **2001**, *8*, 858 – 863.
[18] K. Beis, V. Srikannathasan, H. Liu, S. W. B. Fullerton, V. A. Bamford, D. A. R. Sanders, C. Whitfield, M. R. McNeil, J. H. Naismith, *J. Mol. Biol.* **2005**, *348*, 971 – 982.
[19] S. K. Partha, K. E. van Straaten, D. A. R. Sanders, *J. Mol. Biol.* **2009**, *394*, 864 – 877.
[20] J. M. Chad, K. P. Sarathy, T. D. Gruber, E. Addala, L. L. Kiessling, D. A. R. Sanders, *Biochemistry* **2007**, *46*, 6723 – 6732.
[21] V. Jayalakshmi, N. R. Krishna, *J. Magn. Reson.* **2002**, *155*, 106 – 118.
[22] V. Jayalakshmi, N. R. Krishna, *J. Magn. Reson.* **2004**, *168*, 36 – 45.
[23] Y. Yuan, D. Bleile, X. Wen, D. Sanders, K. Itoh, H. Liu, B. Pinto, *J. Am. Chem. Soc.* **2008**, *130*, 3157 – 3168.
[24] Y. Yuan, X. Wen, D. A. R. Sanders, B. M. Pinto, *Biochemistry* **2005**, *44*, 14080 – 14089.
[25] X. Wen, Y. Yuan, D. A. Kuntz, D. R. Rose, B. M. Pinto, *Biochemistry* **2005**, *44*, 6729 – 6737.
[26] M. G. Szczepina, D. W. Bleile, B. M. Pinto, *Chem. Eur. J.* **2011**, *17*, 11446 – 11455.
[27] X. Yao, D. Bleile, Y. Yuan, J. Chao, K. Sarathy, D. Sanders, B. Pinto, M. O'Neill, *Proteins Struct. Funct. Bioinf.* **2009**, *74*, 972 – 979.
[28] W. Kabsch, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 125 – 132.
[29] B. W. Matthews, *J. Mol. Biol.* **1968**, *33*, 491 – 497.
[30] The Collaborative Computing Project Number 4, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1994**, *50*, 760 – 763.
[31] R. M. Keegan, M. D. Winn, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2008**, *64*, 119 – 124.
[32] G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1997**, *53*, 240 – 255.
[33] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 2126 – 2132.
[34] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 213 – 221.
[35] V. B. Chen, W. B. Arendall III, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66*, 12 – 21.
[36] M. Mayer, B. Meyer, *Angew. Chem.* **1999**, *111*, 1902 – 1906; *Angew. Chem. Int. Ed.* **1999**, *38*, 1784 – 1788.
[37] O. Trott, A. J. Olson, *J. Comput. Chem.* **2010**, *31*, 455 – 461.
[38] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, R. J. Woods, *J. Comput. Chem.* **2005**, *26*, 1668 – 1688.
[39] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, *Proteins Struct. Funct. Genet.* **2006**, *65*, 712 – 725.
[40] K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. González-Outeiriño, C. R. Daniels, B. L. Foley, R. J. Woods, *J. Comput. Chem.* **2008**, *29*, 622 – 655.
[41] E. Vanquelef, S. Simon, G. Marquant, E. Garcia, G. Klimerak, J. C. Delepine, P. Cieplak, F. Y. Dupradeau, *Nucleic Acids Res.* **2011**, *39*, W511 – W517.
[42] C. I. Bayly, P. Cieplak, W. Cornell, P. A. Kollman, *J. Phys. Chem.* **1993**, *97*, 10269 – 10280.
[43] P. Cieplak, W. D. Cornell, C. Bayly, P. A. Kollman, *J. Comput. Chem.* **1995**, *16*, 1357 – 1377.
[44] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, *J. Chem. Theory Comput.* **2008**, *4*, 435 – 447.

[45] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926 – 935.
[46] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*, 10089 – 10092.
[47] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577 – 8593.
[48] G. Bussi, D. Donadio, M. Parrinello, *J. Chem. Phys.* **2007**, *126*, 014101.
[49] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Dinola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684 – 3690.
[50] B. Hess, *J. Chem. Theory Comput.* **2008**, *4*, 116 – 122.
[51] V. V. Krishnan, M. Cosman, *J. Biomol. NMR* **1998**, *12*, 177 – 182.
[52] N. Rama Krishna, V. Jayalakshmi, *Prog. Nucl. Magn. Reson. Spectrosc.* **2006**, *49*, 1 – 25.

Received: October 11, 2013
Published online on December 2, 2013