Organic & Biomolecular Chemistry

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/obc

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

Computational and experimental study of O-glycosylation. Catalysis by human UDP-GalNAc polypeptide:GalNAc transferase-T2.

Hansel Gómez,[a,b] Raúl Rojas,[c] Divya Patel,[c] Lawrence A. Tabak,*[c] José M. Lluch,[a,b] and Laura Masgrau*[a]

Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX
DOI: 10.1039/b000000x

It is estimated that >50% of proteins are glycosylated with sugar tags that can modulate protein activity through what has been called the sugar code. Here we present the first QM/MM calculations of the full enzyme to study the O-glycosylation reaction catalysed by a retaining glycosyltransferase and combine our results with new experiments. In particular, we focus on human ppGalNAc-T2, which catalyses O-glycosylation and starts the biosynthesis of O-glycans. Importantly, we have characterized a hydrogen bond between the β-phosphate and the backbone amide group from the Thr7 of the sugar acceptor that promotes catalysis, and that we propose could be a general catalytic strategy used in peptide O-glycosylation by retaining glycosyltransferases. Focussing on ppGalNAc-T2 catalysis, other important substrate-substrate interactions have been identified, like that between the β-phosphate of UDP with the attacking hydroxyl group from the acceptor substrate and with the substituent at the C2‘ position of the transferred sugar. Our results support a front-side attack mechanism for this enzyme, with a barrier height of ~20 kcal/mol at the QM(M05-2X/TZVP//BP86/SVP)/ CHARMM22 level, in reasonable agreement with the experimental kinetic data. Experimental and *in silico* mutations show that transferase activity is very sensitive to changes in residues Glu334, Asn335 and Arg362. Additionally, our calculations for different donor substrates suggest that human ppGalNAc-T2 would be inactive if 2’-deoxy-Gal or 2’-oxymethyl-Gal were used, while UDP-Gal is confirmed as a valid sugar donor. Finally, the analysis herein presented highlight that both the substrate-substrate and the enzyme-substrates interactions are mainly concentrated on stabilizing the negative charge developing at the UDP leaving group as the transition state is approached, identifying this as a key aspect in retaining glycosyltransferases catalysis.

.

A. Introduction

O-glycans are responsible for a number of unique structural features in mucin glycoproteins and numerous membrane receptors,[1-3] and also impart resistance to thermal change and proteolytic attack in a number of diverse proteins.[4,5] Moreover, O-linked carbohydrate side chains function as ligands for receptors (e.g. in host-microbial interactions,[6] lymphocyte and leukocyte homing[7,8]) and as signals for protein sorting.[9-13] It has been estimated than >50% of proteins are glycosylated and that this sugar tag can modulate their activity acting like an analog switch.[14]

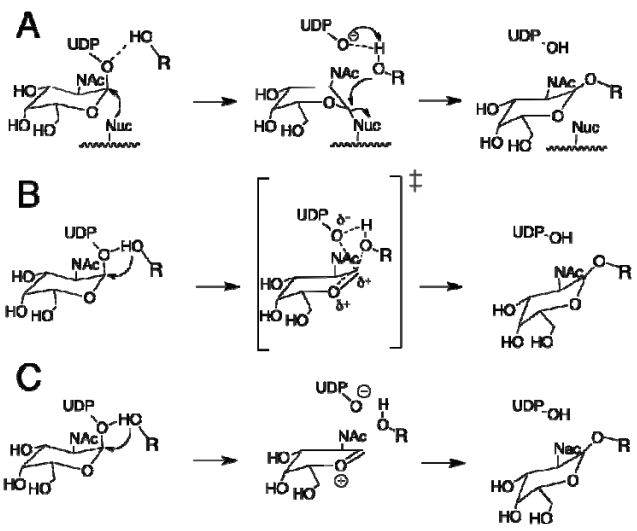
The enzymes UDP-N-acetylgalactosamine polypeptide: N-acetylgalactosaminyl-transferases (ppGalNAcTs, EC 2.4.1.41) catalyze the transfer of GalNAc from the sugar donor UDP-GalNAc to serine and threonine residues, in what is the first committed step in mucin biosynthesis to form the Tn antigen (GalNAcα1-O-Ser/Thr).[15] Subsequent extension of O-glycan formation proceeds step-wise.[16] Several experiments indicate that there is a hierarchical addition of core GalNAc residues to apomucins, implying that the complete glycosylation of certain substrates is dependent on the coordinated action of multiple ppGalNAcTs.[15] Up to 20 members have been identified in humans for this large and evolutionarily conserved family (family 27 in the CAZy[17] database). Thus, understanding the catalytic mechanism of ppGalNAcTs would have important practical implications and would shed light into the O-glycosylation process. More generally, the mechanism for retaining glycosyl transfer stereospecificity has been a matter of debate in glycobiology for the last decades[18] and is hampering the rational design of specific drugs/inhibitors for this class of enzymes.

Two main mechanisms have been proposed in the literature for retaining glycosyltransferases. Initially, and by analogy with retaining glycosidases[19-20] (despite the lack of evolutionary relatedness[21]) a double-displacement mechanism was proposed, with formation and subsequent cleavage of a covalent glycosyl–enzyme (CGE) intermediate involving a nucleophilic residue of the enzyme (Scheme 1A). Although some experiments might support the existence of a CGE,[22,23] so far, there is no conclusive evidence for it. In a recent computational study we showed that,

for bovine α-1,3-galactosyltransferase (α3GalT, GT6 family), the formation of a CGE is consistent with the experimental kinetic data, even though its formation was calculated to be quite endoergic, the minima was not very stable and it may require the presence of the acceptor substrate to be formed at kinetically relevant rates.[24] Others have subsequently confirmed that, computationally, this CGE can be characterized for this enzyme.[25] However, as new crystal structures of retaining glycosyltransferases have been solved, it has become apparent that most retaining GTs do not present a well-positioned residue in the active site to act as the nucleophile.[18] In fact, only family GT6 glycosyltransferases have one. Therefore, the double-displacement mechanism might not be a universal mechanism for retaining glycosyltransferases, and alternative mechanisms have been advocated. The most favoured mechanism is a front-side attack of the acceptor nucleophile, that is located on the same side as the leaving group, resulting in the formation of oxocarbenium species that could correspond to a single transition state (Scheme 1B) or to an oxocarbenium-phosphate short-lived ion pair intermediate, with the two corresponding transition states (Scheme 1C). The latest theoretical and experimental work on retaining GTs[24,26-32] are giving support to this front-side attack mechanism for those retaining GTs where no good nucleophile is suitably positioned to form the CGE. In fact, even in the case of α3GalT we were able to describe a front-side attack mechanism with a comparable potential energy barrier to the one calculated for the double-displacement.[24]

In the case of ppGalNAc-T2 (GT-A fold for the catalytic domain), the available crystallographic structures show that the nearest acidic residues that might function as nucleophiles in a double-displacement mechanism (i.e. Asp224 of the DXH motif binding Mn^{2+} and Glu334) are ~7 Å away from the β-phosphate oxygen. Consequently, a double displacement looks unlikely or would require a large conformational change. The latest is not observed on the timescale of the MD simulations performed on the Michaelis complex of human ppGalNAc-T2 by Milac et al.[33] On the other hand, their results are more consistent with a front-side mechanism, since the distance between the glycosidic oxygen and the nucleophilic hydroxyl group is about 3 Å and is maintained nearly constant during the simulation, which would at least structurally be consistent with a nucleophilic role of the acceptor.[33] These results, together with the available X-ray structures and site-directed mutagenesis data, point to a single-displacement mechanism as the most likely.

In our previous work on retaining glycosyltransferases,[24,30,31] we have emphasized the importance that intra- and inter-substrate interactions have in catalyzing this reaction. In most cases, these interactions involve the β-phosphate group of UDP with the O2' hydroxyl of the transferred monosaccharide, with hydroxyl groups of the acceptor molecule and, most importantly, a hydrogen bond with the hydrogen of the attacking hydroxyl. We have also noted that the particular interactions used by each enzyme:substrate system depend on the chemical identity of the substrates and on their relative binding orientation in the active site (which in turn depends on the specificity of the new glycosidic linkage). Thus, the present study is the first one to analyze how these substrate-substrate interactions act in the case of transferring GalNAc to a peptide acceptor.



Scheme 1. Proposed mechanisms for the retaining GTs. (A) Double-displacement mechanism with formation of a covalently bound glycosyl-enzyme intermediate. Front-side attack with (B) an oxocarbenium ion-like transition state or (C) a short-lived oxocarbenium-phosphate ion pair intermediate.

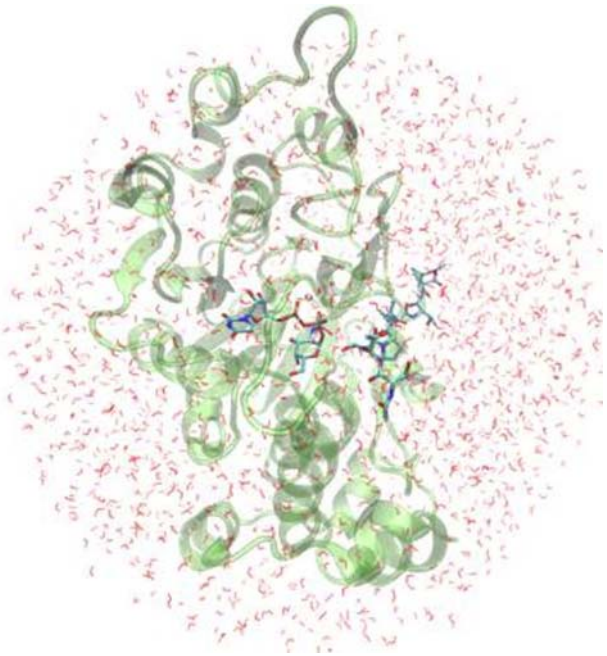
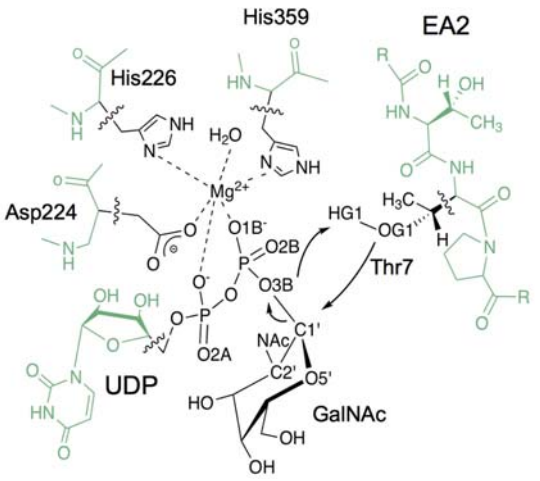


Figure 1. Model system of ppGalNAc-T2: EA2 used in the QM/MM calculations.

We present here a combined computational and experimental work on the catalytic mechanism of human ppGalNAc-T2 and take it as a model of O-glycosylation by retaining glycosyltransferases. Hybrid quantum mechanical/molecular mechanical (QM/MM) calculations on the full enzyme have been used (Figure 1). Key factors supporting catalysis have been identified and, when possible, experiments have been performed to test the computational findings.



Scheme 2. QM/MM partition considered in the present work. QM (MM) atoms are depicted in black (grey). Wavy lines indicate the boundary between the QM and MM regions. The arrows indicate the distances considered in the reaction coordinates and the atoms involved are labeled.

B. Computational Methods

An initial fully solvated ternary complex modelled by Milac et al.[33] was used as starting point in the reactivity study. This ternary complex had been built by taking the coordinates of the catalytic domain of ppGalNAc-T2 and of the acceptor peptide (EA2; sequence PTTDSTTPAPTTK) from the PDB Code 2FFU,[34] and modelling the donor substrate (UDP-GalNAc) in the active site using as a template the human GalNAcT-10 (PDB Code 2D7I[35]), which contains hydrolyzed UDP-GalNAc. For the present study, all water atoms in this solvated ternary complex more than 30 Å away from the anomeric centre ($C1'_{GalNAc}$) were deleted. This procedure resulted in a system with ~12630 atoms, including ~2170 TIP3P water molecules (see Figure 1). The Mn^{2+} ion present in the original X-ray structure was modelled by the computationally more convenient Mg^{2+}, an approach that has been validated in previous studies of related systems.[30,36] Moreover, some experiments have shown that GalNAc-Ts can also be active with Mg^{2+} and other divalent cations.[37] The system was then divided into a QM and an MM zone (Scheme 2). The charge of the QM region was -1 and included 80 atoms: those from the GalNAc ring, the side chain of Thr7 from the acceptor substrate (peptide EA2), Mg^{2+} and its first coordination sphere (phosphate groups from UDP and the side chains of residues Asp224, His226, His359 and one crystallographic water). Five hydrogen link atoms were added to treat the QM/MM boundary with the charge shift model.[38,39] An electronic embedding scheme[40] was adopted in the QM/MM calculations and no cutoffs were introduced for the nonbonding MM and QM/MM interactions. Notice that only residues and water molecules within 15 Å of the anomeric centre (~2080 atoms) were allowed to move during the QM/MM calculations.

This model of the Michaelis complex was then submitted to a NVT QM(SCC-DFTB[41,42])/MM(CHARMM22[43,44]) molecular dynamics simulation using the dynamics module within ChemShell.[45] The SHAKE procedure[46] was applied at every step for the O-H bonds of the water molecules. A 10 ps MD equilibration run was followed by 80 ps of production MD. Two randomly selected snapshots from this simulation were used in QM/MM geometry optimizations with QM = (BP86[47-50]/SVP[51]), a method that we have successfully applied before to study other glycosyltransferases.[24,30]

Starting at these optimized reactant structures, reaction paths were scanned by performing constrained QM(BP86/SVP)/CHARMM optimizations along suitably defined reaction coordinates in steps of 0.2 Å. This provided us with starting structures for subsequent full optimization of transition states and products. Frequency calculations were performed for the QM region to confirm that the optimized TS structures are indeed characterized by one imaginary frequency and a suitable transition vector. Additional single-point energy calculations were carried out at the M05-2X[52]/TZVP[53] level which has proven to properly describe retaining GT systems.[30,36,54] For the purpose of comparison, additional single-point energies were calculated at the BP86/TZVP, B3LYP[47,48,55-57]/SVP and B3LYP/TZVP levels of theory.

The electrostatic stabilization provided by different residues to the QM(M05-2X/TZVP)/CHARMM energy was examined by setting their point charges to zero in additional single-point energy calculations. A Natural Bond Orbital (NBO) analysis[58-61] was also performed for some of the stationary points using the NBO program v3.1[62] included in Gaussian09.[63]

All QM/MM calculations were performed with the modular program package ChemShell using TURBOMOLE[64] or Gaussian09 at the DFT level (BP86, B3LYP and M05-2X functionals) or MNDO[65] at the SCC-DFTB level. MM energies and gradients were retrieved from DL_POLY,[66] using the CHARMM force field. Energy minimizations were done with the low-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm[67,68] and the TS searches were performed with the microiterative TS optimizer that combines L-BFGS and the partitioned rational function optimizer (P-RFO).[69,70] Both L-BFGS and P-RFO algorithms are implemented in the HDLCopt[71] module of ChemShell.

C. Results and Discussion

Our goal here is to determine if retaining O-linked glycosylation can be achieved via the controversial front-side attack mechanism, as there is no strong nucleophile in the vicinity of the anomeric center in ppGalNAc-T2, and to reveal the factors that allow for it.

A. Catalytic Mechanism

The reaction mechanism was modeled by using both a double (RC = [$d(O3B_{UDP}$–$C1'_{\alpha\text{-GalNAc}})$ - $d(OG1_{T7}$–$C1'_{\alpha\text{-GalNAc}})$]) or a triple (RC = [$d(O3B_{UDP}$–$C1'_{\alpha\text{-GalNAc}})$ - $d(OG1_{T7}$–$C1'_{\alpha\text{-GalNAc}})$ - $d(HG1_{T7}$–$O3B_{UDP})$]) reaction coordinate, which are the ones we have used in our studies of LgtC[30] and α3GalT,[24,31] respectively. In both cases, the calculated potential energy profiles were very similar.

Similar results were also obtained for the two frames from the molecular dynamics simulations considered in the QM/MM reactivity studies. For simplicity, the results presented in the main text will refer to frame 1 (see Supporting Information, SI, for the results of frame 2).

The calculated potential energy profile is depicted in Figure 2, altogether with the evolution of key distances along this front-side attack mechanism.

As can be seen in the Figure, a potential energy path reproducing the front-side attack mechanism has been obtained and with a reasonable energy maximum at ~16 kcal/mol (a phenomenological free energy barrier of 17.3 kcal/mol can be derived from the experimental k_{cat} value of 3.7 s^{-1} at 310 K[34]). The distances depicted show that the reaction starts with the breakage of the UDP-GalNAc glycosidic bond. In fact, the energy required to break this linkage accounts for nearly all the potential energy barrier associated to the whole process (See SI, Figure S1, for the WT enzyme with different substrates and for some mutant enzymes, see below). The potential energy surface is quite planar at the region corresponding to the potential energy maximum for the transferase reaction (with 2.8 Å < d(O3B$_{UDP}$–C1'$_{\alpha\text{-GalNAc}}$) < 3.2 Å and 2.6 Å > d(OG1$_{T7}$–C1'$_{\alpha\text{-GalNAc}}$) > 2.2 Å), which could seem to indicate a S_Ni-like mechanism (See Scheme 1). However, no ion-pair intermediate (IP) could be characterized so that the S_Ni term may be more appropriate. Notice, though, that the differences between these two alternatives of front-side attack mechanism can be very subtle in this kind of potential energy surfaces. In fact, the topology of this surface conditioned that we were also unable to find the corresponding transition state. In what follows then, for the QM(BP86/SVP)/MM(CHARMM22) level we will be considering the TS guess (i.e. ?TS^i; structure corresponding to the maximum potential energy value along the RC reaction coordinate) as the effective TS for analysis.

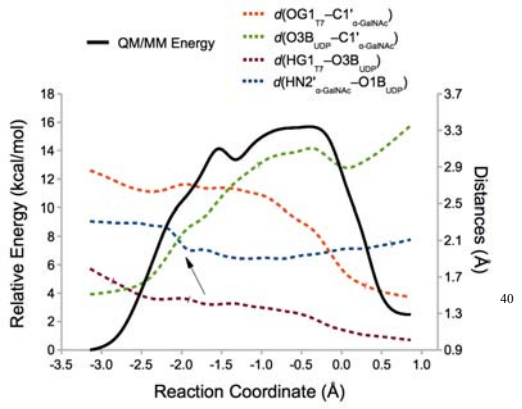


Figure 2. QM(BP86/SVP)/MM(CHARMM22) energy profile for the front-side attack mechanism in ppGalNAc-T2. Reaction coordinate (RC) = [d(O3B$_{UDP}$–C1'$_{\alpha\text{-GalNAc}}$) - d(OG1$_{T7}$–C1'$_{\alpha\text{-GalNAc}}$) - d(HG1$_{T7}$–O3B$_{UDP}$)]. The variation of several interatomic distances involved in the reaction is also depicted. HN2' is the amine hydrogen of the NAc group. The arrow indicates the moment when the NAc group from the α-GalNAc gets properly oriented to favour the O3B$_{UDP}$–C1'$_{\alpha\text{-GalNAc}}$ bond-breaking process (see main text).

Table 1. QM/MM potential energy barriers and reaction energies (in kcal/mol)[a] at different levels of theory for the front-side attack mechanism for frame 1.

	BP86		B3LYP		M05-2X	
	SVP	TZVP	SVP	TZVP	SVP	TZVP
R	0.0	0.0	0.0	0.0	0.0	0.0
?TS^i	15.7	10.7	19.3	14.0	26.3	19.8
P	-2.3	-0.5	1.8	-1.1	3.6	0.3

[a] The calculations were carried out on the corresponding QM(BP86/SVP)/MM(CHARMM22) geometries of reactants (R), transition state guess (?TS^i) and products (P).

At the QM(SCC-DFTB)/MM(CHARMM22) level a TS was easily identified with an imaginary frequency consistent with the reaction under study (see SI Tables S1-S2 and Figures S2-S3 for the structural and energetic results). An estimation of the free energy profile was also done by umbrella sampling calculations at the QM(SCC-DFTB)MM(CHARMM22) level of theory (SI, Figure S3). A qualitative comparison with the potential energy barrier suggests that entropic effects might be relatively small for the present system, as was the case in our previous study of LgtC.[30]

The evolution of distances along the reaction depicted in Figure 2 shows that the reaction starts readily with the breakage of the O3B$_{UDP}$–C1'$_{\alpha\text{-GalNAc}}$ bond, as the HOG1$_{T7}$–O3B$_{UDP}$ hydrogen bond (which we have shown before to be essential in assisting the leaving group departure[24]) is already present in the reactants.

The QM/MM energy barriers and reaction energies calculated at different levels of theory for frame 1 are shown in Table 1. The potential energy barrier at the reference level (QM = M05-2X/TZVP) is 19.8 kcal/mol, again in qualitative agreement with the experimentally derived one, suggesting that the front-side attack is actually feasible. The reaction turns out to be almost isoergic (0.3 kcal/mol) at this level of theory, but slightly exoergic at others; we are not aware of any experimental data on reaction energies to compare with for GalNAc-Ts.

The structures of the stationary points are depicted in Figure 3 (key distances and atomic charges are listed in SI, Table S4). Some common trends are observed when comparing to the results obtained in our previous work on retaining glycosyltransferases.[24,30,31] Notably, the ?TS^i is highly dissociative (d(O3B$_{UDP}$–C1'$_{\alpha\text{-GalNAc}}$) = 3.10 Å), which explains why a nucleophilic substitution by the same side is possible, and proton transfer from the attacking nucleophile (OG1$_{T7}$) to the leaving group oxygen (O3B$_{UDP}$) takes place very late in the reaction allowing then the final formation of the new glycosidic bond. The dissociative character of the TS is also reflected on the positive charge development at the C1' and O5' atoms (Δq(C1'+O5') = 0.25 a.u) in going from the R to the ?TS^i. The closest residue on the β-face of the donor sugar substrate that could stabilize this positive charge is Ala307, whose carbonyl group is 4.09 Å from the anomeric centre at the reactants and, mainly as a result of the change in puckering of the ring, gets closer at the ?TS^i (3.12 Å). On the other side, the Δq(O3B$_{UDP}$) = -0.25 a.u., and different interactions are observed that could favour this increment of the negative charge between R and ?TS^i. All this will be analysed in the following sections (only for frame 1).

B. Inter- and intra- substrate interactions.

In our previous work on other retaining glycosyltransferases (i.e. α1,3-GalT[24] and LgtC[30]) we have characterized a series of inter- and intra-substrate interactions that seem to be vital to explain the catalytic efficiency of this family of enzymes. These interactions were mainly the hydrogen bonds of the β-phosphate of UDP with the C2' hydroxyl group of α-Gal, or with hydroxyl groups of the acceptor substrate. In order to identify equivalent interactions in the case of human ppGalNAc-T2 we have performed an NBO analysis by considering the reactants and the transition state guess. Notice that for the reaction catalysed by ppGalNAc-T2, the transferred sugar has an N-acetyl group (NAc) at position C2' instead of an hydroxyl group, and the acceptor substrate is a peptide instead of another sugar like it was the case in our previous studies.

As described for LgtC and α1,3-GalT, the interactions between molecular orbitals involving the leaving group oxygen $O3B_{UDP}$ and the attacking nucleophilic group from the acceptor substrate (here $(HG1 - OG1)_{T7}$) are the most significant ones (SI, Table S6, and Figure 4A,B). Secondarily, the interactions between the orbitals of $OG1_{T7}$ and an antibonding molecular orbital of $(C1'\text{-}O5')_{GalNAc}$ are also contributing to the stabilization of the transition state. Interestingly, a new inter-substrate interaction involving the leaving group occurs in this system: a hydrogen bond with the backbone amide of the attacking residue (i.e. Thr7 from peptide EA2, Figure 4C,D). This interaction could be a general strategy used by retaining GTs transferring the monosaccharide to a peptide acceptor to help stabilize the transition state. For sugar acceptors, an analogous interaction involving the hydroxyl group neighbouring the attacking oxygen was also described in LgtC.[30] Notice from Figure 4A,C that the two inter-substrates interactions described for ppGalNAc-T2 were already present in the Michaelis complex so that the two substrates are bound in the active site optimally oriented for the specificity of the reaction to be catalysed by the enzyme. Finally, another distinctive trait of ppGalNAc-T2 is the interaction between the 2'-N-acetyl group from the donor substrate and the UDP leaving group, which is only present at the transition state (Figure 4E,F). In the reactants, NAc is interacting with Asp224, one of the residues coordinating the metallic cofactor (i.e. Mg^{2+}), an interaction that is maintained throughout the 40 ns of MD simulation performed by Milac et al. on the Michaelis complex.[33] Along the reaction, and due to the change in the sugar ring puckering (i.e.; from a distorted $^4C_1$ (puckering parameters ϕ = 137.3°, θ = 15.0°) to a ^{4}E-like ring conformation (ϕ = 242.9°, θ = 31.3°) in R and ?TS^i, respectively), NAc gets reoriented and forms a stabilizing interaction with the UDP leaving group. The distance $d(HN2'_{\alpha\text{-GalNAc}}\text{–}O1B_{UDP})$ gets shorter in the process, from 2.31 Å in the Michaelis complex to 1.91 Å in the ?TS^i and, as a result, $N2'_{\alpha\text{-GalNAc}}$ gets hydrogen-bonded to $O1B_{UDP}$ (See Figure 2).

It is known that some hydrolases like OGA (a glycosidase involved in O-GlcNAcylation cycling) employ the NAc group from GlcNAc itself as a nucleophile to cleave the monosaccharide from serine/threonine.[72] However, in the case of glycosyltransferases a catalytic role by a NAc group from the donor substrate has only very recently been described for an inverting glycosyltransferase. In that work, Tvaroška et al.[73] performed a QM/MM study on O-GlcNAc transferase (i.e.; uridine diphospho-N-acetylglucosamine:poly-peptide β-N-acetylaminyltransferase, OGT) and found a substrate-assisted mechanism by the NAc group. More specifically, they described a rotation of the C2'-N2' bond that approaches the HN2' proton to the oxygen of the breaking glycosidic linkage, thus stabilizing the leaving group negative charge and assisting its departure.

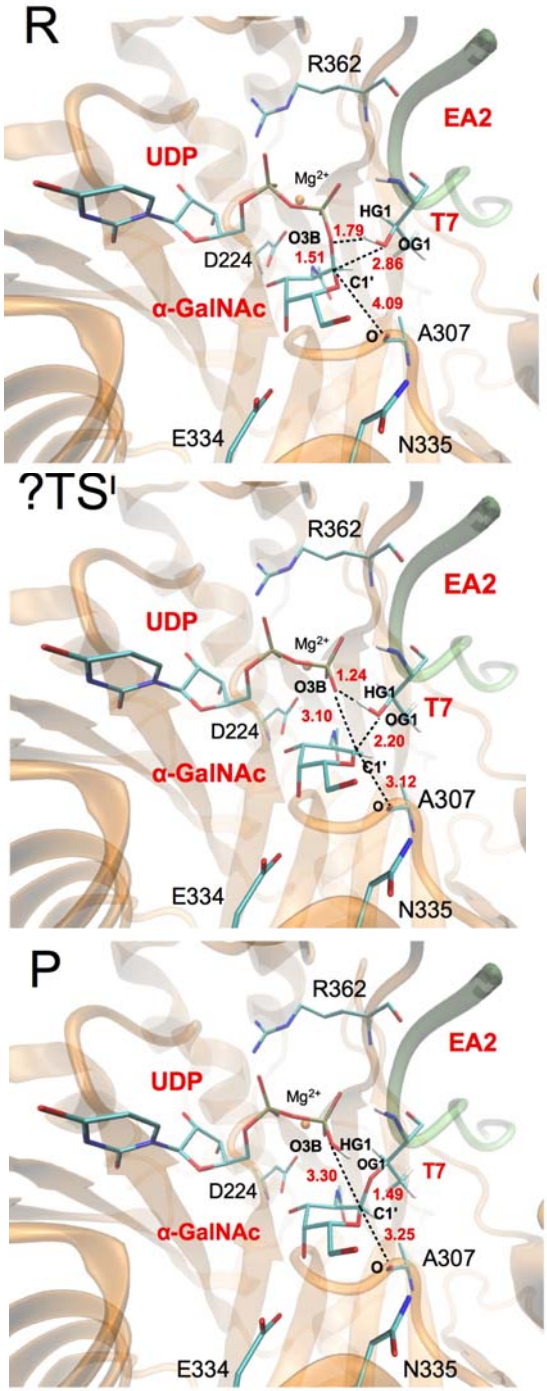


Figure 3. QM(BP86/SVP)/MM(CHARMM22) optimized reactants (R), transition state guess (?TS^i) and products (P) for the front-side attack mechanism. The donor and part of the acceptor substrate, together with some relevant residues in the active site, are represented as sticks. Selected distances (in Å) are also indicated.

The authors hypothesize that OGT could require a mechanism like this to account for the lack of stabilization provided by an absent metal cofactor. This does not appear to be the case for ppGalNAc-T2, but still the NAc group appears as important in catalysis.

To shed more light on the relevance of the NAc group in catalysis, alternative donor substrates were considered *in silico*.

We substituted this NAc group by OH, H or OCH_3 in the original Michaelis complex, which corresponds to consider UDP-Gal, 2'-deoxy-Gal and 2'-oxymethyl-Gal as donor substrates. We assumed that no significant changes in the binding occur and we focused solely in the catalytic process itself. The effects of this functional group substitution on the energy and reaction barriers are summarized in Table 2.

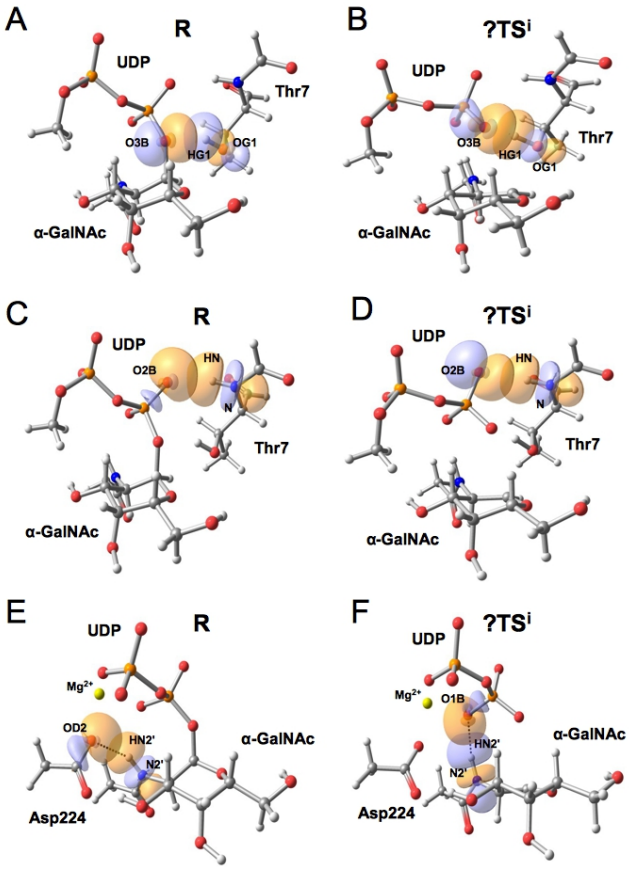


Figure 4. Relevant molecular orbitals interactions between the substrates in GalNAc transfer by ppGalNAc-T2 according to a NBO analysis. These interactions involve: (**A**) $O3B_{UDP}$ and the incoming $OG1_{T7}$ in the reactants, R, and (**B**) in the front-side attack transition state guess, ?TS^i; (**C**) the backbone amide group of Thr7 in R and (**D**) in the ?TS^i; (**E**) the NAc group of the donor substrate with Asp224 in R and (**F**) with UDP in the ?TS^i. For clarity, just a fraction of the QM atoms is shown. The corresponding second order interaction energies are depicted in SI Table S6.

Table 2. QM(M05-2X/TZVP//BP86/SVP)MM(CHARMM22) potential energy barriers ($\Delta V^{\ddagger}$)[a] and reaction energies (ΔV_R), in kcal/mol, for ppGalNAc-T2 with different donor substrates.

	UDP-GalNAc	UDP-Gal	UDP-2'-deoxy-Gal	UDP-2'-oxy-methyl-Gal
$\Delta V^{\ddagger}$	19.8	20.9	26.5	28.5
ΔV_R	0.3	-0.3	8.6	-2.2

[a] $\Delta V^{\ddagger}$ calculated using a transition state guess.

Using UDP-Gal as a donor substrate results in a potential energy barrier very similar to the one obtained with the original UDP-GalNAc substrate. Inspection of the structures shows that in the Michaelis complex the 2'-hydroxyl group of Gal is also predominantly interacting with Asp224, and that it gets reoriented along the reaction, thus behaving similarly to the NAc group in UDP-GalNAc (SI, Figure S4). Interestingly, in the case of LgtC and α1,3-GalT, which use UDP-Gal as donor substrate, the interaction of the 2'-hydroxyl group with UDP was already present in the reactants (SI, Figure S5). The NAc-Asp224 interaction in the reactants for ppGalNAc-T2 prevailed during a 100 ps QM(SCC-DFTB)/MM(CHARMM22) MD simulation of the Michaelis complex. Altogether, our results suggest that human ppGalNAc-T2 may be able to transfer Gal to the peptide EA2, which is consistent with the experimental results obtained for another acceptor peptide (i.e. Muc2; sequence PTTTPISTTTMVTPTPTPTC).[74] In that work, the V_{max} values corresponding to the transfer of Gal-NAc and Gal were estimated in 46.1 and 79.9 pmol/min, respectively, which implies a difference of less than 1 kcal/mol between the two donor substrates. Moreover, the authors concluded that giving the relatively small difference between the K_M values for UDP-GalNAc and UDP-Gal (10 and 27 μM, respectively), UDP-Gal might actually be a naturally relevant substrate of ppGalNAc-T2. In that sense, notice that even if our results predict a slightly higher difference (energy barrier is ~1 kcal/mol higher for UDP-Gal), it falls within the order of error that could be expected for the methods used.

A bigger effect is observed for the other two alternative donor substrates. As can be seen in Table 2, the energy barrier increases by ~7 and ~9 kcal/mol for 2'-deoxy-Gal and 2'-oxymethyl-Gal, respectively, and also the reaction energies are more affected. These results support the idea that supressing the interaction between the 2'-NAc (OH) and UDP in UDP-GalNAc (UDP-Gal) would significantly disrupt catalysis in ppGalNAc-T2. According to our theoretical results, negligible or no detectable residual activity should be expected when 2'-deoxy-Gal or 2'-oxymethyl-Gal are used as donor substrates. Unfortunately, no experimental data can be provided to test this hypothesis as these compounds are not available.

C. Enzyme-substrates interactions; key enzyme residues.

Since the intermediates of the reactions catalyzed by retaining glycosyltransferases are significantly charged, the electrostatic stabilization of the transition state provided by the enzyme residues is expected to be significant. To assess these contributions in the present system, we carried out an analysis for the residues in the active space by switching off the charge of the residue and recalculating the QM/MM interaction energy. The analysis identifies four residues displaying a significant effect: Arg362, Glu334, Ala307 and Trp331, for which electrostatic stabilizations energies of 18.6, 11.5, 2.8 and 2.3 kcal/mol were estimated, respectively, for the ?TS[i] as compared to the reactants (QM = M05-2X/TZVP). The most stabilizing residue (i.e. Arg362) interacts with UDP (Figure 3), as we described in our previous studies of LgtC and α1,3-GalT, and participates in the conformational change that accompanies the binding of the donor substrate. This confirms the key role that the stabilization of the negative charge on the leaving group has in the catalytic efficiency of retaining glycosyltransferases. Arg362 is also involved in acceptor binding via a hydrogen bond that gets slightly elongated along the reaction between its carbonyl backbone and the side chain of Thr6. On another hand, Glu334 (which is located on the β-face of the sugar ring and is negatively charged), is hydrogen-bonded to the donor substrate and is also involved in substrate binding (See Figure 3). The carbonyl group of Ala307 is suitably positioned to stabilize the positive charge development in $C1'_{\alpha\text{-GalNAc}}$ as reflected in the 2.83 kcal/mol of stabilization that it provides. Tyr331 is another example of interaction with the leaving group, in this case through a hydrogen bond with $O3B_{UDP}$.

Interestingly, our analysis would suggest that Asn335, a residue lying on the β-face of the donor sugar substrate that could be suggested to be the putative nucleophile in a double-displacement mechanism in ppGalNAc-T2, does not have a very significant effect on the stabilization of the oxocarbenium species. This is not surprising as it is located at a distance of $d(OD1_{N335}–C1'_{\alpha\text{-GalNAc}})$ = 7.05 Å in the optimized Michaelis complex (See Figure 3) and of 6.96 ± 0.43 Å during the simulations of this complex performed by Milac et al.[33] Moreover, the carbonyl side chain of Asn335 is pointing away from the anomeric carbon, whereas the amide nitrogen is hydrogen bonded to the Ala307 backbone carbonyl.

D. Experimental and *in silico* mutants.

Once the most important residues of the enzyme (from the catalytic point of view) were identified, several mutated forms of ppGalNAc-T2 were tested *in silico* and/or experimentally. The position 335 was also considered, given its potential relevance in catalysis upon mutation. Conservative mutations, which are presumed to preserve the structural role of the residue while targeting the chemistry in question, were applied in most cases.

Recombinant ppGalNAc-T2, mutated at positions E334Q, N335A, N335D, N335H, N335S and R362K, were expressed (see Experimental Section, Table S7 and Figure S8 in SI). The ability of these mutated forms of ppGalNAc-T2 to transfer GalNAc to EA2 peptide was tested *in vitro* (Figure 5) and the results were compared to *in silico* determinations.

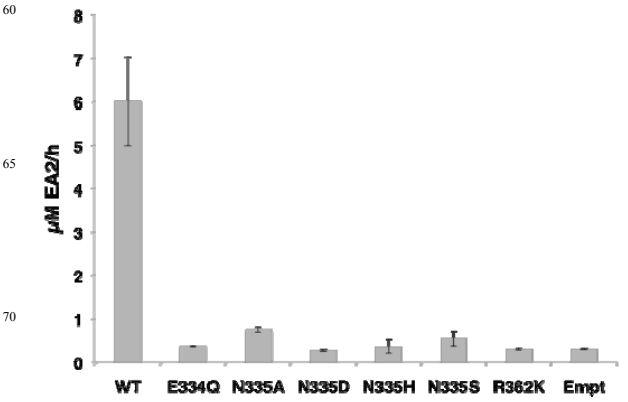


Figure 5. UDP-GalNAc transfer activity by wild type and mutant human ppGalNAc-T2 onto the EA2 peptide. Values are the average of two experiments run by triplicate. WT: wild-type ppGalNAc-T2 and Empt: pIMFK4 empty vector.

Table 3. QM(M05-2X/TZVP//BP86/SVP)MM(CHARMM22) potential energy barriers ($\Delta V^{\ddagger}$)[a] and reaction energies (ΔV_R), in kcal/mol, for wild-type (WT) ppGalNAc-T2 and considered mutants with UDP-GalNAc and EA2 as substrates.

	WT	R362K	E334Q	N335A	N335D
$\Delta V^{\ddagger}$	19.8	23.1	27.1	20.6	12.1
ΔV_R	0.3	0.3	3.7	-0.3	-4.6

[a] $\Delta V^{\ddagger}$ calculated using a transition state guess.

Models of R362K, E334Q, N335A and N335D were also built *in silico* and the energy profile for the front-side attack mechanism was calculated to assess the effect of such mutations in catalysis. Again it is important to highlight that our models of the mutants were built by just replacing the side chain of the original residue, that is, with the purpose of evaluating the effect of the mutations in the catalytic mechanism itself, assuming no significant structural perturbations of the enzyme and a negligible effect on the binding of the substrates. These may not be very good assumption for some of the mutants (even if conservative mutations have been done), since we are considering residues that are directly implicated in the binding and/or could also have a structural role in the active site. However, our goal was to estimate if the catalytic performance of the mutants can be explained by only considering the role of the specific residues in the reaction, while important inconsistencies with the experimental kinetic results may indicate that the overall structure of the active site and/or the binding of the substrates is also affected.

The potential energy barrier and reaction energies associated to these mutants are given in Table 3. The corresponding potential energy profiles were equally planar at their maximum (SI, Figure S6) and, therefore, we did not perform a proper TS search but used the maximum of the potential energy profiles as a TS guess for the analysis.

As can be seen in Table 3, for the mutant R362K we calculated an increase of ~3 kcal/mol in the potential energy

barrier, which is quite significant taking into account that a lysine in such position would still stabilize the developing negative charge on the leaving group, although to a less extent (SI, Figure S7C). The closer distance between Arg362 and the UDP leaving group can explain the differences found (i.e.; $d(NH2_{R362}–O2A_{UDP})/(d(NZ_{K362}–O2A_{UDP})$ = 3.23/4.97 Å and 2.72/4.62 Å in R and ?TS[i], respectively). An increment of ~3 kcal/mol in the energy barrier would imply less than 0.04 % of residual activity. This is in agreement with the experimental result we have obtained for this mutant, for which no significant transferase activity has been measured. A similar mutation (R365K) was reported for bovine α3GalT, resulting in a small variation of K_M and an bigger effect on k_{cat} but maintaining enzyme activity.[75] Therefore, ppGalNAc-T2 seems to be more sensitive to a change at this position.

An even higher effect is obtained for the less conservative mutation E334Q (energy barrier increase by ~7 kcal/mol). In the reactants, the carboxylic group of Glu334, located on the β-face of α-GalNAc, is 6 Å away from the anomeric carbon; yet mutating Glu to Asn significantly reduces the electrostatic stabilization role in catalysis of position 334 (Figure 3 and SI, Figure S7A). Moreover, this is a key residue in the binding of the donor substrate via a hydrogen bond to the OH4 of UDP-GalNAc, a recurrent interaction in retaining GTs. Therefore, mutation of this residue will probably affect both the K_M for the donor substrate and the k_{cat} values. Moreover, since retaining GTs bind their substrates in an ordered and interdependent way, which have also been certified in the case of GalNAcTs,[76] an increase in the K_M value for the acceptor substrate could also be expected. In fact, the experimental E334Q mutant in ppGalNAc-T2 renders the enzyme inactive for the transfer reaction (Figure 5). In the case of murine GalNAcT-1, E319Q (being Glu319 the equivalent of Glu334 in human ppGalNAc-T2) exhibits a residual activity of 0.04 %,[77] and this is also consistent with our findings.

For the *in silico* mutant N335A, the energy barrier remains practically unaffected (~1 kcal/mol higher, within the order of error of the methods, Table 3), which is consistent with our expectations since Asn335 was not found to be an important residue in terms of electrostatic stabilization (SI, Figure S7B). Mutations to alanine in the equivalent asparagine residue in the murine isoform (i.e. ppGalNAcT-1) just had a little effect on catalysis,[77] consistent with our predictions. However, the experimental data obtained here for the N335A mutation in GalNAc-T2 shows no significant transferase activity for the mutant. In fact, all the recombinant mutants tested at this position (N335A, N335D, N335H and N335S) are enzymatically inactive.

The other mutation at position 335 that was tested *in silico* was N335D. In that case, a drop in the energy barrier of ~7 kcal/mol was obtained. This suggests that having a negatively charged residue on the β-face of the donor sugar substrate would turn 335 into a key position, even if it is too far away from the anomeric center to participate in a double-displacement mechanism ($d(OD1_{N335}–C1'_{α\text{-}GalNAc})$ = 7.05 Å in reactants). According to our *in silico* model, the presence of a strong nucleophile like an Asp at this position would facilitate the leaving group departure, basically because of a better stabilization of the positive charge on the α-GalNAc ring (SI, Figure S7B) and, more importantly, would also delay the nucleophilic attack of the incoming hydroxyl group (SI, Figure S6A). The latter could lead to an increase in the probability of hydrolysis, thus, competing with the transferase activity. As mentioned, though, experimentally the N335D mutant does not present transferase activity. The disagreement between our *in silico* results for position 335 and the experimental ones, most likely indicate that the mutation provokes significant changes in the structure or the mechanism that are not captured by our present model. The modification of similar residues in the β face of the sugar ring has also lead to unexpected results for other retaining GTs like LgtC, where formation of a glycosyl-enzyme complex with a neighboring Asp residue has been reported.[22] A deeper understanding of the effect of such mutations would require much more extensive computational work and, probably, also more experimental data; but this is out of the scope of the present study. In any case, it is clear from the experimental results that ppGalNAc-T2 transferase activity is very sensitive to any change at position 335 and, more generally, to any mutation of the residues highlighted by our analysis.

Conclusions

The presence of O-linked carbohydrates in the surface of many proteins has an important biological role and serve to modulate the physico-chemical properties of the glycosylated molecules. The enzymes responsible of forming this O-linkage are glycosyltransferases. In particular, mucin-type O-glycosylation is initiated by ppGalNAc-Ts. Here, we present a combined computational and experimental approach to investigate the catalytic activity of the human retaining glycosyltransferase ppGalNAc-T2. Hybrid QM/MM calculations on the full enzyme have been used for the first time to identify the key factors supporting catalysis in protein O-glycosylation by retaining glycosyltransferases and experiments have been carried out to validate the findings.

A front-side attack mechanism has been described, with a estimated potential energy barrier of ~20 kcal/mol (QM = M05-2X/TZVP), in reasonable agreement with the experimental kinetic data. The analysis of factors contributing to catalysis highlighted two key amino acids in the active site of the enzyme: Arg362 and Glu334. Tyr331 and the backbone of Ala307 were also found to stabilize the transition state, but to a lesser extent. Experimental and *in silico* mutation of residues in positions 334 and 362, and of Asn335 (which is situated on the β-face of the GalNAc ring), confirm that transferase activity is very sensitive to mutation at these positions.

Substrate-substrate interactions that contribute to catalysis by stabilizing the developing negative (positive) charge in UDP (α-GalNAc) have also been identified. Taken together, the interactions that predominate are those that stabilize the negative charge developing at the UDP as the transition state is approached, showing that this is a key aspect in retaining glycosyltransferases catalysis. Very interestingly, a new interaction that promotes catalysis has between characterised, that is, a hydrogen bond between the UDP and the amide group from the accepting Thr7. We propose this as a general feature in peptide O-glycosylation by retaining glycosyltransferases. Finally, an intra-substrate interaction involving the 2' NAc group of α-GalNAc has also been described to stabilize the transition

state. Complementary studies *in silico* with alternative donor substrates suggest that human ppGalNAc-T2 would be inactive if 2'-deoxy-Gal or 2'-oxymethyl-Gal were used, while UDP-Gal is confirmed as a valid substrate.

Acknowledgements

The authors want to acknowledge Dr. A. Milac for providing with the coordinates of the starting model for the calculations and also Dr. T. Fritz for helpful discussions. We thank the financial support from the Spanish "Ministerio de Economía y Competitividad" through project CTQ2011–24292 and the "Ramon y Cajal" program (L.M.), and from the "Generalitat de Catalunya", project 2009SGR409, and the intramural program of NIDCR, NIH.

Notes and references

[a] *Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, 08193 Bellaterra (Cerdanyola del Vallès), Barcelona (Spain). E-mail: Laura.Masgrau@uab.cat*
[b] *Departament de Química, Universitat Autònoma de Barcelona, 08193 Bellaterra (Cerdanyola del Vallès), Barcelona (Spain)*
[c] *Section on Biological Chemistry, National Institute of Dental and Craniofacial Research, National Institutes of Health, Department of Health and Human Services, Bethesda, MD 20892 (USA) E-mail: Lawrence.Tabak@nih.gov*

† Electronic Supplementary Information (ESI) available: Details on the experimental methods, as well as supplementary results are given in the ESI. See DOI: 10.1039/b000000x/

1 N. Jentoft, *Trends Biochem. Sci.*, 1990, **15**, 291.
2 A. M. Moody, D. Chui, P. A. Reche, J. J. Priatel, J. D. Marth and E. L. Reinherz, *Cell*, 2001, **107**, 501.
3 Z. Xu and A. Weiss, *Nat. Immunol.*, 2002, **3**, 764.
4 J. Sauer, B. W. Sigurskjold, U. Christensen, T. P. Frandsen, E. Mirgorodskaya, M. Harrison, P. Roepstorff and B. Svensson, *Biochim. Biophys. Acta* **2000**, *1543*, 275.
5 B. Garner, A. H. Merry, L. Royle, D. J. Harvey, P. M. Rudd and J. Thillet, *J. Biol. Chem.*, 2001, **276**, 22200.
6 L. V. Hooper and J. I. Gordon, *Glycobiology*, 2001, **11**, 1R.
7 J. C. Yeh, N. Hiraoka, B. Petryniak, J. Nakayama, L. G. Ellies, D. Rabuka, O. Hindsgaul, J. D. Marth, J. B. Lowe and M. Fukuda, *Cell*, 2001, **105**, 957.
8 W. S. Somers, J. Tang, G. D. Shaw and R. T. Camphausen, *Cell*, 2000, **103**, 467.
9 M. Alfalah, R. Jacob, U. Preuss, K. P. Zimmer, H. Naim and H. Y. Naim, *Curr. Biol.*, 1999, **9**, 593.
10 Y. Altschuler, C. L. Kinlough, P. A. Poland, J. B. Bruns, G. Apodaca, O. A. Weisz and R. P. Hughey, *Mol. Biol. Cell*, 2000, **11**, 819.
11 L. Breuza, M. Garcia, M.-H. H. Delgrossi and A. Le Bivic, *Exp. Cell Res.*, 2002, **273**, 178.
12 H. Y. Naim, G. Joberty, M. Alfalah and R. Jacob, *J. Biol. Chem.*, 1999, **25**, 17961.
13 X. Zheng and J. E. Sadler, *J. Biol. Chem.*, 2002, **277**, 6858.
14 G. W. Hart and R. J. Copeland, *Cell*, 2010, **143**, 672.
15 K. G. Ten Hagen, T. A. Fritz and L. A. Tabak, *Glycobiology*, 2003, **13**, 1R.
16 G. J. Strous, *Proc. Natl. Acad. Sci. U. S. A.*, 1979, **76**, 2694.
17 B. L. Cantarel, P. M. Coutinho, C. Rancurel, T. Bernard, V. Lombard and B. Henrissat, *Nucleic Acids Res.*, 2009, **37**, D233.
18 L. L. Lairson, B. Henrissat, G. J. Davies and S. G. Withers, *Annu. Rev. Biochem.*, 2008, **77**, 521.
19 A. Bottoni, G. P. Miscione and M. De Vivo, *Proteins: Struct., Funct., Bioinf.*, 2005, **59**, 118.
20 A. L. Bowman, I. M. Grant and A. J. Mulholland, *Chem. Commun.*, 2008, **37**, 4425.
21 G. J. Davies and S. G. Withers in *Comprehensive Biological Catalysis*, (Ed. M. L. Sinnott), London, 1998, pp. 119-208.
22 L. L. Lairson, C. P. Chiu, H. D. Ly, S. He, W. W. Wakarchuk, N. C. Strynadka and S. G. Withers, *J. Biol. Chem.*, 2004, **279**, 28339.
23 N. Soya, Y. Fang, M. M. Palcic and J. S. Klassen, *Glycobiology*, 2011, **21**, 547.
24 H. Gómez, J. M. Lluch and L. Masgrau, *J. Am. Chem. Soc.*, 2013, **135**, 7053.
25 V. Rojas-Cervellera, A. Ardèvol, M. Boero, A. Planas and C. Rovira, *Chemistry*, 2013, **19**, 14018.
26 I. Tvaroška, *Carbohydr. Res.*, 2004, **339**, 1007.
27 J. C. Errey, S. S. Lee, R. P. Gibson, C. Martinez-Fleites, C. S. Barry, P. M. Jung, A. C. O'Sullivan, B. G. Davis and G. J. Davies, *Angew. Chem.*, 2010, **49**, 1234.
28 S. S. Lee, S. Y. Hong, J. C. Errey, A. Izumi, G. J. Davies and B. G. Davis, *Nat. Chem. Biol.*, 2011, **7**, 631.
29 A. Ardèvol and C. Rovira, *Angew. Chem.*, 2011, **50**, 10897.
30 H. Gómez, I. Polyak, W. Thiel, J. M. Lluch and L. Masgrau, *J. Am. Chem. Soc.*, 2012, **134**, 4743.
31 H. Gómez, J. M. Lluch and L. Masgrau, *Carbohydr. Res.*, 2012, **356**, 204.
32 B. Schuman, S. V. Evans and T. M. Fyles, *PLOS ONE*, 2013, **8**, e71077.
33 A. L. Milac, N. V. Buchete, T. A. Fritz, G. Hummer and L. A. Tabak, *J. Mol. Biol.*, 2007, **373**, 439.
34 T. A. Fritz, J. Raman and L. A. Tabak, *J. Biol. Chem.*, 2006, **281**, 8613.
35 T. Kubota, T. Shiba, S. Sugioka, S. Furukawa, H. Sawaki, R. Kato, S. Wakatsuki and H. Narimatsu, *J. Mol. Biol.*, 2006, **359**, 708.
36 J. Kón□a and I. Tvaros□ka, *Chem. Pap.*, 2009, **63**, 598.
37 A. Elhammer and S. Kornfeld, *J. Biol. Chem.*, 1986, **261**, 5249.
38 P. Sherwood, A. H. de Vries, S. J. Collins, S. P. Greatbanks, N. A. Burton, M. A. Vincent and I. H. Hillier, *Faraday Discuss.*, 1997, **106**, 79.
39 A. H. de Vries, P. Sherwood, S. J. Collins, A. M. Rigby, M. Rigutto and G. J. Kramer, *J. Phys. Chem. B*, 1999, **103**, 6133.
40 D. Bakowies and W. Thiel, *J. Phys. Chem.*, 1996, **100**, 10580.
41 M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai and G. Seifert, *Phys. Rev. B*, 1998, **58**, 7260.
42 T. Frauenheim, G. Seifert, M. Elstner, T. Niehaus, C. Köhler, M. Amkreutz, M. Sternberg, Z. Hajnal, A. D. Carlo and S. Suhai, *J. Phys.: Condens. Matter*, 2002, **14**, 3015.
43 A. D. MacKerell, D. Bashford, Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz- Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586.
44 A. D. MacKerell, M. Feig and C. L. Brooks, *J. Am. Chem. Soc.*, 2004, **126**, 698.
45 P. Sherwood, A. H. de Vries, M. F. Guest, G. Schreckenbach, C. R. A. Catlow, S. A. French, A. A. Sokol, S. T. Bromley, W. Thiel, A. J. Turner, S. Billeter, F. Terstegen, S. Thiel, J. Kendrick, S. C. Rogers, J. Casci, M. Watson, F. King, E. Karlsen, M. Sjovoll, A. Fahmi, A. Schäfer and C. Lennartz, *J. Mol. Struct. (Theochem.)*, 2003, **632**, 1.
46 J. P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comp. Phys.*, 1977, **23**, 327.
47 J. C. Slater, *Phys. Rev.*, 1951, **81**, 385.
48 A. D. Becke, *Phys. Rev. A*, 1988, **38**, 3098.
49 S. H. Vosko, L. Wilk and M. Nusair, *Can. J. Phys.*, 1980, **58**, 1200.
50 J. P. Perdew, *Phys. Rev. B*, 1986, **33**, 8822.
51 A. Scha□fer, H. Horn and R. Ahlrichs, *J. Chem. Phys.*, 1992, **97**, 2571.
52 Y. Zhao, N. E. Schultz and D. G. Truhlar, *J. Chem. Theory Comput.*, 2006, **2**, 364.
53 A. Schäfer, C. Huber and R. Ahlrichs, *J. Chem. Phys.*, 1994, **100**, 5829.
54 I. C. Gábor, D. F. Alfred, P. J. Glenn and A. S. Carlos, *J. Chem. Theory Comput.*, 2009, **5**, 679.
55 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 1372.

56 P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623.
57 C. T. Lee, W. T. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter*, 1988, **37**, 785.
58 J. P. Foster and F. Weinhold, *J. Am. Chem. Soc.*, 1980, **102**, 7211.
59 A. E. Reed and F. Weinhold, *J. Chem. Phys.*, 1983, **78**, 4066.
60 A. E. Reed, R. B. Weinstock and F. Weinhold, *J. Chem. Phys.*, 1985, **83**, 735.
61 A. E. Reed, L. A. Curtiss and F. Weinhold, *Chem. Rev.*, 1988, **88**, 899.
62 E. D. Glendening, A. E. Reed, J. E. Carpenter and F. Weinhold *Program NBO Version 3.1*.
63 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. Montgomery, J. A., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian09, revision A.1; Gaussian, Inc., Wallingford CT*, 2009.
64 R. Ahlrichs, M. Bär, M. Häser, H. Horn and C. Kölmel, *Chem. Phys. Lett.*, 1989, **162**, 165.
65 W. Thiel *Program MNDO2009, Version 7.0*, Max-Planck-Institut für Kohlenforschung: Mu□lheim, 2005.
66 W. Smith and T. R. Forester, *J. Mol. Graph.*, 1996, **14**, 136.
67 J. Nocedal, *Math. Comp.*, 1980, **35**, 773.
68 D. C. Liu and J. Nocedal, *Math. Programming*, 1989, **45**, 503.
69 A. Banerjee, N. Adams, J. Simons and R. Shepard, *J. Phys. Chem.*, 1985, **89**, 52.
70 J. Baker, *J. Comput. Chem.*, 1986, **7**, 385.
71 S. R. Billeter, A. J. Turner and W. Thiel, *Phys. Chem. Chem. Phys.*, 2000, **2**, 2177.
72 D. J. Vocadlo and S. G. Whiters, *Biochemistry*, 2005, **44**, 12809.
73 I. Tvaros□ka, S. Kozmon, M. Wimmerová and J. Koča, *J. Am. Chem. Soc.*, 2012, **134**, 15563.
74 H. H. Wandall, H. Hassan, E. Mirgorodskaya, A. K. Kristensen, P. Roepstorff, E. P. Bennett, P. A. Nielsen, M. A. Hollingsworth, J. Burchell, J. Taylor-Papadimitriou and H. Clausen, *J. Biol. Chem.*, 1997, **272**, 23503.
75 H. Jamaluddin, P. Tumbale, S. G. Withers, K. R. Acharya and K. Brew, *J. Mol. Biol.*, 2007, **369**, 1270.
76 S. Wragg, F. K. Hagen and L. A. Tabak, *J. Biol. Chem.*, 1995, **270**, 16947.
77 F. K. Hagen, B. Hazes, R. Raffo, D. deSa and L. A. Tabak, *J. Biol. Chem.*, 1999, **274**, 6797.