CHEMISTRY
A EUROPEAN JOURNAL

FULL PAPER

DOI: 10.1002/chem.201300383

A QM/MM Investigation of the Catalytic Mechanism of Metal-Ion-Independent Core 2 β1,6-*N*-Acetylglucosaminyltransferase**

Igor Tvaroška,*[a, b] Stanislav Kozmon,[b, c, d] Michaela Wimmerová,[b, c] and Jaroslav Koča[b, c]

Abstract: β1,6-GlcNAc-transferase (C2GnT) is an important controlling factor of biological functions for many glycoproteins and its activity has been found to be altered in breast, colon, and lung cancer cells, in leukemia cells, in the lymhomonocytes of multiple sclerosis patients, leukocytes from diabetes patients, and in conditions causing an immune deficiency. The result of the action of C2GnT is the core 2 structure that is essential for the further elongation of the carbohydrate chains of *O*-glycans. The catalytic mechanism of this metal-ion-independent glycosyltransferase is of paramount importance and is investigated here by using quantum mechanical (QM) (density functional theory (DFT))/molecular modeling (MM) methods with different levels of theory. The structural model of the reaction site used in this report is based on the crystal structures of C2GnT. The entire enzyme–substrate system was subdivided into two different subsystems: the QM subsystem containing 206 atoms and the MM region containing 5914 atoms. Three predefined reaction coordinates were employed to investigate the catalytic mechanism. The calculated potential energy surfaces discovered the existence of a concerted S_N2-like mechanism. In this mechanism, a nucleophilic attack by O6 facilitated by proton transfer to the catalytic base and the separation of the leaving group all occur almost simultaneously. The transition state for the proposed reaction mechanism at the M06-2X/6-31G** (with diffuse functions on the O1′, O5′, O_{Glu}, and O6 atoms) level was located at C1−O6 = 1.74 Å and C1−O1 = 2.86 Å. The activation energy for this mechanism was estimated to be between 20 and 29 kcal mol^{-1}, depending on the method used. These calculations also identified a low-barrier hydrogen bond between the nucleophile O6H and the catalytic base Glu320, and a hydrogen bond between the *N*-acetamino group and the glycosidic oxygen of the donor in the TS. It is proposed that these interactions contribute to a stabilization of TS and participate in the catalytic mechanism.

Keywords: computer chemistry · density functional calculations · molecular modeling · reaction mechanisms · transferases · transition state

Introduction

Protein O-glycosylation is a post-translational modification implicated in a wide range of physiological processes including cell adhesion and trafficking, T-cell apoptosis, cell signaling, endocytosis, and pathogen-host interaction.[1] Mucin-type *O*-glycans are expressed on the cell surface of leukocytes and have been shown to play a vital role during the early course of the inflammatory cascade.[2] Glycans with a β1,6-GlcNAc branch can be extended and capped with a Lewis^x antigen.[3] This represents the minimum carbohydrate epitope recognized by P-, E-, and L-selectins, which mediate the tethering and rolling of leukocytes on activated endothelial cells.[4]

The key enzyme in the biosynthesis of branched *O*-glycans is the Golgi enzyme UDP-GlcNAc: Galβ1-3GalNAc-(GlcNAc to GalNAc) β1,6-GlcNAc-transferase (Core 2 GnT, C2GnT). C2GnT transfers GlcNAc to the core 1 structure on Galβ1-3GalNAcβ-O-Ser/Thr glycoproteins and forms the core 2 structure (Scheme 1), GlcNAcβ1-6(Galβ1-3)GalNAcβ-O-Ser/Thr.[5] The formation of this branched structure is essential for the further elongation of carbohydrate chains of *O*-glycans. The C2GnT enzyme is a key controlling factor in the biological functions of various glycoproteins and its activity has been found to be altered in breast, colon, and lung cancer cells,[6] in leukemia cells,[7] in the lymphomonocytes of multiple sclerosis patients,[8] leuko-

[a] Dr. I. Tvaroška
Institute of Chemistry, Slovak Academy of Sciences
845 38 Bratislava (Slovak Republic)
Fax: (+421) 2-5941-0222
E-mail: chemitsa@savba.sk

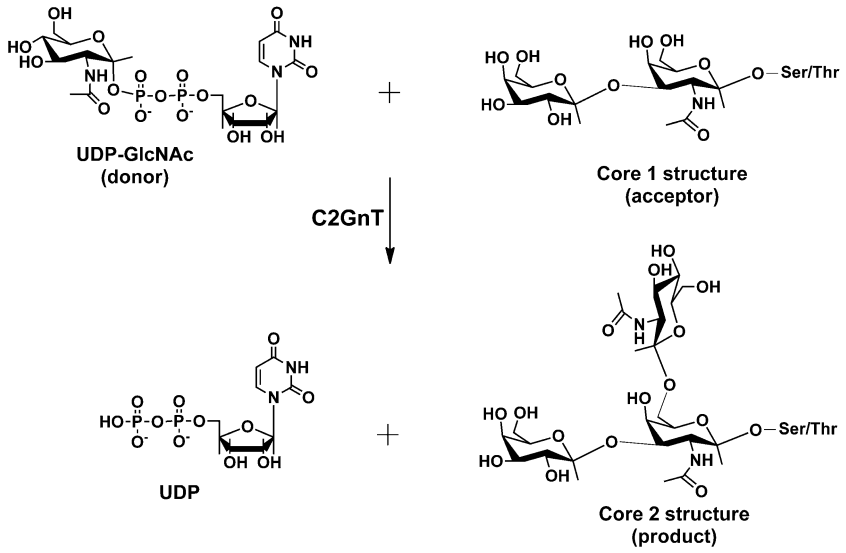
[b] Dr. I. Tvaroška, Dr. S. Kozmon, Prof. M. Wimmerová, Prof. J. Koča
Central European Institute of Technology (CEITEC)
Masaryk University, 625 00 Brno (Czech Republic)

[c] Dr. S. Kozmon, Prof. M. Wimmerová, Prof. J. Koča
Faculty of Science, National Centre for, Biomolecular Research
Masaryk University, 625 00 Brno (Czech Republic)

[d] Dr. S. Kozmon
On leave from the Institute of Chemistry
Slovak Academy of Sciences, 845 38 Bratislava (Slovak Republic)

[**] QM/MM = quantum mechanical/molecular modeling

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/chem.201300383.



Scheme 1. Schematic diagram of the enzymatic reaction catalyzed by C2GnT.

cytes from diabetes patients,[9] and in conditions causing an immune deficiency.[10] Three isoforms of C2GnT have been identified and cloned. Two of them, the leukocyte-type (C2GnT-1 or C2GnT)[11] and thymus-associated enzyme (C2GnT-3 or C2GnT-T)[12] exhibit exclusive core 2 acceptor specificity, whereas C2GnT-2 (or C2GnT- M),[13] which is distributed in mucin-secreting tissues along the gastrointestinal tract, can synthesize core 2, core 4, and I-branched structures. Though the C2GnT enzymes differ in their tissue distribution, they exhibit a significant homology.[14] The members of the C2GnT group share nine conserved cysteine residues. It was found that eight cysteines are involved in disulfide bonds, Cys59–Cys413, Cys100–Cys172, Cys151–Cys199, and Cys372–Cys381[15] and one, Cys235 was a free thiol or involved in intermolecular (dimer) formation. The presence of seven conserved cysteines (C59, C151, C199, C172, C372, C381, and C413) is required for full activity of the enzyme, whereas the C100S mutant had 10% activity. The monomer is assumed to be the active form of the enzyme.[16]

In terms of primary sequences, C2GnT belongs to the GT14 family in the CAZY database[17] and is a metal-ion-independent enzyme. C2GnT operates through an inverting mechanism. In general, this mechanism follows a S_N2-like reaction in which the enzyme provides a catalytic base that activates the nucleophile (in C2GnT, this is the C6 hydroxyl group from a GalNAc residue) to displace the uridine 5′-diphosphate (UDP) leaving group from the UDP-GlcNAc donor in a concerted process.[18] It is assumed that C2GnT follows the ordered bi-bi catalytic mechanism with the UDP-GlcNAc bindings first, followed by the disaccharide acceptor binding over the nucleotide sugar.

The first three-dimensional model of C2GnT (amino acid 117–428) was proposed based on the distribution of the disulfide bonds and threading/homology modeling by using the crystal structure of T4 bacteriophage GT.[15] The X-ray crystal structure of murine C2GnT (residues 38–428) in the absence (PDB code 2GAK) and presence of the acceptor substrate Galβ1-3GalNAc (PDB code 2GAM) were resolved[19] at resolution of 2.0 and 2.7 Å, respectively. Recently, the C217S mutant was prepared and its X-ray structure in a complex with UDP was solved.[20] It was found that the monomer units of the observed dimer C2GnT possesses the GT-A fold and are connected by a disulfide bond between the Cys235 residues. However, the enzyme does not contain the metal-ion-binding DXD pattern typical for inverting glycosyltransferases with the GT-A fold. The structure of C2GnT contains two regions; the first (38–121) is composed of α-helices. The second region, which corresponds to the catalytic domain (122-428), is an α/β/α structure consisting of a central six-stranded mixed β-sheet. Four disulfide bonds were found in each monomer (Cys151–Cys199, Cys372–Cys381, Cys59–Cys413, and Cys100–Cys172); the remaining Cys217 was unpaired and located in the donor binding site. It is noteworthy that C2GnT may occur in an "open" conformation, and a "closed" conformation though these two conformations do not resemble the loop ordering observed in metal-ion-dependent inverting glycosyltransferases.[21] The location of the C2GnT-conserved Glu320 residue structurally corresponds to the catalytic base found in other glycosyltransferases with the GT-A fold.

The recently resolved crystal structures of C2GnT shed some light on the catalytic mechanism of C2GnT. However the catalytic mechanism of C2GnT at the microscopic level is not yet understood. The majority of inverting glycosyltransferases from the GT-A family employ divalent cation coordinated by the DXD motif to facilitate the departure of the negatively charged leaving group.[18a] The catalytic mechanism of the inverting glycosyltransferases of the GT-A fold was investigated by using a cluster model of the active site[22] and by using hybrid QM/MM methods.[23] These calculations supported a direct displacement S_N2-like catalytic mechanism involving a near simultaneous nucleophilic attack facilitated by proton transfer to the catalytic base and leaving-group dissociation. However, the C2GnT enzyme does not contain the metal-ion-binding DXD pattern typical for inverting glycosyltransferases with the GT-A fold. Recently, the catalytic mechanism of metal-ion-independent glycosyltransferase *O*-GlcNAc transferase (OGT) was investigated by hybrid QM/MM method,[24] however, OGT possess the GT-B fold. To our knowledge, a theoretical investigation of the mechanism of metal-ion-independent glycosyltransferases with the GT-A fold has not been accomplished.

This prompted us to examine the catalytic mechanism of the metal-ion-independent C2GnT enzyme at the atomic level. The determination of the transition-state structure is a valuable outcome of this study, which could serve as a guide in designing potent and specific inhibitors of C2GnT with therapeutic potential for the treatment of overzealous inflammatory responses that lead to a pathological condition such as rheumatoid arthritis, asthma, and inflammatory bowel disease.

Models and Methods

Model preparation: For modeling the catalytic reaction, we followed the approach used in our recent study.[24] The coordinates of the C2GnT(residues 56–428)-Galβ1-3GalNAc complex in a "closed" conformation, which represents the active substrate complex,[20] were obtained from the PDB database under the code 2GAM, and prepared with Modeller,[25] UCSF Chimera,[26] Schrodinger's Maestro,[27] and Protein Preparation Wizard[28] from the Schrodinger suite of programs as follows. All water oxygen atoms were removed from the structure before the missing residues were added. The four missing amino acids residues Glu81–Lys84 were added by using Modeller so that all atoms from the original PDB data were fixed to their original coordinates and only the added amino acid residue in the loops were optimized. Hydrogen atoms were then added, and protonation states were assigned based on the residue pK_a value at normal pH (7.0). The orientations of the added hydrogen atoms and protonation of amino acid residues were based on the positions and types of the neighboring atoms by using the method implemented in UCSF Chimera. The protonated structure was used for docking the missing active donor substrate UDP-GlcNAc. A docking grid was generated with Maestro GUI[27] and Glide.[29] Atom types and partial charges were assigned according to the OPLS_2005 force field also known as OPLS-AA.[30] The structure of UDP-GlcNAc was obtained by energy minimization by using the program Jaguar v9.0[31] at the DFT B3LYP[32] level with the 6-31+G* basis set prior to docking. The calculated ESP charges were used as input partial charges for ligand atoms in the docking calculations. A docking grid was generated for the C2GnT-L-Galβ1-3GalNAc structure with the center of the cubic grid box placed on the centroid of the bound Galβ1-3GalNAc. The box size was set to 12 Å in all three dimensions. The UDP-GlcNAc molecule was the docked into generated grid, in which positions of the amino acid side chains were held fixed and UDP-GlcNAc was allowed to rotate around the rotatable single bonds (flexible ligand/rigid receptor docking procedure). During the docking, the sampling of the GlcNAc ring conformation was not allowed and was docked in the $^4C_1$ ring conformation. Up to five thousand docking poses were evaluated. The pose was considered as different if the root-mean-square (rms) would be bigger than 0.5 Å, and the maximum atomic displacement was greater than 1.3 Å. Each observed pose was minimized in site and the docking score was calculated. The other Glide settings were not changed and their default parameters for a standard precision docking were used. The best 25 poses with the highest score were kept and analyzed by hand. A docked pose was chosen for UDP-GlcNAc at a reasonable distance from the GlcNAc anomeric carbon from the acceptor and O6 hydroxyl from the Galβ1-3GalNAc disaccharide. The C2GnT-UDP-GlcNAc-Galβ1-3GalNAc ternary complex was overlaid with the original crystal structure. In the X-ray structure, all oxygen atoms with low B-factors were visualized in a 10 Å area around the UDP-GlcNAc. A comparison of the docked UDP-GlcNAc with the X-ray structure of the C217S mutant[20] form of C2GnT revealed that the position of the UDP residue in both structures is remarkably similar. All oxygen atoms in place of docked UDP-GlcNAc were deleted, and the possible interactions of the remaining crystallographic water oxygen atoms were analyzed by hand. The eight water oxygen atoms with potential non-covalent interactions that could be made with both ligands or between the C2GnT and the ligands (water residues 2, 8, 9, 21, 38, 45, 49, and 74) were retained in the active site. Water oxygen atoms were protonated, and then water molecules were rotated to make appropriate interactions with the neighboring atoms.

QM/MM model: The QM/MM calculations were carried out using Schrodinger's QSite program.[33] In Qsite, the QM/MM methodology (an additive scheme) was used with frozen atoms and an electrostatic treatment at the interface between QM and MM regions by using the Gaussian charge distribution represented on a grid. The entire enzyme–substrate system (C2GnT-UDP-GlcNAc-Galβ1-3GalNAc ternary complex), consisting of 379 amino acids, eight water molecules, donor, and acceptor (altogether 6120 atoms), was partitioned into two separate subsystems: the QM and the MM regions (see Scheme 2). The QM region, containing 206 atoms, was formed of the complete donor UDP-GlcNAc and complete Galβ1-3GalNAc acceptor, and the side chains of amino acids critical for the catalytic activity; Cys217, Glu243, Lys251, Arg254, Glu320 (suggested catalytic base), Arg378, and Lys401. The MM region (5914 atoms) was composed of the remaining C2GnT atoms and eight molecules of water in the model complex structure. Prior to potential energy surface calculations, a geometry optimization of the whole system was performed to obtain a refined location of the donor. The structure of the C2GnT-UDP-GlcNAc-Galβ1-3GalNAc complex, generated by the above procedure, was optimized, at the beginning using dynamic constraints of 3.0 and 2.0 Å on the O6(nucleophile)···C1 and H6···O_{Glu} (catalytic base Glu320) distances, respectively, and then fully relaxed. In this optimization, the QM subsystem was treated using density functional theory (DFT) with the B3LYP[32] functional and 6-31G** basis set with diffuse functions (6-31+G**) on the O1′, O5′, O_{Glu}, and O6 atoms from the Jaguar library,[31] and the MM subsystem was characterized with an OPLS_2005 all-atom force field.[30] The calculated structure was then used as the starting structure (Michaelis complex, ES) for modeling the reaction catalyzed by C2GnT, and is shown in Figure 1.

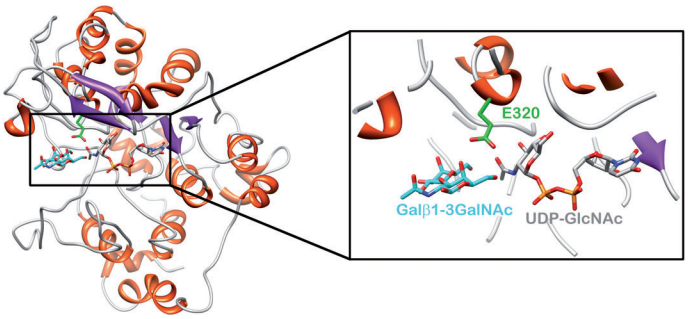
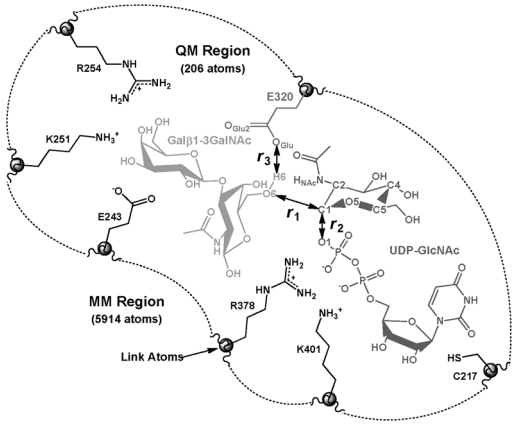


Figure 1. Ribbon representation of the overall structure of the modeled C2GnT-UDP-GlcNAc-Galβ1-3GalNAc complex calculated by using the QM/MM/M06-2X method. Water molecules and hydrogen atoms are not shown for clarity. Residues included in the QM region for the QM/MM calculations are shown in stick representation. Left: overall view of the C2GnT complex; right: close-up of the active site with the acceptor shown in cyan and the catalytic base in green.

Reaction mechanism: The catalytic site chemistry of C2GnT involves the creation of a new glycosidic linkage between the GalNAc oxygen of the disaccharide acceptor (O6) and the anomeric carbon C1 of the donor GlcNAc, cleavage of the donor glycosidic linkage between GlcNAc and UDP, and a transfer of the H6 proton from the acceptor hydroxyl O6H to the catalytic base (Glu320). The reaction mechanism was monitored, similarly as in our recent publication,[24] by using three reaction coordinates. The first reaction coordinate r_1 was defined as the distance between the anomeric carbon C1 and the nucleophilic oxygen O6 of the acceptor hydroxyl group. This reaction coordinate represented the nucleophilic attack of the GlcNAc acceptor on the anomeric carbon of the donor UDP-GlcNAc and designates the formation a new β-glycosidic linkage. The second reaction coordinate r_2 represented the dissociation of the glycosidic bond C1−O1, and was defined as the distance between the

anomeric carbon C1 and glycosidic oxygen O1 of the donor UDP-GlcNAc. The third reaction coordinate r_3 describes the transfer of the proton from the nucleophile hydroxyl to the oxygen of the catalytic base (Glu320) and is defined as the distance between the proton of the GalNAc hydroxyl group H6 and the oxygen of Glu320 O_{Glu}. All three coordinates are depicted in Scheme 2, which only shows the QM region.



Scheme 2. Schematic representation of the active site model dividing into QM and MM regions for the hybrid QM/MM calculations. The QM region contains complete UDP-GlcNAc (donor), whole residues of the GlcNAc-Galβ1-3GalNAc disaccharide (acceptor), and side chains of amino acids essential for catalytic activity: Cys217, Glu243, Lys251, Arg254, Glu320 (catalytic base, E320), Arg378, and Lys401. The reaction coordinates r_1, r_2, and r_3 are indicated with arrows and linked atoms with a sphere.

The reaction potential energy surface (PES) was determined by so-called adiabatic mapping. The reaction coordinates were varied by increments of 0.2 Å, between 3.5 and 1.4 Å for r_1 and r_2, and between 2.0 and 1.0 Å for r_3, respectively. In the vicinity of the barriers, the increment was decreased to 0.1 and then 0.05 Å. Geometrical variables of all 6120 atoms were optimized except for the reaction coordinates. The two-dimensional PES was calculated by using density functional theory (DFT) with the B3LYP[32] functional and 6-31G** basis set, with diffuse functions (6-31+G**) on O1′, O5′, O6, and O_{Glu} from the Jaguar library[31] for the QM region and the OPLS 2005 all-atom force field[30] for the MM region. In these calculations, to avoid extraordinarily extensive calculations, the geometry optimizations were stopped after 150 optimization steps, providing that the gradient rms was lower than 0.0005 Hartree $Bohr^{-1}$. The CPU time required for one geometry optimization step on a cluster with 24 processors (Intel i7-3300 GHz) was between 4–5 h depending on the used functional. Thus, the calculation of one point on the PES that required 150 optimization steps lasted 2–3 weeks. Once the energy profile was obtained, the structure of the energy maximum was used to search for the transition state (TS) by using the default criteria for this computation in QSite.[33] Although the B3LYP functional has become extremely popular in DFT calculations, there are several reports that examine the reliability of hybrid GGA functional B3LYP.[34] Recently, based on a comparison of relatively large TS structures and energies, a procedure for large systems was recommended[34b] in which hybrid GGA functionals such as B3LYP are used in geometry optimization and then additional single-point calculations are performed by using more expensive hybrid meta GGA functional with ultrafine grid. We have used more prudent approach. In all calculations we used ultrafine grid, with the QSITE parameters gdftfine = −14, gdftgrad = −14, and gdftmed = −14. For the S_N2-like reaction pathways, the structures of the Michaelis complex (ES), transitions state (TS), and product complex (P) were optimized using the B3LYP, M06-2X, [35] MPW1K,[36] PWB6K,[37] and M05-2X[38] functionals with the basis set defined as above. In addition, single-point calculations were carried out for these stationary points by using various DFT functionals: B3LYP, M06-2X, MPW1K, PWB6K, and M05-2X.

Results and Discussion

To examine the active site chemical reaction, we calculated the potential energy surface (PES) as a function of the three reaction coordinates. These represent the creation of a new β-glycosidic linkage (r_1) between the acceptor and donor, cleavage of the donor C1−O1 glycosidic bond (r_2), and proton transfer (r_3) from the acceptor nucleophile to the catalytic base. The possible reaction pathways were explored by means of two two-dimensional maps, namely (r_1, r_2) and (r_1, r_3), and the number of optimized structures exceeded 400. All geometrical parameters, apart from the constrained reaction coordinates were optimized during these calculations, for example, for each point on the (r_1, r_2) map, only r_1 and r_2 were constrained to the value representing a given point on the calculated PES. The calculated potential energy surfaces of the catalytic reactions are represented in the form of contour diagrams in Figures 2 and 3. Figure 4 shows the M06-2X optimized structures of the active site models in the ES, TS, and P complexes. Selected geometrical parameters and ESP charges for ES, TS, and P are listed in Table 1 (see also Table 1 S in the Supporting Information), and the relative energies of stationary structures calculated at various levels of the theory are given in Table 2.

Structure of the C2GnT-UDP-GlcNAc-Galβ1–3GalNAc ternary complex in ES: A comparison of the X-ray[19,20] and the ES complex structures shows that C2GnT does not undergo significant structural changes upon optimization by the QM/MM methods, and all relevant interactions observed in the X-ray structures can be seen in the calculated ternary complex. Furthermore, the calculated structure provides insight into a binding mode of the GlcNAc residue, which is not available in X-ray structures, and exposes interactions involved in its binding. Thus in ES, the assumed catalytic base Glu320 is involved in a bidentate hydrogen bond with both the O4 and O6 of the GalNAc moiety of the acceptor with the H6−O_{Glu} and O4−O_{Glu2} distances (at the M06-2X level) being 1.961 and 1.72 Å, respectively. The interaction with O6 is important for the acceptor binding, the removal of O6 led to a significant decrease in acceptor binding affinity[39] (K_M = 0.08 vs. K_M = 0.56 mM). The O4 hydroxyl of GalNAc also interacts through a hydrogen bond with Arg254; the O4···HN_{Arg254} distance being 1.68 Å. The distances of 1.66 and 1.63 Å between the O4 and O6 protons of Gal with Glu243 also suggests hydrogen-bond interactions. These interactions also seem to be relevant to the acceptor binding, since the 6-deoxy acceptor analogue is a poor substrate.[5b] The hydrogen-bond interactions of Tyr358 with the NH group of GalNAc and O2 of Gal, with proton–oxygen distances of approximately 2.2 Å, further stabilize the complex. The UDP-GlcNAc donor is buried in the binding site, and its location is stabilized by a combination of hydrogen

Table 1. QM/MM calculated geometrical parameters (distances in Å, angles in degrees) and selected ESP charges (Q, in e) of the Michaelis complex (ES), transition state (TS), and product complex (P) observed on the reaction pathway described in Figure 3 at B3LYP, MPW1K, and M06-2X levels and 6-31G** basis set with diffuse functions on O1′, O5′, O6, and O_{Glu320} (6-31+G**) atoms.

	ES			TS			P		
	B3LYP	MPW1K	M06-2X	B3LYP	MPW1K	M06-2X	B3LYP	MPW1K	M06-2X
r_1(C1−O6)	3.044	3.204	3.030	1.813	1.778	1.735	1.442	1.421	1.428
r_2(C1−O1)	1.424	1.404	1.415	2.868	2.887	2.866	3.101	3.055	3.008
r_3(O_{Glu}−H6)	1.984	1.938	1.961	1.442	1.417	1.382	1.008	0.996	1.007
O6−O_{Glu}	2.903	2.884	2.907	2.451	2.424	2.411	2.614	2.575	2.602
C1−O5	1.389	1.380	1.391	1.315	1.302	1.322	1.407	1.387	1.399
H_{Nac}−O1	2.489	2.479	2.484	1.795	1.840	1.852	1.732	1.718	1.820
HO4−O_{Glu2}	1.720	1.745	1.727	1.702	1.675	1.647	1.861	1.831	1.851
O5-C1-O6	94.7	94.4	94.6	106.8	106.4	106.9	105.6	105.8	105.8
O5-C1-O1	112.1	112.1	1117	100.9	98.8	98.7	113.6	113.8	110.1
O5-C1-H1	106.4	106.7	106.9	111.9	112.4	112.3	110.7	111.0	110.6
O5-C1-C2	109.9	109.4	109.6	121.4	120.9	120.4	113.8	113.3	113.9
O6-C1-O1	146.7	147.1	147.8	147.6	151.1	150.8	135.3	135.6	139.0
C4-C5-O5-C1	−3.6	−4.6	54.7	26.1	29.9	29.4	65.6	66.5	66.3
C5-O5-C1-H1	176.7	176.5	175.6	135.7	132.8	130.7	77.8	75.9	76.6
C5-O5-C1-O1	55.8	55.3	54.3	72.7	68.3	67.9	37.3	34.5	34.6
C5-O5-C1-O6	−144.4	−144.7	−144.7	−124.3	−126.0	−126.5	−164.7	−166.1	−167.7
H2-C2-N-HN	−152.4	−152.7	−153.9	−163.8	−165.5	−166.1	−163.8	−167.8	−164.5
Q(C1)	1.195	1.199	1.201	0.7382	0.6928	0.7124	0.4145	0.2623	0.1923
Q(O1)	−0.570	−0.581	−0.590	−0.8835	−0.9175	−0.9124	−0.8497	−0.8697	−0.8831
Q(O6)	−0.714	−0.711	−0.711	−0.4704	−0.4413	−0.4208	−0.2007	−0.1201	−0.1646
Q(O5)	−0.927	−0.956	−0.950	−0.7013	−0.7123	−0.7338	−0.6043	−0.5958	−0.5909
Q(C2)	−1.115	−1.171	−1.108	−0.2918	−0.3041	−0.3075	0.0126	−0.0994	−0.0002
Q(C5)	1.007	1.085	1.087	0.7421	0.7916	0.7958	0.4042	0.4588	0.4280
Q(O_{Glu})	−0.725	−0.762	−0.742	−1.0010	−1.0263	−1.0361	−0.8241	−0.8286	−0.8657
Q(H1)	0.027	0.039	0.040	0.1244	0.1684	0.1450	0.1254	0.1915	0.2488
Q(HO6)	0.613	0.630	0.623	0.6993	0.7242	0.7242	0.5464	0.5370	0.6241
Q(HN_{Ac})	0.367	0.391	0.394	0.3671	0.4076	0.3992	0.2919	0.31332	0.3105

bonds and van der Waals interactions. The uracil moiety of UDP-GlcNAc is located between the imidazol ring of His130 and Val185 at an approximate distance of 4 Å. Also, the distance of 1.76 Å between the HN3 proton of the uracil moiety and Asp155 suggests a stabilizing hydrogen-bond interaction. The ribose position seems to be stabilized by interactions of the O2 and O3 oxygen atoms with protons from Cys217; with approximate distances of 2.21 Å. The ribose HO3 proton is involved in the interaction with Arg192, with a distance of 2.2 Å. The diphosphate group in GT-A glycosyltransferases is usually chelated by a bound divalent metal. However, this is not the case with C2GnT, and the metal-ion-coordinating DXD motif is not present in the active site with C2GnT. Instead, the negative charges of the diphosphate moiety are stabilized by interactions with positively charged Lys401 and Arg378; with phosphate oxygen-lysine proton distances between 1.57 Å and 1.90 Å. The interactions with Lys401 may play a significant role in the binding of UDP-GlcNAc, as suggested by a loss of binding of the K401A mutant to a UDP-hexanolamine column.[20] The orientation of the GlcNAc residue of the donor is stabilized by several hydrogen-bond interactions. The oxygen of Asp319 forms 1.71 and 1.70 Å hydrogen-bond interactions with the O4′ and O6′ protons of GlcNAc and the O3′ oxygen is the acceptor of a proton from the backbone nitrogen of Ala287 with a distance of 2.15 Å. An *N*-acetyl group is located in the apolar pocket created by Val354 and Val380 that also contributes to the orientation of GlcNAc.

PES as a function of distances r_1 (r_{C1-O6}) and r_3 ($r_{H6-OGlu}$): The potential energy surface in Figure 2 characterizes the nucleophilic attack (r_1) of the O6 oxygen from the acceptor on the anomeric carbon C1 of the donor UDP-GlcNAc and the proton transfer (r_3) from the O6 hydroxyl (acceptor) to the catalytic base (Glu320). The vertical axis describes the proton transfer and the horizontal axis designates nucleophilic attack.

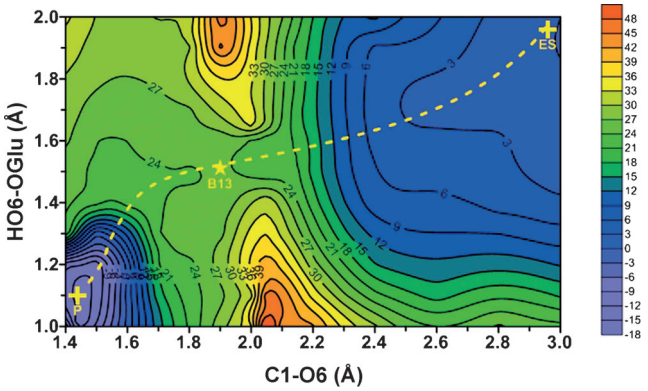


Figure 2. The B3LYP calculated PES with the 6-31G** basis set and diffuse functions (6-31+G**) on the O1′, O5′, O_{Glu}, and O6 atoms by using the distances r_1 and r_3 (in Å) as reaction coordinates. The yellow dashed line from the upper right corner to the bottom left corner indicates the S_N2-like reaction pathway from the ES to the P through B13. Contours are in kcal mol^{-1}.

The PES given in Figure 2 shows the existence of one reaction pathway with a single transition barrier (B13) in the central region of the map that connects the upper right corner with the lower-left corner. The Michaelis complex is located in the upper-right corner of the PES. A thorough examination of the geometrical parameters of the lowest minimum in the lower-left corner on the PES clearly found, as was observed in our previous studies,[22,24,40] that the structure of this minimum corresponds to the product (P). In other words, the C1−O1 length (r_2) varies in a continuous manner with the C1−O6 distance (r_1) and the proton transfer (r_3). In this pathway, from the Michaelis complex (ES) to the products (P), the presence of only one transition barrier (B13) indicates the existence of a concerted S_N2-like mechanism, in which the nucleophilic attack by O6 (r_1) facilitated by the proton transfer to the catalytic base (r_2) and the separation of the leaving group (r_3) all occur almost simultaneously. The key factor in this S_N2-like mechanism appears to be the nucleophilic attack, with the nucleophilicity of GalNAc hydroxyl O6 increased by the proton transfer to Glu320, which functions as the catalytic base. The pathway follows the diagonal of the two-dimensional map and a transition barrier (B13) of ≈23 kcal mol^{-1} located at $r_1 \approx 1.9$ and $r_3 \approx 1.5$ Å was detected in the central area of the contour map. An analysis of the geometrical parameters showed that the C1−O1 bond was lengthened to $r_2 \approx 2.9$ Å.

PES as a function of distances r_1 (r_{C1-O6}) and r_2 (r_{C1-O1}): Figure 3 shows the PES calculated in terms of the distance (r_1) between the C1 carbon and the O6 oxygen of the attacking GalNAc residue and the distance (r_2) between the anomeric carbon C1 and the glycosidic oxygen O1 of the donor. In this contour diagram, the horizontal axis represents the creation of a new C−O glycosidic bond, whereas the vertical axis represents scission of the C1−O1 glycosidic linkage of the donor. The Michaelis complex is located in the lower right corner of the PES. The structures in the lower left corner correspond to the structure with a pentacoordinated carbon, in which the C2, O1, O5, O6, and H1 atoms are bound to the anomeric carbon C1. As expected, this structure is unstable with a relative energy higher than 100 kcal mol^{-1}. In the upper left corner, we anticipated structures representing the product with the protonated nucleophilic oxygen O6, and, therefore, a high energy region. However, a low minimum was found in this area. Moreover, the relative energy (≈−18 kcal mol^{-1}) of this minimum is equivalent to the one in Figure 2 that represents the product. A thorough examination of the geometrical parameters of this minimum in Figure 3 clearly revealed, in contrast to what was observed in our previous studies,[22,24] that the H6 proton was transferred to the catalytic base Glu320 and the minimum actually corresponds to the product. A clear pathway that connects the ES and the product can be seen in this contour map. The pathway follows the diagonal of the two-dimensional map, from the lower-right corner to the upper-left corner, with a transition barrier (B12) of ≈24 kcal mol^{-1}. The barrier was detected in the central area of the

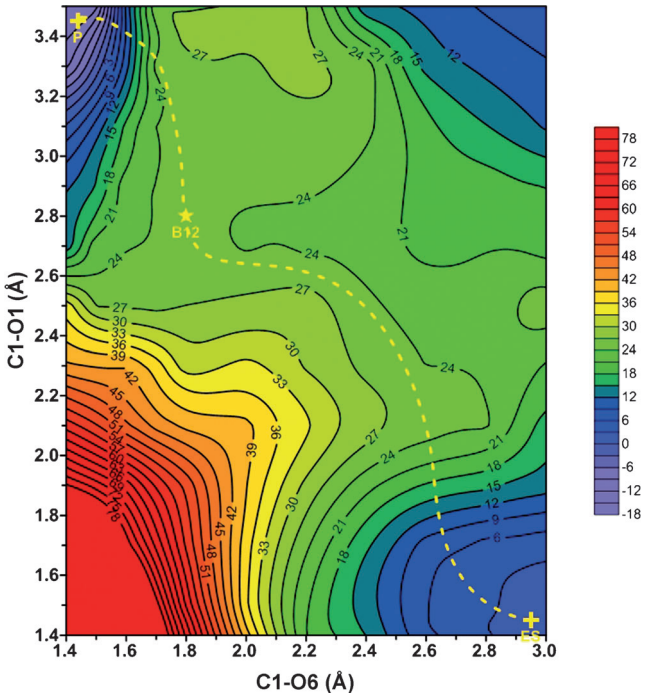


Figure 3. The B3LYP calculated PES with the 6-31G** basis set and diffuse functions (6-31+G**) on the O1′, O5′, O_{Glu}, and O6 atoms by using the distances r_1 and r_2 (in Å) as reaction coordinates. The yellow dashed line from the bottom-right corner to the upper-left corner indicates the S_N2-like reaction pathway from the ES to the P through B12. Contours are in kcal mol^{-1}.

contour map at $r_1 \approx 1.85$ Å and $r_2 \approx 2.9$ Å. An inspection of the geometrical parameters showed that the H6 proton was located at $r_3 \approx 1.5$ Å. It is noteworthy that the location of the B12 barrier coincides with the C1−O6, C1−O1, and H6−O_{Glu} distances determined in the transition barrier structure B13 observed in Figure 2. The structures at the energy maximum (B13 and B12) were used to refine the structure of the transition state (TS) with no geometry constraints. (Atomic coordinates for the stationary points calculated at the M06-2X level are available in the Supporting Information).

On the catalytic mechanism of C2GnT: Structure and experimental data[19,20] on C2GnT support the ordered bi-bi catalytic mechanism, with Glu320 proposed as the catalytic base for C2GnT. In this mechanism, the UDP-GlcNAc binds first and then the acceptor binds over the nucleotide sugar. An ordered bi-bi mechanism with UDP-GlcNAc binding first and UDP leaving last, is supported by the observation that C2GnT can bind the acceptor in the absence of the donor. However, this acceptor binding changes the conformation of C2GnT to one in which the binding of the donor is blocked.[20]

Previous calculations[23] performed on metal-ion dependent inverting glycosyltransferases of the GT-A family support a S_N2-like mechanism, in which the enzyme provides a catalytic base that activates the nucleophile to replace the UDP leaving group from the nucleotide sugar (donor) in a

concerted process and the metal presented facilitates the reaction by stabilizing the developing negative charge on the diphosphate leaving group. This study is the first calculation for an inverting metal-ion-independent glycosyltransferase from the GT-A family. The calculated two-dimensional maps of the potential energy surface suggest that C2GnT employs a S_N2-like mechanism, in which a nucleophilic attack by O6, the proton transfer to the catalytic base, and the separation of the leaving group all occur almost simultaneously. The energy barriers calculated at different levels of theory (Table 2) are between 20 and 29 kcal mol^{-1}. Unfortunately, the lack of experimental data on C2GnT does not enable a direct comparison. However, these values are in a good agreement with an activation barrier of 21 kcal mol^{-1} calculated from $k_{cat}=0.22$ min^{-1} determined for OGT,[41] which is also a metal-ion-independent enzyme, though with a GT-B fold. The single-point barriers are between 17 and 29 kcal mol^{-1}. The barrier calculated at the M06-2X level is slightly lower and at the MPW1K slightly higher than those obtained with others functionals.

Table 2. QM/MM energies (Hartree) of the Michaelis complex (ES) and the relative energies (kcal mol^{-1}) of the transition state (TS), and product complex (P) for the S_N2 reaction pathway calculated at various levels of theory by using the 6-31G** basis set with diffuse functions (6-31+G**) on O1′, O5′, O6, and O_{Glu320} atoms and with the MM subsystem characterized with an OPLS_2005 all-atom force field.

Method	Geometry	ES	TS	P
B3LYP	B3LYP	−6 283.668185	22.15	−15.93
MPW1K		−6 282.217352	27.44	−15.91
M05-2X		−6 283.067278	22.87	−18.94
M06-2X		−6 281.542623	22.95	−19.35
PWB6K		−6 286.707841	26.69	−17.65
MPW1K	MPW1K	−6 282.246039	28.80	−17.70
B3LYP		−6 283.643455	24.99	−16.80
M05-2X		−6 283.065913	21.93	−20.75
M06-2X		−6 281.539580	20.61	−21.91
PWB6K		−6 286.746165	25.91	−20.72
M05-2X	M05-2X	−6 283.077359	22.08	−19.67
B3LYP		−6 283.661677	23.88	−12.53
MPW1K		−6 282.236591	28.31	−12.71
M06-2X		−6 281.551464	19.83	−22.98
PWB6K		−6 286.733962	26.53	−17.68
M06-2X	M06-2X	−6 281.552218	19.54	−20.00
B3LYP		−6 283.662188	21.15	−6.06
MPW1K		−6 282.234912	25.12	−8.30
M05-2X		−6 283.075322	16.91	−16.79
PWB6K		−6 286.730800	22.96	−14.84
PWB6K	PWB6K	−6 286.740639	21.58	−22.73
B3LYP		−6 283.624046	19.25	−20.73
MPW1K		−6 282.236341	27.11	−21.69
M05-2X		−6 283.053336	18.97	−25.26
M06-2X		−6 281.526245	17.24	−26.70

Structural changes associated with the active site chemical reaction (ES→TS→P) are reflected in the structure of the points along the reaction pathways. The values in Table 1 clearly shows that all methods led essentially to the same structures of ES, TS, and P, respectively. The geometry of the Michaelis complex model (ES) at the B3LYP (MPW1K, M06-2X) level is described by values of 1.429 (1.411, 1.418) Å and 1.398 (1.380, 1.390) Å for the C1−O1 and C1−O5 bond lengths, respectively. The pyranoid ring of the GlcNAc in the ES is in the $^4C_1$ chair conformation depicted by the ring-puckering parameters $\phi=66.6°$, $\theta=12.4°$, and $Q=0.51$ (61.2, 12.1, 0.51°; 61.8, 12.1, 0.52°) within normal $^4C_1$ values for theta, but a bit flatter than usual for Q. The structure of the product (P) is described by values of 1.442 (1.421, 1.428) Å, 3.101 (3.055, 3.008) Å, and 1.407 (1.387, 1.399) Å for the C1−O6, C1−O1, and C1−O5 bond lengths, respectively. The ring structure of the transferred GlcNAc in P is depicted by the ring-puckering parameters $\phi=313.6°$, $\theta=31.1°$, and $Q=0.48$ (311.7, 28.4, 0.48°; 314.1, 30.0, 0.49°) and represent the distorted $^4C_1$ chair conformation.

Along the concerted reaction path ES→TS→P, the conformation of the pyranoid ring continuously changes from a $^4C_1$ chair in the ES through a $^4H_3/^4E$ half-chair in the TS (224.4, 36.9, 0.37; 231.3, 36.5, 0.38°; 230.3, 36.4, 0.39°) back to a $^4C_1$ chair conformation. In this process, the C1−O1 bond length between the anomeric carbon C1 and the leaving group, UDP, gradually elongates from 1.429 (1.411, 1.418) to 2.868 (2.887, 2.866) Å as the distance between the anomeric carbon and the attacking oxygen r_1 decreases, and the H1 atom moves, from an equatorial position through a position in the plane defined by the C2−C1−O5 atoms to an axial position. These modifications in the six-membered ring conformation of GlcNAc are accompanied by changes in the position of the leaving and attacking groups with respect to the six-membered ring. As the ring shape shifts to the half-chair conformation, the atoms attached to the anomeric carbon become coplanar with a sp^2 character at the reaction center C1. The delocalization of the ring-oxygen lone-pair electrons into the empty p orbital at the C1 atom stabilizes the oxocarbenium ion-like nature of GlcNAc.[42] The creation of such an oxocarbenium ion requires alterations to the GlcNAc ring conformation, from chair to half-chair, to satisfy its partial double-bond character. A consequence of the charge delocalization is the shortening of the C1−O5 bond length from its equilibrium value of 1.407 (1.387, 1.399) Å observed in the ES to 1.315 (1.302, 1.322) Å in the TS, in which the C1−O5 bond developed a partial double bond character. The length of the new glycosidic linkage C1−O6 (r_1) in the TS is characterized by value of 1.813 (1.778, 1.735) Å. The TS structure calculated at the M06-2X level is shown in Figure 4B.

In metal-ion-dependent glycosyltransferases, the donor is anchored into the active site through electrostatic interactions with a divalent cation that is coordinated by a DXD motif. The metal ion is assumed to reduce the reaction barrier through an electrostatic stabilization of the developing negative charge on the diphosphate moiety. It was hypothesized[20] that the positively charged side chains of Lys401 and Arg378 may play this role. The calculations support to some extent a stabilizing effect of these two amino acids. A comparison of the calculated distances between the Lys401 and Arg378 and diphosphate oxygens showed that phosphate oxygen–lysine proton distances are in the TS between 1.48

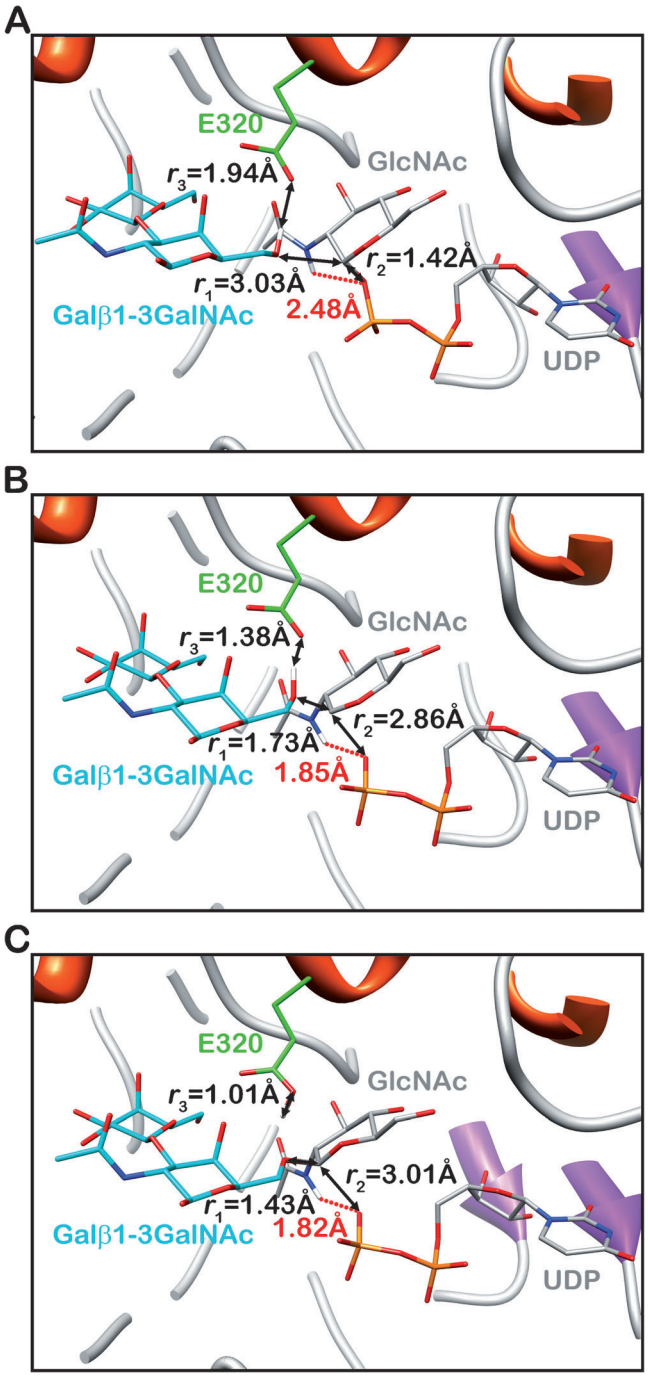


Figure 4. Close-up of the active site models for the A) Michaelis complex (ES), B) transition state (TS), and C) product complex (P) calculated at M06-2X level with the 6-31G** basis set and diffuse functions (6-31+G**) on the O1′, O5′, O_{Glu}, and O6 atoms. The acceptor is shown in cyan and the catalytic base in green.

and 1.89 Å and are slightly smaller than in the ES (1.57 Å and 1.90 Å).

The QM/MM calculations of the inverting glycosyltransferase OGT, which is a metal-ion-independent enzyme with GT-B fold, discovered other interactions that allow the catalytic reaction. In this substrate-assisted mechanism, the proton from the donor NAc group interacts with the oxygen from the diphosphate group and stabilizes the TS. A comparison of the distance between the H_{NAc} proton of the *N*-acetyl group and glycosidic oxygen O1 in C2GnT revealed similar interactions. In the TS, the H_{NAc}···O1 distance decreases from 2.48 in the ES to 1.85 Å in the TS. This shortening of the H_{NAc}···O1 distance is accompanying by a small conformational change of the *N*-acetyl group around the C2−N linkage. Although the *N*-acetyl group remains in its most stable conformation, called the *Z*-trans,[41] in the ES its conformation is characterized by the dihedral angle $\chi = -153.7°$ (χ = H2-C2-*N*-H_{NAc}), whereas in the TS the *N*-acetyl group rotates to a conformation with $\chi = -166.1°$. To evaluate how this hydrogen-bond interaction affects the stability of TS, we have calculated the energy of the ES and TS with the *N*-acetyl group replaced by a methyl group. In this analogue, the H_{NAc}···O1 hydrogen-bond interactions are not likely. The single-point calculations showed that reaction barrier in the 2-methyl analogue was, depending on the functional used, between 33 and 36 kcal mol^{-1}. Assuming the calculated barriers for the UDP-GlcNAc donor (Table 2), a crude estimate of the TS stabilization by the H_{NAc}···O1 hydrogen bonding is 11–14 kcal mol^{-1}. It is not yet clear whether this is a coincidence or a common mechanism used by metal-ion-independent GnTs. Also, the hydrogen-bond interaction between the acceptor hydroxyl group O6 and the carboxylate oxygen O_{Glu} of Glu320 seems to be important for the reaction. A decrease in the O6···O_{Glu} distance along the reaction pathway, from 2.94 in the ES to 2.41 Å in the TS indicates a stronger hydrogen-bond than in the ES. We suppose that this low-barrier-hydrogen-bond between H6 and O_{Glu} also contributes to the stabilization of TS.

Conclusion

C2GnT is a key enzyme in the biosynthesis of branched *O*-glycans and its activity has been found to be altered in various diseases. However, the mechanism of C2GnT, which could be used as a guide in developing effective inhibitors with therapeutic potential, has not yet been characterized. In the theoretical study presented here, QM/MM calculations were performed on the inverting metal-ion-independent *N*-acetylglucosaminyltransferase C2GnT to understand the catalytic mechanism of the GlcNAc transfer to the GalNAc residue of the core 1 structure on glycoproteins. The structural model of the reaction site used in this report is based on the crystal structures of C2GnT.[20] The entire enzyme–substrate system (C2GnT-UDP-GlcNAc-Galβ1-3GalNAc ternary complex, altogether 6120atoms), was partitioned into two separate subsystems: the QM subsystem, containing 206 atoms, and the MM region containing 5914atoms. The calculated potential energy surfaces clearly support a concerted S_N2-like displacement mechanism. In this mechanism, the nucleophilic attack, dissociation of the C1−O1 glycosidic linkage and proton transfer from the nucleophile oxygen to the catalytic base all occur simultaneously. Similarly to the inverting glycosyltransferase OGT

case, which is a metal-ion-independent enzyme with GT-B fold, we found that the hydrogen-bond interaction between the H_{NAc} proton of GlcNAc and the glycosidic oxygen O1 of the donor facilitates the breaking of the glycosidic linkage and the departure of the leaving group (UDP). The calculations also showed a key role played by the low-barrier-hydrogen-bond between the nucleophilic oxygen O6 and catalytic base Glu320. This QM/MM allowed for the first time a clear understanding of the catalytic mechanism of an inverting metal-ion-independent glycosyltransferase from the GT-A family with atomistic details.

Acknowledgements

This research was financially supported by the European Union from ERDF as part of the FP7 Program of the CEITEC (CZ.1.05/1.1.00/02.0068 project), SYLICA (Contract No. 286154 under the "Capacities" program), and SoMoPro (No. 2SGA2747, FP7 grant agreement No. 229603 with the South Moravian region contribution) projects. The authors thank the Czech National Supercomputing Centre, METACENTRUM, for providing computational resources (research intent LM2010005). This work was also supported by the Scientific Grant Agency of the Slovak Academy of Sciences (grants VEGA-02/0159/12) and Research & Development Operational Programs funded by ERDF (CEGreenI, Contract No. 26240120001 and CEGreenII, Contract No. 26240120025). We wish to thank Shinya Fushinobu for access to the Cremer–Pople parameter calculator (hi-ho).

[1] A. Varki, *Cell* **2006**, *126*, 841–845.
[2] a) M. Fukuda, *Biochimica Et Biophysica Acta-General Subjects* **2002**, *1573*, 394–405; b) J. B. Lowe, *Immunol. Rev.* **2002**, *186*, 19–36.
[3] C. Foxall, S. R. Watson, D. Dowbenko, C. Fennie, L. A. Lasky, M. Kiso, A. Hasegawa, D. Asa, B. K. Brandley, *J. Cell Biol.* **1992**, *117*, 895–902.
[4] a) K. D. Patel, S. L. Cuvelier, S. Wiehler, *Semin. Immunol.* **2002**, *14*, 73–81; b) D. Vestweber, J. E. Blanks, *Physiol. Rev.* **1999**, *79*, 181–213.
[5] a) H. Schachter, I. Brockhausen, in *Glycoconjugates. Composition, Structure, and Function* (Eds.: H. J. Allen, E. C. Kisailus), Marcel Dekker, Inc., New York, **1992**, pp. 263–332; b) D. Williams, H. Schachter, *J. Biol. Chem.* **1980**, *255*, 11247–11252.
[6] a) I. Brockhausen, J. Yang, M. Lehotay, S. Ogata, S. Itzkowitz, *Biol. Chem.* **2001**, *382*, 219–232; b) I. Brockhausen, J. M. Yang, J. Burchell, C. Whitehouse, J. Taylorpapadimitriou, *Eur. J. Biochem.* **1995**, *233*, 607–617; c) E. Machida, J. Nakayama, J. Amano, M. Fukuda, *Cancer Res.* **2001**, *61*, 2226–2231; d) K. Shimodaira, J. Nakayama, N. Nakamura, O. Hasebe, T. Katsuyama, M. Fukuda, *Cancer Res.* **1997**, *57*, 5201–5206.
[7] a) I. Brockhausen, W. Kuhns, H. Schachter, K. L. Matta, D. R. Sutherland, M. A. Baker, *Cancer Res.* **1991**, *51*, 1257–1263; b) O. Saitoh, F. Piller, R. I. Fox, M. Fukuda, *Blood* **1991**, *77*, 1491–1499.
[8] A. Orlacchio, P. Sarchielli, V. Gallai, A. Datti, C. Saccardi, C. A. Palmerini, *J. Neurol. Sci.* **1997**, *151*, 177–183.
[9] R. Chibber, B. M. Ben-Mahmud, D. Coppini, E. Christ, E. M. Kohner, *Diabetes* **2000**, *49*, 1724–1730.
[10] a) E. A. Higgins, K. A. Siminovitch, D. L. Zhuang, I. Brockhausen, J. W. Dennis, *J. Biol. Chem.* **1991**, *266*, 6280–6290; b) F. Piller, F. Ledeist, K. I. Weinberg, R. Parkman, M. Fukuda, *J. Exp. Med.* **1991**, *173*, 1501–1510.
[11] a) M. F. A. Bierhuizen, M. Fukuda, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 9326–9330; b) M. F. A. Bierhuizen, K. Maemura, S. Kudo, M. Fukuda, *Glycobiology* **1995**, *5*, 417–425.
[12] T. Schwientek, J. C. Yeh, S. B. Levery, B. Keck, G. Merkx, A. G. van Kessel, M. Fukuda, H. Clausen, *J. Biol. Chem.* **2000**, *275*, 11106–11113.
[13] a) T. Schwientek, M. Nomoto, S. B. Levery, G. Merkx, A. G. van Kessel, E. P. Bennett, M. A. Hollingsworth, H. Clausen, *J. Biol. Chem.* **1999**, *274*, 4504–4512; b) J. C. Yeh, E. Ong, M. Fukuda, *J. Biol. Chem.* **1999**, *274*, 3215–3221.
[14] A. Vanderplasschen, N. Markine-Goriaynoff, P. Lomonte, M. Suzuki, N. Hiraoka, J. C. Yeh, F. Bureau, L. Willems, E. Thiry, M. Fukuda, P. P. Pastoret, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5756–5761.
[15] T.-Y. Yen, B. A. Macher, S. Bryson, X. Chang, I. Tvaroska, R. Tse, S. Takeshita, A. M. Lew, A. Datti, *J. Biol. Chem.* **2003**, *278*, 45864–45881.
[16] X. Yang, W. Qin, M. Lehotay, D. Toki, P. Dennis, J. S. Schutzbach, I. Brockhausen, *Biochim. Biophys. Acta Proteins Proteomics* **2003**, *1648*, 62–74.
[17] B. L. Cantarel, P. M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, B. Henrissat, *Nucleic Acids Res.* **2009**, *37*, D233–D238.
[18] a) L. L. Lairson, B. Henrissat, G. J. Davies, S. G. Withers, *Annu. Rev. Biochem.* **2008**, *77*, 521–555; b) I. Tvaroska, *Trends Glycosci. Glycotechnol.* **2005**, *17*, 177–190; c) I. Tvaroska, *Mini-Rev. Org. Chem.* **2011**, *8*, 263–269; d) R. Zhang, V. L. Y. Yip, S. G. Withers, in *Comprehensive Natural Products II* (Eds.: M. Editors-in-Chief: Lew, L. Hung-Wen), Elsevier, Oxford, **2010**, pp. 385–422.
[19] J. E. Pak, P. Arnoux, S. H. Zhou, P. Sivarajah, M. Satkunarajah, X. K. Xing, J. M. Rini, *J. Biol. Chem.* **2006**, *281*, 26693–26701.
[20] J. E. Pak, M. Satkunarajah, J. Seetharaman, J. M. Rini, *J. Mol. Biol.* **2011**, *414*, 798–811.
[21] P. K. Qasba, B. Ramakrishnan, E. Boeggeman, *Trends Biochem. Sci.* **2005**, *30*, 53–62.
[22] I. Tvaroska, I. Andre, J. P. Carver, *Glycobiology* **2003**, *13*, 559–566.
[23] a) S. Kozmon, I. Tvaroska, *J. Am. Chem. Soc.* **2006**, *128*, 16921–16927; b) M. Krupicka, I. Tvaroska, *J. Phys. Chem. B* **2009**, *113*, 11314–11319.
[24] I. Tvaroska, S. Kozmon, M. Wimmerova, J. Koca, *J. Am. Chem. Soc.* **2012**, *134*, 15563–15571.
[25] a) *MODELLERv9.9*, *A Program for Protein Structure Modeling 9.9*, University of California, San Francisco, **2011**; b) A. Sali, T. L. Blundell, *J. Mol. Biol.* **1993**, *234*, 779–815.
[26] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605–1612.
[27] *Maestro 9.2*, Schrodinger, LLC, New York, **2011**.
[28] a) *Epik 2.2*, Schrodinger, LLC, New York, **2011**; b) *Impact 5.7*, Schrodinger, LLC, New York, **2011**; c) *Prime 3.0*, Schrodinger, LLC, New York, **2011**; d) *Schrodinger Suite 2011*, *Protein Preparation Wizard 2011*, Schrodinger, LLC, New York, **2011**.
[29] a) *Glide 5.7*, Schrodinger, LLC, New York, NY, **2011**; b) R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, P. S. Shenkin, *J. Med. Chem.* **2004**, *47*, 1739–1749; c) R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin, D. T. Mainz, *J. Med. Chem.* **2006**, *49*, 6177–6196; d) T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard, J. L. Banks, *J. Med. Chem.* **2004**, *47*, 1750–1759; e) M. S. Park, C. Gao, H. A. Stern, *Proteins Struct. Funct. Bioinf.* **2011**, *79*, 304–314.
[30] W. L. Jorgensen, D. S. Maxwell, J. TiradoRives, *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
[31] *Jaguar 7.8*, Schrodinger, LLC, New York, **2011**.
[32] a) A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648–5652; b) C. T. Lee, W. T. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789.
[33] a) *QSite 5.7*, Schrodinger, LLC, New York, NY, **2011**; b) R. B. Murphy, D. M. Philipp, R. A. Friesner, *J. Comput. Chem.* **2000**, *21*, 1442–1457; c) D. M. Philipp, R. A. Friesner, *J. Comput. Chem.* **1999**, *20*, 1468–1494.
[34] a) J. Kóňa, I. Tvaroška, *Chem. Pap.* **2009**, *63*, 598–607; b) L. Simón, J. M. Goodman, *Org. Biomol. Chem.* **2011**, *9*, 689–700; c) X. F. Xu,

I. M. Alecu, D. G. Truhlar, *J. Chem. Theory Comput.* **2011**, *7*, 1667–1676; d) J. J. Zheng, Y. Zhao, D. G. Truhlar, *J. Chem. Theory Comput.* **2009**, *5*, 808–821.
[35] Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.* **2008**, *120*, 215–241.
[36] B. J. Lynch, P. L. Fast, M. Harris, D. G. Truhlar, *J. Phys. Chem. A* **2000**, *104*, 4811–4815.
[37] Y. Zhao, D. G. Truhlar, *J. Phys. Chem. A* **2005**, *109*, 5656–5667.
[38] Y. Zhao, N. E. Schultz, D. G. Truhlar, *J. Chem. Theory Comput.* **2006**, *2*, 364–382.
[39] a) O. Hindsgaul, K. J. Kaur, G. Srivastava, M. Blaszczyk-Thurin, S. C. Crawley, L. D. Heerze, M. M. Palcic, *J. Biol. Chem.* **1991**, *266*, 17858–17862; b) W. Kuhns, V. Rutz, H. Paulsen, K. L. Matta, M. A. Baker, M. Barner, M. Granovsky, I. Brockhausen, *Glycoconjugate J.* **1993**, *10*, 381–394.
[40] I. Tvaroska, I. Andre, J. P. Carver, *J. Am. Chem. Soc.* **2000**, *122*, 8762–8776.
[41] M. B. Lazarus, Y. S. Nam, J. Y. Jiang, P. Sliz, S. Walker, *Nature* **2011**, *469*, 564–U168.
[42] G. Davies, M. L. Sinnott, S. G. Withers in *Comprehensive Biological Catalysis* (Ed.: M. L. Sinott), Academic Press Limited, New York, **1997**, pp. 119–208.
[43] P. Fowler, B. Bernet, A. Vasella, *Helv. Chim. Acta* **1996**, *79*, 269–287.

Received: January 31, 2013
Revised: March 15, 2013
Published online: April 24, 2013

