O-/N-/S-specificity in glycosyltransferase catalysis: From mechanistic understanding to engineering

Tezé, David; Coines, Joan; Fredslund, Folmer; Dubey, Kshatresh D.; Bidart, Gonzalo N.; Adams, Paul D.; Dueber, John E.; Svensson, Birte; Rovira, Carme; Welner, Ditte Hededam

Published in:
ACS Catalysis

Link to article, DOI:
10.1021/acscatal.0c04171

Publication date:
2021

Document Version
Peer reviewed version

Link back to DTU Orbit

Citation (APA):
O-/N-/S-specificity in Glycosyltransferase Catalysis: from Mechanistic Understanding to Engineering.

1,2David Teze*, 3Joan Coines‡, 2Folmer Fredslund, 3Kshatresh D. Dubey‡, 2Gonzalo N. Bidart, 4,5,6Paul D. Adams, 4,7John E. Dueber, 1Birte Svensson, 3,8Carme Rovira*, 1,2,5,6Ditte H. Welner*

1Department of Biotechnology and Biomedicine, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark. 2The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark. 3Departament de Química Inorgànica i Orgànica and Institut de Química Teòrica i Computacional (IQTCUB), Universitat de Barcelona, Barcelona, Spain. 4Department of Bioengineering, University of California, Berkeley, California, USA. 5Joint BioEnergy Institute, Emeryville, California, USA. 6Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA. 7Biological Systems & Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA. 8Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluis Companys 23, 08018, Barcelona, Spain.

Keywords: Enzymes • O-glycosylation • N-glycosylation • S-glycosylation • Carbohydrates • Glycosyltransferases • Quantum mechanics/Molecular mechanics

ABSTRACT: Glycosyltransferases (GTs) catalyze the formation of glycosidic bonds in carbohydrates and glycoconjugates, with various outcomes depending not only on the acceptor molecules they bind, but also on the type of glycosidic bond they form (C-O, C-N, C-S or C-C). Here we show that PtUGT1 from Polygonum tinctorium catalyzes either N-, O- or S-glycosylation with similar rates. We solve the structure of the enzyme in complex with its donor and acceptor substrates, and elucidate the molecular basis of N-, O- and S-specificities using experimental mutagenesis and QM/MM simulations, revealing distinct mechanisms for N-, O- and S-glycosylation. We also show that the active site can be engineered to increase or favor one of the three glycosylation activities over another. These results will foster the design of more active and specific enzyme variants for glycosides production.

Glycosides are ubiquitous in the biosphere, and encompass a myriad of compounds. The formation of glycosidic bonds is typically catalyzed by glycosyltransferases (GTs). These enzymes transfer a monosaccharide from an activated sugar donor to an acceptor substrate. Understanding the factors that control acceptor specificity would increase the value of GTs as biotechnological catalysts as well as enhance our insight in the important biological processes involving glycosides. As such, structure-function relationships governing the specificity of their substrates are a focus of enormous interest.

GTs, like most carbohydrate-active enzymes, are classified in families within the CAZy database3,4 (www.cazy.org). Two-thirds of GT families (74/111 on Sept 22nd, 2020) catalyze glycosidic bond formation with inversion of the anomeric configuration (inverting GTs), while the others results in retention of the anomeric configuration (retaining GTs). Whereas two different mechanisms have been proposed and thoroughly investigated for retaining GTs,5,6 both experimental and theoretical evidence supports a common mechanism for inverting GTs7–9 (Scheme 1). It consists of an S_{2}2-like reaction, involving the deprotonation of an acceptor by a catalytic base. Barring few exceptions,10–14 this base is an Asp, Glu or His residue.7

Glycosyltransferases from family GT1 are particularly interesting for their ability to catalyze the formation of several types of glycosidic bonds. These enzymes act on a vast array of acceptors in plants, animals and bacteria. Bifunctional O-/N-1,7,15 and O-/S-transf

Yet, the limited number of structures of GT1 enzymes in complex with donors and/or acceptors15 hinders dissecting their mechanism of action in detail. In addition, a direct kinetic comparison on structurally similar acceptors is not available, thus the determinants of the enzyme’s specificity for R-OH, R-NH2 and R-SH substrates is poorly understood and it is assumed that all of them employ a similar mechanism.16

Mutating the acid-base residue of O-active retaining glycoside hydrolases, while providing activated glycosyl donors and thiolated acceptors, represents the thioglycosidase approach, which allows for the formation of S-glycosidic bonds.19 This approach has been recently expanded to O- and N-glycosylation, providing that the chosen acceptors are good nucleophiles themselves. O-GTs have also been used for the synthesis of S-linked oligosaccharides, but the corresponding reactions exhibit drastically lower reaction rates compared to the natural O-glycosylation activity.22

Here we report a GT1 member that is able to catalyze O-, N- and S-glycosylation with similar rates. Our combined structural, mutational, kinetic and computational study is focused on the GT1 enzyme from the indigo plant Polygonum tinctorium (PtUGT1). Its natural O-glycosylation activity against indoxyl (Scheme 1) has been recently employed in an eco-friendly indigo dyeing strategy.23 PtUGT1 belongs to a phylogenetic subgroup of GT124–26 that includes enzymes featuring an active site histidine, interacting with an aspartate, as their catalytic base (Scheme 1, top).27 We reveal that PtUGT1 is able to catalyze O-, N- and S-glycosylation with similar rates on aryl acceptors that differ solely in their nucleophilic
atom. We also show that monofunctional N- or S-glycosylating activities, as well as >400% increased N-glycosylation activity can be achieved by appropriate mutation of active site residues. We rationalize our observations through X-ray structures in complexes with donor and acceptor, as well as QM/MM metadynamics simulations, providing a molecular basis for enzyme specificity. Finally, we establish that there are significant mechanistic differences between N-, O- and S-glycosylation activities.

Scheme 1. Inverting GT1 mechanism presented according to the catalyst (magenta), donor (black) and acceptors (grey) used in this study.

To understand catalysis and substrate specificity in PtUGT1, we undertook the kinetic evaluation of PtUGT1-wt and mutants of its catalytic residues against aromatic acceptors that vary solely in the identity of the nucleophilic atom: 3,4-dichlorophenol (DCP, hereafter also named O-acceptor), 3,4-dichloroaniline (DCA, N-acceptor) and 3,4-dichlorothiophenol (DCT, S-acceptor). Mutations were rationally designed to assess how chemical properties (basicity, size, aromaticity, hydrogen bond donor/acceptor) contribute to specificity, focusing on the catalytic dyad. The results obtained, summarized in Figure 1 and Table S1, show that PtUGT1-wt catalyzes O-, N- or S-glycosylation with similar rates, but acceptor specificity can be drastically altered by appropriate mutations. In particular, essentially monospecific N-glycosylation is achieved for PtUGT1-His26Ala and monospecific S-glycosylation for PtUGT1-Asp122Asn, while PtUGT1-His26Glu displays improved N-glycosylation ($k_{cat}$ increases from 71.5 to 292.6 min$^{-1}$).

Figure 1 also shows that some mutations can disable only one of the three glycosylation functions. For instance, PtUGT1-His26Gln exhibits N- and S-glycosylation activities similar as PtUGT1-wt, while being devoid of O-glycosylation (Figure 1). This suggests that O-glycosylation needs the His26/Asp122 pair to deprotonate the phenol group, whereas N- and S-glycosylation proceed without acceptor deprotonation. In fact, N-glycosylation operating without deprotonation was proposed for the bi-functional O-/N- transferase AtrUGT72B1,15 which glycosylate monoglycosyls37 and shares 61% sequence identity with PtUGT1. It was suggested that aniline compounds would form glycosylated aniliniums ($pK_a(DCAH^+)$ = 3)$^{28}$, which would later be deprotonated by the solvent.15

Of PtUGT1’s three functionalities, O-glycosylation is the most sensitive to modification of the catalytic machinery, since all variants display a complete loss of activity with the O-acceptor (Figure 1). S-glycosylation is conversely mostly unaffected as long as a heteroatom (O or N) is present in the side chain of residue 26; variants His26Ala and His26Phe being the only ones with severely decreased activities. Overall, our results suggest that the factors governing glycosylation are distinct for O-, N- or S-acceptors.

Our results suggest that the S-acceptor performs nucleophilic attack in its thiophenolate form at neutral pH ($pK_a(DCT) = 5.47$), since S-glycosylation is mostly independent on the nature of residue 26. By contrast, the O-acceptor probably attacks in its phenolic form ($pK_a(DC) = 8.22^{29/8.59}$) and requires deprotonation, consistently with the lack of O-glycosylation activity for all mutants. Moreover, no recovery of O-activity was observed when these variants were tested at pH=8.0, 8.5, 9.0 and 9.5, where the dichlorophenolate species would be increasingly predominant in solution. PtUGT1-wt retains most of its activity at these pH values (Figure S1). This suggest the active site environment is acidic enough to protonate DCP.

From these kinetic results, the trifunctionality and the differential disruption of the catalytic machinery of the O-/N-/S-activities by the studied variants have no straightforward explanation. Therefore, we turned to simulation informed by X-ray crystallography, to probe the molecular events leading to our observations.

We solved the crystal structures of the binary complexes of PtUGT1 with either the donor uridine diphosphate glucose (UDP-Glc, PDB ID: 6SU6) or the N-acceptor DCA (PDB ID: 6SU7) bound in the active site, and used these structures to model the enzyme ternary complex (Figure 2a, crystallography statistics can be found in Table S2). The ternary structure is consistent with the canonical binding of UDP-Glc in GT1, while DCA is too far away from the reaction center to be catalytically productive (N$_1^{DCA}$ located 6.2 Å from C1 of 3,4-dichloroaniline (Figure 2a). Therefore, full mechanistic detail could not be inferred directly from the structures. Molecular dynamics (MD) simulations (see methods and Figures S2-S6, S9) were performed starting from the ternary complex structure, showing that the N-acceptor quickly...
moves into a catalytically competent position (N1DCA located 4.7 Å from C1UDP-Glc, and H1DCA at 2.0 Å from NεHis26, Figure 2b). Favorable configurations were also obtained after in silico mutagenesis of His26 to Asp (Figure 2c) or Glu (Figure 2d). In contrast, the His26Phe mutation led to a stacking of Phe26 and the glucose moiety, preventing DCA from approaching the reaction center (Figures 2e and S5d. This explains PtUGT1-His26Phe N-glycosylation activity loss (Figure 1). Therefore, our results suggest that the ability of the acceptor to come close to the donor is the critical factor for N-glycosylation activity.

**Figure 2.** N-Glycosylation. (a) Ternary complex modelled by superimposing the binary complexes (PtUGT1:UDP-Glc (PDB ID: 6SU6) in black sticks and green density, and PtUGT1:DCA (PDB ID: 6SU7) in grey sticks and blue density) obtained by X-ray crystallography. Selected snapshots along the MD simulations of PtUGT1 variants in presence of DCA. (b) wild-type; (c) His26Asp; (d) His26Glu; (e) His26Phe.

Initial structures for the complexes with the O- and S-acceptors were obtained by substitution of the nucleophilic atom by O and S, respectively. MD simulations of PtUGT1-wt show that these acceptors reach catalytically active configurations in the active site, consistent with the observed high O- and S-glycosylation activities. However, simulations with His26 mutated to Asp or Glu show that both DCP and DCT also reach the active site, adopting orientations that should in principle be favorable for catalysis. This was surprising since only the O-acceptor (DCP) shows activity when the catalytic base is mutated (Figure 1). This prompted us to go one step forward in accuracy and predictive power by using first principles methods, thus considering electronic effects and charge rearrangements in the active site. In particular, we used quantum mechanics/molecular mechanics (QM/MM) metadynamics methods,30,31 based on Density Functional Theory (DFT)12–35 to model O- and S-glycosylation by PtUGT1.

QM/MM MD simulations show that the proton on the hydroxyl group of DCP jumps back and forth between DCP and His26 within the PtUGT1-wt active site, increasing its nucleophilicity (Figure S7). It is also clear from the computed Michaelis complex (R in Figure 3a) that the acceptor is well positioned for the SN2 reaction to take place. In contrast, the phosphate-C1-O-acceptor angle in the His26Asp mutant departs significantly from the ideal 180° (Figure S8), while the proton acceptor ability of Asp26 is limited by its interaction with Asp122. As a consequence, whereas the computed energy barrier for the glycosylation reaction in PtUGT1-wt (Figure 3a) is indicative of a feasible reaction (17.2 kcal·mol⁻¹, experimentally 17.6 kcal·mol⁻¹), that of the His26Asp mutant is very high (39 kcal·mol⁻¹), consistent with kinetic observations. Therefore, our QM/MM results reveal that the orientation of the acceptor in the active site, only achieved when His26 is present, is crucial for O-glycosylation catalyzed by PtUGT1.

**Figure 3.** O-glycosylation. (a) Free energy profile for O-glycosylation catalyzed by PtUGT1-wt and PtUGT1-His26Asp (dashed line). (b,c) Average structures corresponding to relevant configurations of the active site along the reaction coordinate: R=reagents, TS=transition state, P=products.

The QM/MM simulations also show that, even though His26 does deprotonate the O-acceptor easily, the nucleophilic attack occurs before deprotonation (Figure 3b, do,H = 1.09±0.05 and 1.11±0.10 Å at R and TS, respectively). On the other hand, deprotonation happens before the nucleophilic attack with PtUGT1-His26Asp. These observations are most likely transferable to the His26Glu variant. Only the His26-Asp122 dyad provides the right
catalytic machinery for O-glycosylation by PtUGT1, consistent with its high conservation through evolution.

The mechanistic scenario is very different in the case of S-glycosylation. The DCT acceptor is likely to perform the nucleophilic attack in its thiophenolate form and thus does not need deprotonation. As shown in Figure 4, the acceptor is perfectly poised for nucleophilic attack in PtUGT1-His26Asp (the O\textsubscript{phosphate}-C\textsubscript{1Glc}-S\textsubscript{acceptor} angle is close to 180 degrees, Figure S8). The computed free energy barrier of PtUGT1-His26Asp with DCT is 18.7 kcal mol\(^{-1}\), in good agreement with experiments ($\Delta G^\ddagger = 18.5$ kcal mol\(^{-1}\)). Additionally, the negative charge on the thiophenolate group further stabilizes the positive charge being developed on the C\textsubscript{1Glc} at the reaction TS, compared with the phenol group of DCP. Hence, there is no direct involvement of Asp26 in the catalysis (Figure 4b), which effectively takes place without the participation of any catalytic residue. This explains the limited effects of His26 mutations on S-glycosylation.

The results for the three acceptors indicate that substrate selectivity in PtUGT1 is governed by structural and electrostatic factors: O-glycosylation requires the native His-Asp dyad to catalyze the nucleophilic attack via acceptor deprotonation by His26. In contrast, both N- and S-glycosylation happen without proton transfer with a catalytic residue, and depend on the acceptor positioning relative to the donor.

Further structural analysis shows that the acceptor locates above the donor glucose ring in a hydrophobic pocket constituted by Leu123, Phe124, Tyr149, Phe153, Ile188 and Leu202 (Figure S10), defining a cavity in which acceptors with an aromatic group can be easily accommodated. Out of 1356 sequences sharing 20–80% pairwise identity (average 42±8.5%), none of these hydrophobic residues is conserved, yet charged residues are not observed and polar residues amount for only 1% of observed residues at these positions, which relates to the documented promiscuity of GT1 enzymes with respect to hydrophobic acceptors. Moreover, His26 is 99.4% conserved while neither Glu nor Asp residues are observed. Since His26 and Asp122 are conserved within this GT1 family subgroup, one can expect the herein reported effects of mutations to be transferable to other enzymes, hence providing a template for tailoring sought-after catalysts.

In summary, we characterized a truly trifunctional GT1 enzyme, solved crystal structures with both donor and acceptor, and investigated acceptor specificity with mutagenesis and computer simulations. Our results reveal significant mechanistic differences with aryl acceptor type, which can be exploited to engineer the active site to achieve O-, N- and S-activities selectively. These findings will foster the design of more active and specific GT1 variants for biotechnological applications.

ASSOCIATED CONTENT
Supporting Information
Materials and Methods
Figure S1: pH profile of PtUGT1 activity
Figures S2-S4: RMSD evolution from MD simulations
Figures S5-S9: distance analyses from MD simulations
Figure S10: acceptor subsite of PtUGT1
Table S1: $k_{cat}$ values of PtUGT1-wt and its variants
Table S2: crystallographic statistics

The Supporting Information (PDF file) is available free of charge on the ACS Publications website.

AUTHOR INFORMATION
Corresponding Author
*datez@dtu.dk, diwel@biosustain.dtu.dk and c.rovira@ub.edu

Present Addresses
†(KD) Department of Chemistry, School of Natural Sciences, Shiv Nadar University, U.P. India

Author Contributions
‡(DT, JC) These authors contributed equally to the study.

Notes
The authors declare no competing financial interests.

ACKNOWLEDGMENT
This work was supported by the Novo Nordisk Foundation (grants NNF17OC0025660, NNF10CC1016517 and NNF16OC0019088 to D. T., F. F., G. N. B. and D. W.), the Spanish Ministry of Science and Innovation (AEI/FEDER, UE) (grant CTQ2017-85-496 to C. R.), the Spanish Structures of Excellence Maria de Maeztu (grant MDM-2017-0767 to C. R.), and the Agency for Management of University and Research Grants (AGAUR) (grant SGR2017-1189 to C. R.). J. C. thanks MICINN for a predoctoral fellowship (FPI-
REFERENCES


