Family 1 glycosyltransferases (GT1, UGTs) are subject to dilution-induced inactivation and low chemo stability towards their own acceptor substrates

David Teze*, Gonzalo N. Bidart, Ditte H Welner*

The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark.

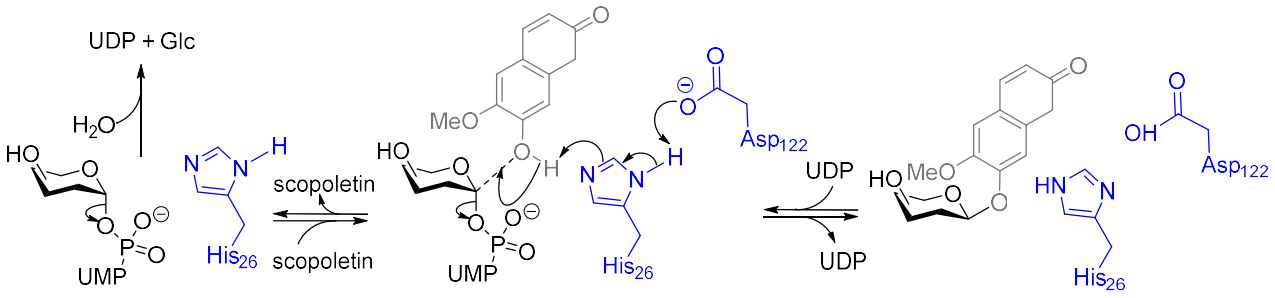
*** Correspondence:**
David Teze (datez@biosustain.dtu.dk) Ditte H. Welner (diwel@biosustain.dtu.dk)

Keywords: Glycosyltransferases, Glycosylation, Biotechnology, Stability, Polyphenols

Abstract. Glycosylation reactions are essential but challenging from a conventional chemistry standpoint. Conversely, they are biotechnologically feasible as glycosyltransferases can transfer a sugar to an acceptor with perfect regio- and stereo- selectivity, quantitative yields, in a single reaction and under mild conditions. Low stability is often alleged to be a limitation to the biotechnological application of glycosyltransferases. Here we show that these enzymes are not necessarily intrinsically unstable, but that they present both dilution-induced inactivation and low chemostability towards their own acceptor substrates, and that these two phenomena are synergistic. We assessed 18 distinct GT1 enzymes against three unrelated acceptors (apigenin, resveratrol and scopoletin – respectively a flavone, a stilbene and a coumarin), resulting in a total of 54 enzyme:substrate pairs. For each pair, we varied catalyst and acceptor concentrations to obtain 16 different reactions conditions. Fifteen of the assayed enzymes (83 %) displayed both low chemostability against at least one of the assayed acceptors at submillimolar concentrations, and dilution-induced inactivation. Further, there is a likely correlation between sensitivity to reaction conditions and thermal stability of the enzymes, the three unaffected enzymes having melting temperatures above 55 °C, whereas the full enzyme panel ranged from 37.4 to 61.7 °C. These results are important for GT1 understanding and engineering, as well as for discovery efforts and biotechnological use.

Introduction

Glycosylation is one of the most common reactions in the biosphere, yet a particularly challenging one for conventional synthetic chemistry. Indeed, the need to control both regio- and stereo-selectivity leads to a succession of reactions, including protecting group manipulations and bond activations, resulting in low chemical yields, poor atom economy and large amounts of waste. Conversely, enzymatic glycosylation occurs in a single reaction with unprotected sugars and acceptors, and lends perfect control over stereoselectivity. Provided with the appropriate enzyme, full control over regioselectivity, as well as quantitative chemical yields are also feasible. In Nature, glycosylation is primarily catalyzed by glycosyltransferases, enzymes that transfer a saccharide from an activated sugar donor to an acceptor molecule. These enzymes are organized in >100 distinct glycosyltransferase families in the CAZy database (Coutinho *et al.*, 2003; Lombard *et al.*, 2014), with all enzymes within a family sharing phylogeny, structural fold, and mechanism. The β-glycosylation of natural products is mainly achieved by enzymes from glycosyltransferase family 1 (GT1) (Louveau and Osbourn, 2019). These GT1s are inverting enzymes using α-nucleotide sugars as donors, most commonly UDP-sugars, and are thus also termed UGTs, for UDP-dependent glycosyltransferases (Ross *et al.*, 2001). They catalyze the formation of *O*-, *N*-, *S*- or *C*-glycosidic bonds. *O*-glycosylations are the most common reactions, and are usually promoted by a His-Asp catalytic dyad sharing a proton abstracted from the acceptor (Scheme 1) (Brazier-Hicks *et al.*, 2007; Teze *et al.*, 2021). The *N*- and *S*- mechanisms are slightly different (Teze *et al.*, 2021), and the *C*-glycosylation mechanism is related but yet to be firmly established (Gutmann and Nidetzky, 2013; Putkaradze *et al.*, 2021). GT1 enzymes are relatively promiscuous, being able to act on a variety of natural products (Offen *et al.*, 2006; Chen *et al.*, 2015; Zhang *et al.*, 2022), and most GT1s are active against polyphenols (Yang *et al.*, 2018).



Scheme 1. Scopoletin glucosylation by GT1s. Amino acids numbers from *Pt*UGT1 (Teze *et al.*, 2021). The enzyme residues are represented in blue, the acceptor in grey and the donor in black. A His activated by an Asp acts as a general base, increasing the nucleophilicity of the acceptor. The predominant reaction is the reversible glycosylation of scopoletin, which is the reaction depicted by the arrows. The enzyme can also catalyze the irreversible hydrolysis of the donor uridine diphosphate glucose (UDP-Glc).

GT1s have received considerable interest as tools for biotechnological glucosylation (Nidetzky, Gutmann and Zhong, 2018; Vasudevan and Lee, 2020). Indeed, the possibility to use sucrose synthase for forming UDP-Glc from UDP and sucrose, and to use lysates from the enzyme's production as UDP provider, makes β-glucosylation an economically feasible process (Wang *et al.*, 2012; Schmölzer *et al.*, 2016; Liu and Nidetzky, 2021). However, their stability – a crucial industrial property – has only been scarcely characterized (Fujiwara *et al.*, 2009; Gao *et al.*, 2020). In a few recent cases (Petermeier *et al.*, 2021; Bidart *et al.*, 2022), we observed instability, seemingly not intrinsic but dependent on experimental conditions, and particularly enzyme and acceptors concentrations. Indeed, a non-linear behavior was observed upon enzyme dilution, particularly at acceptor substrate concentrations in the millimolar range (Petermeier *et al.*, 2021; Bidart *et al.*, 2022). In order to investigate how widespread this peculiar behavior is within GT1-catalyzed reactions, we analyzed the effect of 16 reaction conditions on end-point reaction yields from 18 distinct GT1 enzymes, each against three different polyphenol acceptors.

Materials and Methods

Protein production, purification and storage. Proteins are expressed in One Shot™ BL21 Star™ (DE3) *E. coli* cells (ThermoFisher Scientific, USA) cells transformed with pET28a+ plasmids encoding the various enzymes with a hexahistidine tag and a TEV cleavage site in N-term (plasmids purchased from Genscript, USA). Protein expression is induced by the addition of 200 µM of isopropyl-β-D-galactopyranoside to cultures that had reach an optical density at 600 nm of 0.6 and continued for 16 h at 293 K. The cultures are then centrifuged, and the pellet is resuspended in 50 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7, 300 mM NaCl, and 20 mM imidazole. The cell suspension is lysed in a homogenizer (French Press) Avestin Emulsiflex C5 (ATA Scientific Pty Ltd., Canada), centrifuged and the pellet is discarded. The supernatant is purified by nickel affinity chromatography on an ÄKTA pure (GE Healthcare, U.S.). The fractions containing the purified GT1 are pooled, concentrated, buffer exchanged against 25 mM HEPES pH 7, 50 mM NaCl, and 1 mM dithiothreitol (DTT), then stored at 193 K after flash-freezing in 25 µL aliquots.

Enzymatic reactions and yield determination. All reactions were performed in flat-bottom, low sorption 96-well microtiter plates, in the following conditions: 100 µL volume, no stirring, 20 h at 293 K. The reaction components were 10, 20, 40 or 80 mg/L (circa 0.15–1.2 µM) protein; UDP-Glc 500 µM; 50, 100, 200 or 400 µM aglycon; 25 mM HEPES pH 7. After 20 h, reactions were diluted 25-fold in milli-Q water (10+240 µL), and analyzed by reverse-phase chromatography. Acceptor consumption was monitored using an Ultimate 3000 Series apparatus (Thermo Scientific) and an Eclipse Plus C18 3.5 µm 100x4.6 mm analytical column (Agilent). Milli-Q water containing 0.1% formic acid and acetonitrile were used as mobile phases A and B, respectively. Monitoring and data handling was operated using the Chromeleon software (Thermo Scientific). A combinations of isocratic, immediate ramp and gradients at a flow rate of 1 mL/min was used for the analytes separation:

0–0.5 min, 2% B; 0.5–1.5 min, 35% B; 1.5–3 min, 35–80% B; 3–4.2 min, 98% B; 4.2–5 min, 2% B. Apigenin and scopoletin were monitored at 340 nm, resveratrol at 300 nm. Data points for which acceptor consumption did not match products appearance were discarded.

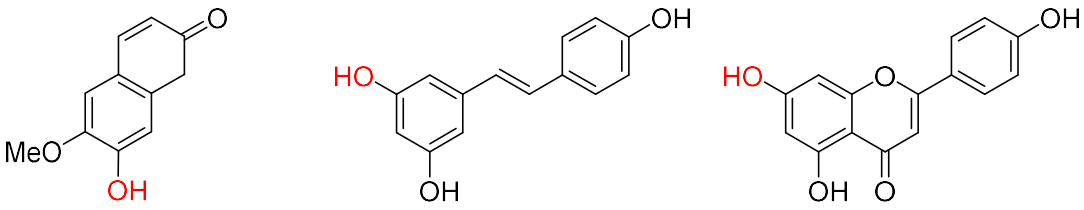
Differential scanning fluorimetry (DSF). Melting temperatures (T_m) of the different UGTs were measured by DSF using the Protein Thermal Shift Dye Kit (ThermoFisher Scientific) and a qPCR QuantStudio5 machine. Dye solution (1000x) and acceptors (resveratrol, scopoletin, apigenin, quercetin, pinoresinol, silibinin, xanthotoxol, genistein and 3,4-dichlorophenol) were diluted in 0.8 equivalents NaOH in H_2O milliQ (*e.g.* 1 mM acceptor in 800 µM NaOH). 10 µL of dye/acceptor solution 2x was mixed with 10 µL of protein samples at 0.8 mg/mL in Buffer 2x (100 mM HEPES pH7) and pipetted in a qPCR 96-wells plate. Final conditions are thus HEPES pH7 50 mM, protein 0.4 mg/mL, acceptor either 0, 400 µM (polyphenols) or 750 µM (3,4-dichlorophenol). The plate was centrifuged 30 seconds at 1000 rpm and transferred to the qPCR machine. The protocol initiates with 2 minutes incubation at 298 K, followed by a temperature increase of 0.05 $K.s^{-1}$ up to 372 K, and a final incubation of 2 minutes at 372 K. Measurements were carried out in triplicate. Raw data was analyzed with Protein Thermal Shift™ Software v1.x.

Results

The 18 GT1 enzymes have 24–40% pairwise identity after multiple sequence alignment via clustal omega (Sievers and Higgins, 2014). Nine of these enzymes have been previously described in the literature: *Pt*UGT1 (Teze *et al.*, 2021), *Zm*UGT708A6 (Ferreyra *et al.*, 2013), *Zm*UGT706F8 (Bidart *et al.*, 2022), the GT1s from *Arabidopsis thaliana* (*At*UGT72E2, *At*71C1, *At*71D1) (Yang *et al.*, 2018), RhGT1 (Wang *et al.*, 2013), Gm88E3 (Liu and Nidetzky, 2021) and *Mt*UGT78G1 (Modolo *et al.*, 2007). Among the 9 GT1 enzymes that were not previously described, five already had designated names (*Zm*71B1, *Os*88C1, *Lc*72B10, *Fi*88A10, *Fe*88J1), and the remaining four are named according to the UGT naming convention (Mackenzie *et al.*, 2005) preceded by two letters referring to genus and

114 species (*e.g. Zm*UGT88C10). Uniprot accession numbers and melting temperatures (T_m) of the 18 GT1
115 enzymes are provided in Table S1.

116 These 18 GT1 enzymes are described here for their activity against three acceptors (Scheme 2)
117 representing different classes of polyphenols of biotechnological interest, *i.e.* flavones (apigenin),
118 stilbenes (*trans*-resveratrol), and coumarins (scopoletin). Interestingly, we found that each of the 18
119 enzymes are active against each of the chosen acceptors, and in most cases (44/54), analytical yields
120 of glycosylation >50% are reached. >90% yield is obtained for at least one condition in about half the
121 enzyme-acceptor pairs (Figures 1 & 2).



scopoletin
coumarin
7-Hydroxy-5-methoxycoumarin

resveratrol
stilbene
trans-3,5,4'-Trihydroxystilbene

apigenin
flavone
4',5,7-trihydroxyflavone

122

123 **Scheme 2. Acceptors assessed in this study.** The position most commonly glucosylated is indicated
124 in red.

125 Note that while only a single glucosylation product is possible for scopoletin, several could be – and
126 are – formed by GT1 enzymes for resveratrol and apigenin. Given that different products or product
127 mixtures are observed for the various enzymatic:substrate pairs, the displayed analytical yields relate
128 to acceptor consumption, and are cross-validated by analyzing the sum of the peak areas observed for
129 products.

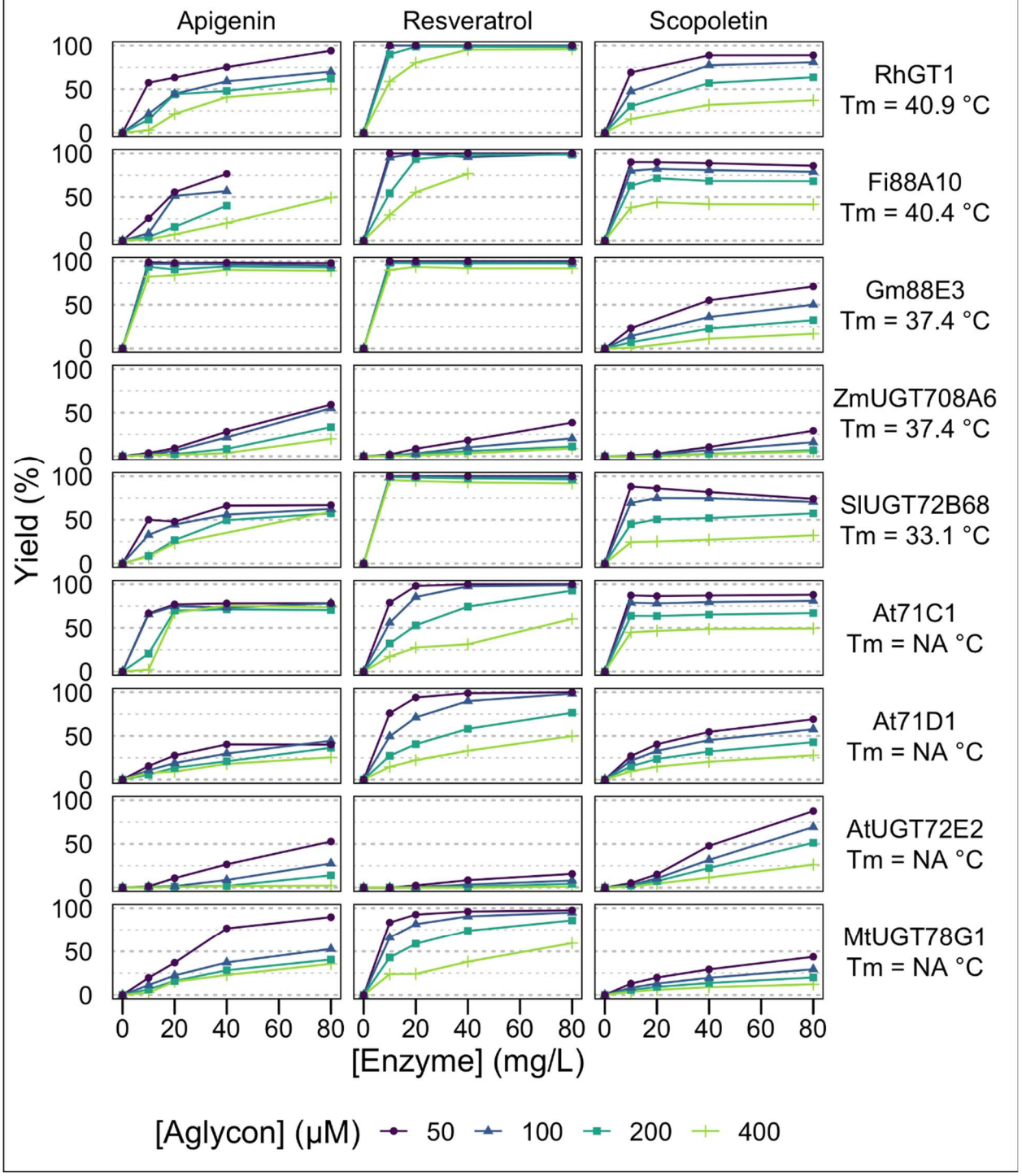


Figure 1: Effect of reaction conditions on glycosylation yields from low-T_m enzymes.
Analytic yields of acceptor conversion are plotted against enzyme concentration, from 10 to 80 mg/L (circa 0.15–1.2 μM). HEPES pH 7, aglycon concentration range 50–400 µM, UDP-Glc 500 µM, 20 h at 293 K, without stirring in 100 µL volume. The 9 GT1s with the lowest T_m or no measured T_m are displayed. NA = Not Available.

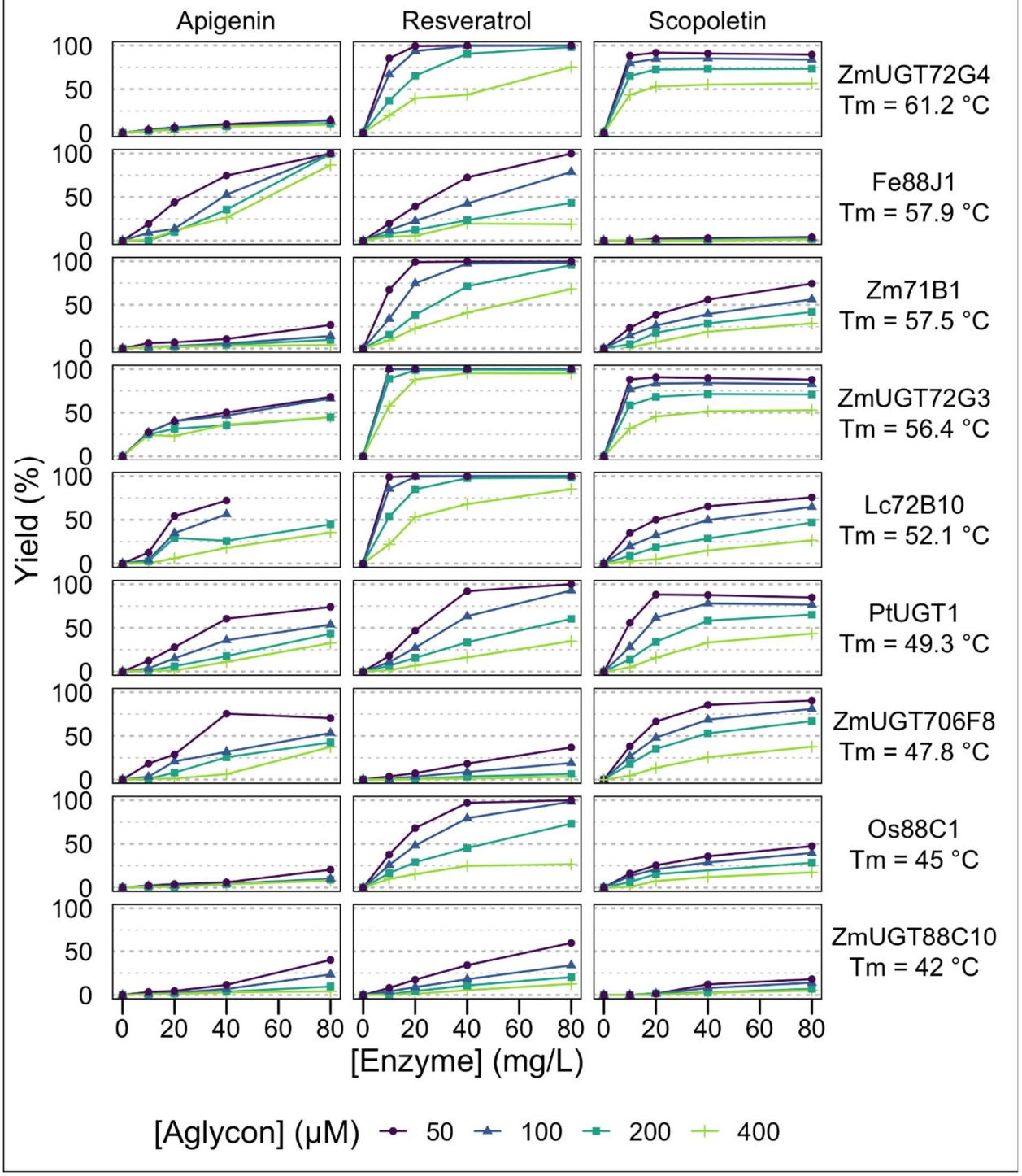


Figure 2: Effect of reaction conditions on glycosylation yields from high-T_m enzymes.
Analytic yields of acceptor conversion are plotted against enzyme concentration, from 10 to 80 mg/L (circa 0.15–1.2 μM). HEPES pH 7, aglycon concentration range 50–400 µM, UDP-Glc 500 µM, 20 h at 293 K, without stirring in 100 µL volume.

130

While most of the curves display the classical dependency on enzyme concentration of a reaction catalyzed by enzymes with low total turnover numbers – *i.e.* a linear or sublinear increase in product as a function of enzyme concentration – half of the enzyme:substrate pairs (27/54) display dilution-induced inactivation behavior with a superlinear dependency on enzyme concentration (Fig. 1 & 2). At low enzyme concentrations (*e.g.* 10 mg/L, circa 150 nM), no-to-little reaction is observed, yet doubling the enzyme concentration far more than double the observed yields. The full dataset of yields as a function of enzyme and acceptor concentrations is available in the supplementary material (Table S2).

Importantly, this behavior is also related to acceptor concentration, being more prevalent at 400 μM than at 50 μM. It is particularly pronounced with apigenin, *e.g.* for *Fe*88J1, *Pt*UGT1, *Zm*UGT88C10, *At*UGT72E2, *At*71C1, *Rh*GT1 and *Zm*UGT708A6 (Figures 1 & 2). It is also observed with resveratrol (*e.g.* *Zm*UGT708A6 or *Zm*UGT706F8) and scopoletin (*e.g.* *Zm*UGT88C10 or *At*UGT72E2). Interestingly, while glycosylation of apigenin and resveratrol regularly (19/36) reaches full conversion of the acceptors, the glucosylation of scopoletin results in an equilibrium (Scheme 1), with a maximum yield depending of the acceptor concentration. At the highest acceptor concentration, nearing donor and acceptor equimolarity (500 and 400 µM, respectively), the maximal yields observed are around 50% (Figures 1 & 2). This allows for the observation of hydrolysis in 5/18 GT1 enzymes in our dataset, being particularly pronounced for SlUGT72B68. Indeed, while the formation of scopoletin-glucoside from UDP-Glc and the formation of UDP-Glc from scopoletin are in equilibrium, the hydrolysis of UDP-Glc by the enzyme is irreversible (Scheme 1). There seems to be a weak correlation between intrinsic stability of the enzyme, represented by its melting temperature (T_m), as the three enzymes seemingly unaffected by the conditions were the relative stable *Zm*UGT72G3 (T_m=56.4±0.1ºC), *Zm*UGT72G4 (T_m=61.2±0.4ºC), and *At*71D1 (ND). Conversely, *Zm*UGT708A6 (T_m=37.4±0.1ºC) and *Zm*UGT88C10 (T_m=42±0.3ºC) were most affected by conditions. Considering ~0.008 kJ/mole/residue

(Rees and Robertson, 2001), and an average length of GT1 enzymes of *c.* 500 residues, a ΔT_m of 1 K roughly equates a stabilization of 1 kcal/mol, thus between the most and least stable enzymes in our dataset a difference as large as 25 kcal/mol is observed. Enzyme-substrate interactions are generally thought to be stabilizing, which is the rationale behind the use of differential scanning fluorimetry as a basis for identifying enzyme-substrate pairs (Niesen, Berglund and Vedadi, 2007). We assessed whether polyphenol acceptors modified the T_m of our proteins, and did not observe a significant change in either direction (Fig. S1). *Zm*UGT708A6, which displays chemostability issues in presence of all three acceptors, would even appear to present slightly higher T_m in presence of resveratrol and apigenin (Fig. S1).

Discussion

In this article, we demonstrate the widespread yet not widely reported phenomena of dilution-induced inactivation and low chemostability towards their own acceptors of GT1 enzymes. These effects are important and can introduce biases in both the kinetic study and discovery efforts for GT1 enzymes. The latter is of particular importance, since one of the major obstacles to a wider biotechnological application of glycosyltransferases is the characterization of their acceptor scope. While one might be enticed to assess acceptors at high concentration to detect catalysts with low affinity (high K_m), or at low enzyme concentrations to be cost-efficient, our results demonstrate that this would result in a significant number of false negatives. While we report the effect, we do not offer a mechanistic explanation. Molecular crowding, occasionally invoked to rationalize dilution-induced inactivation, occurs at much higher concentrations (Miklos *et al.*, 2011; Wang *et al.*, 2012; Cohen and Pielak, 2017). Conversely, the enzyme's adsorption unto equipment (vessel, glassware, tips, etc.) is a concern for trace concentrations or up to the nanomolar range, several orders of magnitude lower than our data and therefore not likely to account for our observations. Further, GT1 enzymes are monomeric, clearly demonstrated by size exclusion chromatography and several crystallographic structures (Wetterhorn *et*

179 *al.*, 2016), ruling out dilution-induced oligomerization disruption as explanation. Here, the synergistic
180 effect with the chemostability at moderately high acceptor concentrations, together with the fact that
181 each enzyme presents various behaviors depending on the acceptor, indicates that specific phenomena
182 related to GT1 enzymes are behind our observations. Conceivably, their relatively large, solvent-
183 exposed hydrophobic acceptor site (Brazier-Hicks *et al.*, 2007; Teze *et al.*, 2021) could be involved.

184

185 **Conflict of Interest**

186 *The authors declare that the research was conducted in the absence of any commercial or financial*
187 *relationships that could be construed as a potential conflict of interest*.

188 **Author Contributions**

189 All authors conceived the study and agreed to the final manuscript. GB and DT performed the
190 enzymatic reactions, which were analyzed by DT. GB performed the T_m measurements. DT drafted
191 the first version of the manuscript, further improved by GB and DW.

192 **Funding**

193 The authors thank the Novo Nordisk Foundation for financial support (grants NNF18OC0034744,
194 NNF10CC1016517, NNF19OC0055620, and NNF20CC0035580).

195 **References**

196 Bidart, G. N. *et al.* (2022) 'Family 1 glycosyltransferase UGT706F8 from Zea mays selectively
197 catalyzes the synthesis of silibinin 7-O-β-D- glucoside', *ACS Sustainable Chemistry & Engineering*,
198 in press(in press), pp. 1–24. Available at: 10.1021/acssuschemeng.1c07593.

199 Brazier-Hicks, M. *et al.* (2007) 'Characterization and engineering of the bifunctional *N*- and *O*-
200 glucosyltransferase involved in xenobiotic metabolism in plants', *Proceedings of the National*
201 *Academy of Sciences*, 104(51), pp. 20238–20243.

202 Chen, D. *et al.* (2015) 'Probing the catalytic promiscuity of a regio- and stereospecific C-
203 glycosyltransferase from Mangifera indica', *Angewandte Chemie - International Edition*, 54(43), pp.
204 12678–12682. doi: 10.1002/anie.201506505.

205 Cohen, R. D. and Pielak, G. J. (2017) 'A cell is more than the sum of its (dilute) parts : A brief
206 history of quinary structure', *protein science*, 26, pp. 403–413. doi: 10.1002/pro.3092.

207 Coutinho, P. M. *et al.* (2003) 'An Evolving Hierarchical Family Classification for
208 Glycosyltransferases', *J. Mol. Biol.*, 2836, pp. 307–317. doi: 10.1016/S0022-2836(03)00307-3.

209 Ferreyra, L. M. F. *et al.* (2013) 'Identification of a Bifunctional Maize *C*- and O-

Glucosyltransferase', *Journal of Biological Chemistry*, 288(44), pp. 31678–31688. doi: 10.1074/jbc.M113.510040.

Fujiwara, R. *et al.* (2009) '*In silico* and *in vitro* Approaches to Elucidate the Thermal Stability of Human UDP-glucuronosyltransferase (UGT) 1A9', *Drug Metab. Pharmacokinet.*, 24(3), pp. 235–244. doi: 10.2133/dmpk.24.235.

Gao, J. *et al.* (2020) 'Investigating the Catalytic Activity of Glycosyltransferase on Quercetin from Tripterygium wilfordii', *ACS Omega*, 5(5), pp. 1414–1421. doi: 10.1021/acsomega.9b02919.

Gutmann, A. and Nidetzky, B. (2013) 'Enzymatic C-glycosylation : Insights from the study of a complementary pair of plant O- and C-glucosyltransferases', *Pure Appl. Chem.*, 85(9), pp. 1865–1877.

Liu, H. and Nidetzky, B. (2021) 'Leloir glycosyltransferases enabled to flow synthesis : Continuous production of the natural *C*-glycoside nothofagin', *Biotechnol Bioeng.*, 118(118), pp. 4402–4413. doi: 10.1002/bit.27908.

Lombard, V. *et al.* (2014) 'The carbohydrate-active enzymes database (CAZy) in 2013.', *Nucleic acids research*, 42(Database issue), pp. D490-5. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3965031&tool=pmcentrez&rendertype=abstract (Accessed: 10 July 2014).

Louveau, T. and Osbourn, A. (2019) 'The Sweet Side of Plant-Specialized Metabolism', *Cold Spring Harb Perspect Biol*, 11, p. a034744.

Mackenzie, P. I. *et al.* (2005) 'Nomenclature update for the mammalian UDP glycosyltransferase (UGT) gene superfamily', *Pharmacogenetics and Genomics*, 15(10), pp. 677–685.

Miklos, A. C. *et al.* (2011) 'Protein Crowding Tunes Protein Stability', *Journal of the American Chemical Society*, 133(133), pp. 7116–7120.

Modolo, L. V. *et al.* (2007) 'A functional genomics approach to (iso) flavonoid glycosylation in the model legume *Medicago truncatula*', *Plant. Mol. Biol.*, 64(64), pp. 499–518. doi: 10.1007/s11103-007-9167-6.

Nidetzky, B., Gutmann, A. and Zhong, C. (2018) 'Leloir Glycosyltransferases as Biocatalysts for Chemical Production', *ACS Catalysis*. American Chemical Society, 8(7), pp. 6283–6300.

Niesen, F. H., Berglund, H. and Vedadi, M. (2007) 'The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability', *Nature protocols*, 2(9), pp. 2212–2221. doi: 10.1038/nprot.2007.321.

Offen, W. *et al.* (2006) 'Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification', *EMBO Journal*, 25(6), pp. 1396–1405. doi: 10.1038/sj.emboj.7600970.

Petermeier, P. *et al.* (2021) 'Exploring the in Vitro Operating Window of Glycosyltransferase PtUGT1 from Polygonum Tinctorium for a Biocatalytic Route to Indigo Dye', *ACS sustainable Chemistry & Engineering*, 9(9), pp. 8497–8506.

Putkaradze, N. *et al.* (2021) 'Natural Product Reports Natural product C-glycosyltransferases – a scarcely characterised enzymatic activity with biotechnological potential', *Natural Product Reports*. Royal Society of Chemistry, 38, pp. 432–443. doi: 10.1039/D0NP00040J.

Rees, D. C. and Robertson, A. D. (2001) 'Some thermodynamic implications for the thermostability of proteins', *protein science*, 10(3), pp. 1187–1194. doi: 10.1110/ps.180101.

Ross, J. *et al.* (2001) 'Higher plant glycosyltransferases.', *Genome biology*, 2(2), p. reviews3004.1-3004.6.

Schmölzer, K. *et al.* (2016) 'Sucrose synthase : A unique glycosyltransferase for biocatalytic glycosylation process development', *Biotechnology Advances*. Elsevier Inc., 34(2), pp. 88–111. doi: 10.1016/j.biotechadv.2015.11.003.

Sievers, F. and Higgins, D. G. (2014) 'Clustal Omega', *Current Protocols in Bioinformatics*, 2014(December), pp. 3.13.1-3.13.16. doi: 10.1002/0471250953.bi0313s48.

Teze, D. *et al.* (2021) '*O*-/*N*-/*S*-Specificity in glycosyltransferase catalysis: from mechanistic understanding to engineering', *ACS Catalysis*, 11(11), pp. 1810–1815.

Vasudevan, U. M. and Lee, E. Y. (2020) 'Flavonoids , terpenoids , and polyketide antibiotics : Role of glycosylation and biocatalytic tactics in engineering glycosylation', *Biotechnology Advances*. Elsevier, 41(41), p. 107550. doi: 10.1016/j.biotechadv.2020.107550.

Wang, L. *et al.* (2013) 'Comparing the acceptor promiscuity of a *Rosa hybrida* glucosyltransferase RhGT1 and an engineered microbial glucosyltransferase OleD^{PSA} toward a small flavonoid library', *Carbohydrate research*, 368, pp. 73–77. doi: 10.1016/j.carres.2012.12.012.

Wang, Y. *et al.* (2012) 'Macromolecular Crowding and Protein Stability', *J. Am. Chem. Soc.*, 134(134), pp. 6–10.

Wetterhorn, K. M. *et al.* (2016) 'Crystal Structure of Os79 (Os04g0206600) from Oryza sativa : A UDP- glucosyltransferase Involved in the Detoxification of Deoxynivalenol', *Biochemistry*, 55(55), pp. 6175–6186. doi: 10.1021/acs.biochem.6b00709.

Yang, M. *et al.* (2018) 'Functional and informatics analysis enables glycosyltransferase activity prediction', *Nature Chemical Biology*. Springer US, 14(12), pp. 1109–1117. Available at: http://dx.doi.org/10.1038/s41589-018-0154-9.

Zhang, L. *et al.* (2022) 'Promiscuity Characteristics of Versatile Plant Glycosyltransferases for Natural Product Glycodiversification', *ACS synthetic biology*, 11(11), pp. 812–819. doi: 10.1021/acssynbio.1c00489.

Acknowledgments

The authors thank Folmer Fredslund for data representation and helpful discussions, and Adrian Frey for helpful discussions. This manuscript is also presented as a preprint on the ChemRxiv server.

Data Availability Statement

283 The dataset generated and analyzed for this study can be found in the Supplementary Materials.
284 Proteins sequences can be retrieved from their Uniprot accession number provided in Supplementary
285 Materials Table TS1.