Alexander Lepak MSc

Elucidation of glycosyltransferase specificities for biocatalysis

DISSERTATION

zur Erlangung des akademischen Grades

Doktor der technischen Wissenschaften

eingereicht an der

Technischen Universität Graz

Betreuer

Univ.-Prof. Dipl.-Ing. Dr. techn. Bernd Nidetzky

Institut für Biotechnologie und Bioprozesstechnik, TU Graz

Graz, März 2019

EIDESSTATTLICHE ERKLÄRUNG

AFFIDAVIT

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZonline hochgeladene Textdokument ist mit der vorliegenden Dissertation identisch.

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present doctoral dissertation.

_____   _____
Datum / Date   Unterschrift / Signature

Acknowledgement

First of all, I would like to thank my supervisor Prof. Dr. Bernd Nidetzky for guiding me through this thesis during the last 5 years. He not only challenged me with new topics, but always encouraged me to develop my sense of critical thinking. Especially his open mind towards new methodology gave me the chance to assemble a broad set of skills.

Special thanks goes to Alexander G. and Alexander D. for their support during my thesis and the stimulating discussions, which helped me to further develop my scientific skills. I am also grateful to Annika Borg, who had to suffer together with me whenever the HPLC broke down, but always kept my mood up. Most of the work would not have been possible without the help of Karin L., Margret S. and Stefan D. who always had a helping hand with experiments.

My mother and my grandfather receive my deepest gratitude for allowing me to pursue on my way through life.

I am grateful to my partner Miriam for cheering me up and always supporting me.

"The first principle is that you must not fool yourself and you are the easiest person to fool. "

<div align="right">Richard P. Feynman</div>

Abstract

The function of many natural products, such as antibiotics, oligosaccharides, polyphenols and proteins is determined by their glycosylation pattern. The site selective glycosylation is in many cases performed by nucleotide diphosphate (NDP) dependent glycosyltransferases (GT) which transfer a sugar moiety from an activated sugar-nucleotide towards a nucleophilic acceptor molecule. However, their industrial application in processes where chemical synthesis lacks selectivity is typically limited by the costs of the nucleotide sugars. To regenerate α-UDP-glucose in-situ from sucrose and UDP, sucrose synthases (SuSy) were explored as recycling system in one-pot reactions. To elucidate the reaction mechanism of inverting GT's, glucosyl fluorides were studied as new glycosyl donors and potential mechanistic probes.

For resveratrol and phloretin, three polyphenol-glucoside preparations from only an acceptor molecule, catalytical amounts of UDP and inexpensive sucrose the improvement in overall reaction performance was shown.The uptake of resveratrol (trans-3,5,4'-trihydroxystilbene) which has drawn great attention as health promoting food ingredient is limited by its vanishingly low water solubility. UGT71A15 from apple was identified as suitable biocatalyst for synthesis of highly soluble resveratrol 3,5-β-D-diglucoside without impairing its antioxidative effectiveness. Phloretin which was identified as a potential treatment for diabetes type 2 is found in form of its 2'-O-glucoside phlorizin in apple. Nothofagin which has a *C*-glycosidic bond instead of the *O*-glycosidic bond found in most of naturally occurring glycosylation products was produced to enhance the stability of the glucoside against hydrolysis.

In a similar approach to "glycosynthases" which are engineered glycoside hydrolases able to synthesize product glycosides from glycosyl fluorides, GT's were tested for their ability to form products from glycosyl fluorides. Out of 6 inverting GTs tested to accept glycosyl fluorides as alternative sugar donors only 3 use β-glucosyl fluoride in presence of catalytic amounts of UDP in place of α-UDP-glucose. In a first rate limiting step α-UDP-glucose is formed in situ and is then used as substrate in the following glucosylation of phloretin. Neither for inverting nor retaining GTs a transfer reaction from α-glucosyl fluorides was observed. In addition, α-mannosyl, galactosyl and glucosyl fluoride were not accepted as sugar donors for the synthesis of the corresponding nucleotide sugars performed by two SuSy's from *Acidophilus bacteriasus* and *Glycine max*. Mutational studies in the active site of a *C*-glycosyltransferase from *Oryza sativa* (OsCGT) revealed that a switch from *O*- to *C*-transferase activity in the I121D mutant increases the formation of α-UDP-glucose from the corresponding fluorid.

Glucosyl fluorides were further used as probes in the study of levoglucosan kinase from *Lipomyces starkeyi* and helped to confirm the predicted reaction mechanism. The improvement of analytical techniques using NMR and whole mass spectroscopy presented strong evidence of an early transition state for the direct transfer of sugar moieties in *O*- and *C*-glycosylation.

List of contents

▪ Peer reviewed publications

Article

"Decoupling of recombinant protein production from Escherichia coli cell growth enhances functional expression of plant Leloir glycosyltransferases"
Martin Lemmerer, Juergen Mairhofer, **Alexander Lepak**, Karin Longus, Rainer Hahn, and Bernd Nidetzky
2019, *Biotechnology and bioengineering*, doi.org/10.1002/bit.26934

Article

"β-Glucosyl Fluoride as Reverse Reaction Donor Substrate and Mechanistic Probe of Inverting Sugar Nucleotide-Dependent Glycosyltransferases"
Alexander Lepak, Alexander Gutmann, and Bernd Nidetzky
2018, *ACS catalysis*, 8(10), 9148-9153, doi.org/10.1021/acscatal.8b02685

Article

"New flavanol O-glycosides in grape and wine"
Marie Zerbib, Jean-Paul Mazauric, Emmanuelle Meudec, Christine Le Guernevé, **Alexander Lepak,** Bernd Nidetzky, Véronique Cheynier, Nancy Terrier, and Cédric Saucier
2018, *Food chemistry*, 266, 441-448, doi.org/10.1016/j.foodchem.2018.06.019

Article

"Biochemical characterization and mechanistic analysis of the levoglucosan kinase from Lipomyces starkeyi"
Christina Rother, Alexander Gutmann, Ramakrishna Gudiminchi, Hansjörg Weber, **Alexander Lepak**, and Bernd Nidetzky
2017, *ChemBioChem,* 19(6), 596-603, doi.org/10.1002/cbic.201700587

Article

"Glycosyltransferase cascades for natural product glycosylation: Use of plant instead of bacterial sucrose synthases improves the UDP-glucose recycling from sucrose and UDP"
Alexander Gutmann, **Alexander Lepak**, Margo Diricks, Tom Desmet, Bernd Nidetzky
2017, *Biotechnology journal*, 12(7), 1600557, doi.org/10.1002/biot.201600557

Article

"Screening of Recombinant Glycosyltransferases Reveals the Broad Acceptor Specificity of Stevia UGT-76G1"
Griet Dewitte, Maarten Walmagh, Margot Diricks, **Alexander Lepak**, Alexander Gutmann, Bernd Nidetzky, Tom Desmet
2016, *Journal of biotechnology*, 233, 49-55, doi.org/10.1016/j.jbiotec.2016.06.034

Article

"Creating a Water-Soluble Resveratrol-Based Antioxidant by Site-Selective Enzymatic Glucosylation"
Alexander Lepak, Alexander Gutmann, Sandra Kulmer, Bernd Nidetzky
2015, *ChemBioChem*, 16(13), 1870-1874, doi.org/10.1002/cbic.201500284

Article

"Whole Molecular Mass Spectroscopy Measurement of Kinetic Isotope Effect in C- and O-glycosyltransferases"
Alexander Lepak, Kenneth B. Jensen, Jürgen Hartler, Alexander Gutmann, Bernd Nidetzky
Manuscript in preparation

▪ Oral and poster presentations

Talk

"Sucrose Synthases Enable nucleotide-Sugar Recycling in Glycosyltransferase Catalyzed Glycosylations"

Alexander Lepak, Alexander Gutmann, Bernd Nidetzky
July 2017 · Eurocarb – Carbohydrate Symposium, Barcelona, Spain

Poster

"Nucleotide-sugar Regeneration in Glycosyltransferase Catalyzed Preparation of Water Soluble Polyphenols"

Alexander Lepak, Alexander Gutmann, Sandra Kulmer, Bernd Nidetzky
April 2017 · Carbohydrate Bioengineering Meeting, Vienna, Austria

Poster

"Site-Selective Enzymatic Glucosylation Creates Water-Soluble Polyphenols"

Alexander Lepak, Alexander Gutmann, Sandra Kulmer, Bernd Nidetzky
November 2016 · European Summit of Industrial Biotechnology, Graz, Austria

Talk

"*Creating a Water-Soluble Resveratrol-Based Antioxidant by Site-Selective Enzymatic Glucosylation"*

Alexander Lepak, Alexander Gutmann, Sandra Kulmer, Bernd Nidetzky
July 2017 · International Conference on Polyphenols, Vienna, Austria

Poster

"Improving the Water Solubility of Polyphenolsthrough Site-Selective Enzymatic Glucosylation"

Alexander Lepak, Alexander Gutmann, Sandra Kulmer, Bernd Nidetzky
June 2016 · EMBO Conference - Biocatalysis, Oulu, Finland

Talk

"Creating a Water-Soluble Resveratrol-Based Antioxidant by Site-Selective Enzymatic Glucosylation"

Alexander Lepak, Alexander Gutmann, Sandra Kulmer, Bernd Nidetzky
June 2016 · Carbohydrate Workshop, Vienna, Austria

Poster

"NMR studies of EBF1 in B-cell development"

Alexander Lepak, Jacob Kuruvilla, Susana Cristobal, Mikael Sigvardsson, Maria Sunnerhagen
June 2013 · Structural Biology Network Annual Conference, Tällberg, Sweden

Chapter 1:

Creating a Water-Soluble Resveratrol-Based Antioxidant by Site-Selective Enzymatic Glucosylation

Creating a Water-Soluble Resveratrol-Based Antioxidant by Site-Selective Enzymatic Glucosylation

Alexander Lepak,[a] Alexander Gutmann,[a] Sandra T. Kulmer,[a] and Bernd Nidetzky*[a, b]

The phytochemical resveratrol (trans-3,5,4′-trihydroxystilbene) has drawn great interest as health-promoting food ingredient and potential therapeutic agent. However, resveratrol shows vanishingly low water solubility; this limits its uptake and complicates the development of effective therapeutic forms. Glycosylation should be useful to enhance resveratrol solubility, with the caveat that unselective attachment of sugars could destroy the molecule's antioxidant activity. UGT71A15 (a uridine 5′-diphosphate α-D-glucose-dependent glucosyltransferase from apple) was used to synthesize resveratrol 3,5-β-D-diglucoside; this was about 1700-fold more water-soluble than the unglucosylated molecule (~0.18 mm), yet retained most of the antioxidant activity. Resveratrol 3-β-D-glucoside, which is the naturally abundant form of resveratrol, was a practical substrate for perfect site-selective conversion into the target diglucoside in quantitative yield (g L^{-1} concentration).

Polyphenols constitute a large and diverse class of plant natural products that are widespread and abundant in fruits, seeds, and vegetables.[1] Interest in natural polyphenols has been high across multiple scientific disciplines, as well as for the public: there are strong claims about the significant health benefits resulting from the regular consumption of polyphenol-containing foods.[2] The stilbene resveratrol (Scheme 1) has received keen attention in particular, because in addition to its cardioprotective[3] and anti-inflammatory[4] effects (long suspected in red wine), resveratrol has also shown promising activities against different forms of cancer.[3b, 4a, 5] It has proved effective against neurodegenerative diseases,[6] in extending the lifespan of model organisms[7] and in attenuating age-related changes in mice.[8] Its positive effects on the skin have further encouraged interesting applications in cosmetics.[9]

A limitation in translating the beneficial effects of resveratrol to humans apparently lies in the molecule's low bioavailability.[10] Rapid metabolism by sulfation or glucuronidation is recognized (although not universally)[11] as major problem for resveratrol plasma stability and target-tissue delivery.[10b, 12] The vanishingly low water solubility of resveratrol is another key problem, mainly for uptake from foods or in pharmaceutical form.[13] Besides classical formulation development,[10b, 13] chemical conjugations, such as sulfation,[14] glycosylation,[15] oligomerization,[16] and other derivations,[15a, 17] hold particular promise, as they offer the possibility of altering the uptake characteristics.[17a,c] Glycosylation is an optimal biocompatible approach for water-solubility enhancement; it is also a proven method in prodrug development.[18] Notably, in an animal model, resveratrol 3,5-β-D-diglucoside was shown to be a potential prodrug for the treatment of inflammatory bowel disease.[17a] Mechanistic data support the notion that biological effects of resveratrol potentially involve the activation of the protein deacetylase SIRT1,[19] and also its action as an antioxidant.[4b] A concern for the attachment of sugars, therefore, is that the antioxidant activity of resveratrol might be impaired. A systematic evaluation of resveratrol glycosides is lacking, however, one can infer from literature that an unmodified 4′-OH group is important for various pharmacological properties of resveratrol.[14b, 15a, 20] This implies that whatever method is chosen for resveratrol glycosylation, site selectivity requires special attention. Selective enzymatic glycosylation of the chemically unprotected substrate would constitute a highly favorable synthetic procedure. No such procedure is currently available.

Because of their high stilbene content, the roots of Japanese knotweed (Polygonum cuspidatum) are the main commercial source of resveratrol and piceid (also known as polydatin and resveratrol 3-O-β-D-glucoside).[21] The piceid concentration exceeds that of resveratrol by five- to tenfold in P. cuspidatum,[21] and therefore piceid presents an evident point of reference for the development of new resveratrol glycosides.[17a, 22] Like resveratrol, piceid is biologically active.[17a, 22d] Natural glycosylation of polyphenolic substrates is accomplished by nucleoside diphosphate sugar-dependent GTs

Scheme 1. Glycosylation of phenolic compounds, such as resveratrol, by glycosyltransferases (GTs, EC 2.4) typically involves transfer of glucosyl moieties from UDP-glucose and formation of O-β-D-glucosides.

[a] A. Lepak, Dr. A. Gutmann, S. T. Kulmer, Prof. Dr. B. Nidetzky
Institute of Biotechnology and Biochemical Engineering
Graz University of Technology, NAWI Graz
Petersgasse 12, 8010 Graz (Austria)
E-mail: bernd.nidetzky@tugraz.at

[b] Prof. Dr. B. Nidetzky
Austrian Centre of Industrial Biotechnology
Petersgasse 14, 8010 Graz (Austria)

Supporting information for this article is available on the WWW under
http://dx.doi.org/10.1002/cbic.201500284.

2

(Scheme 1), thus strongly suggesting these enzymes for the synthesis of well-defined glycosidic products. With resveratrol, unfortunately, none of various GTs examined showed site selectivity to a degree suitable for synthetic applications.[15b, 23] The naturally available piceid was often the main product of resveratrol glucosylation, thus rendering enzymatic glucosylation of limited use.

Here we report a highly efficient single-step synthesis of a unique resveratrol 3,5-β-D-diglucoside; it exhibited almost 1700-fold increased water solubility compared to resveratrol (100-fold compared to piceid), yet retained most of the antioxidant activity of the unglucosylated molecule. Based on a detailed characterization of its uridine 5′-diphosphate (UDP)-α-D-glucose-dependent reaction with resveratrol, the glycosyltransferase UGT71A15 (GenBank: DQ103712) from apple was identified to catalyze site-selective 5-O-β-D-glucosylation of piceid. We also show that piceid was an efficient acceptor substrate in an overall high-yielding enzymatic transformation with UDP-glucose as donor, thereby giving the target diglucoside in concentrations suitable for the isolation of pure product on a preparative scale ($\geq 5\,\mathrm{g\,L^{-1}}$).

Recombinant UGT71A15, containing N-terminal *Strep*-tag II for single-step affinity purification, was obtained from *Escherichia coli* cultures by well-established protocols for the preparation of (closely) related plant GTs.[24] Although the enzyme tended to form inclusion bodies, a sufficient amount of UGT71A15 was recovered from the soluble fraction of bacterial cell extracts (Figure S1 in the Supporting Information). UGT71A15 was initially selected because of its broad acceptor scope in the glycosylation of flavonoids.[25] We were pleased that resveratrol was a good substrate for UGT71A15: specific activity $\sim 380\,\mathrm{mU\,mg^{-1}}$ in the presence of 20% DMSO, which had to be added in all conversions to avoid precipitation of resveratrol. Glucosylation of resveratrol was analyzed by reversed-phase HPLC adapted for the determination of resveratrol and different glycosides thereof (Supporting Information). Figure 1 A shows a representative UV-HPLC trace from the late stage of the enzymatic reaction. Three products appeared to have been formed from the conversion of resveratrol. Authentic resveratrol glucosides eluted at positions corresponding exactly to the product peaks in the sample (Figure 1 A), thus enabling unambiguous assignment as well as peak area calibration. Only piceid is available as a commercial standard; the other glucosides were isolated by preparative HPLC and identified by NMR analysis as resveratrol 4′-O-β-D-glucoside (resveratroloside) and resveratrol 3,5-di-O-β-D-glucoside.

Figure 1 B shows a full time course of resveratrol glucosylation. The monoglucosides piceid and resveratroloside initially formed in a ratio of about 3:1. Unlike resveratroloside, which reached a maximum concentration upon complete consumption of resveratrol and remained constant afterwards, piceid accumulated in the first phase of reaction (up to 4 h), but then decreased due to its gradual transformation into the 3,5-diglucoside. After extended reaction times, therefore, 3,5-diglucoside was the main product of resveratrol conversion. Interestingly, despite effective glucosylation at 3-OH and 4′-OH of free resveratrol, no 3,4′-diglucoside was synthesized by transforma-
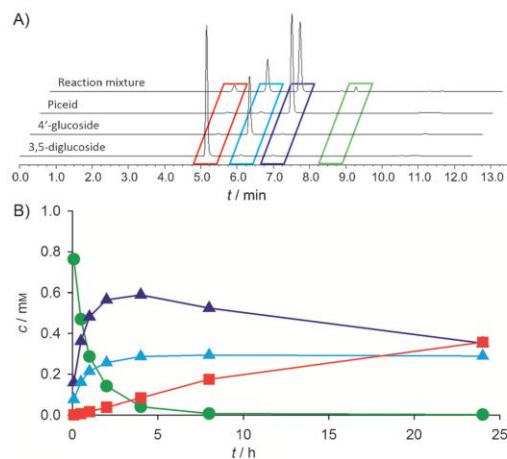
Figure 1. A) HPLC analysis (320 nm) of products formed by resveratrol glucosylation by UGT71A15 and standards (resveratrol glucosides). B) Time course of product formation at pH 7.0 from resveratrol (1 mM) with UGT71A15 (100 μg mL⁻¹) and UDP-glucose (2 mM). Resveratrol (green), piceid (purple), resveratroloside (blue) and 3,5-diglucoside (red).

tion of the intermediaries piceid or resveratroloside (Scheme 2). Resveratrol triglucoside was also not formed. As pH could influence resveratrol glucosylation (by affecting reaction kinetics or thermodynamics),[26] we carried out reactions over pH 6.0–8.0: optimum conversion was around pH 8.0. The pH-dependent distribution of glucosylated products (Figure S2) can be explained by the increased reaction rate at high pH. Piceid prevailed at low pH, at which the 3,5-diglucoside was only a minor by-product. At high pH, the relative amount of 3,5-diglucoside increased, apparently at the expense of piceid formation, because of the accelerated conversion.

Initial-rate analysis with 5 mM resveratrol at pH 8.0 revealed 2.5-fold faster formation of piceid than of resveratroloside (Scheme 2). Resveratrol showed pronounced substrate inhibition, resulting in reduced activities above 2.5 mM (Figure S3). The half-saturation constant for UDP-glucose when glucosylating resveratrol was around 0.25 mM, and no substrate inhibition was detected (Figure S4). Enzyme selectivity for glucosylation of the 3-OH compared with 4′-OH did not change with different concentrations of resveratrol or UDP-glucose.

The results shown in Figure 1 B suggest that UGT71A15 might be a useful catalyst for the conversion of piceid into resveratrol 3,5-diglucoside. This transformation is of interest not only because of the assumed high product selectivity, but also because piceid is more effective and less expensive than resveratrol.[21] The pH dependence of piceid glucosylation by UGT71A15 showed a broad optimum over pH 7.5–8.5 (Figure S5). Kinetic analysis at pH 8.0 revealed that piceid was about sevenfold less efficiently glucosylated than resveratrol (Scheme 2, Figure S6). However, piceid exhibited lower substrate inhibition compared to resveratrol. At 40 mM piceid, UGT71A15 still showed 60% of the maximum activity at about

3
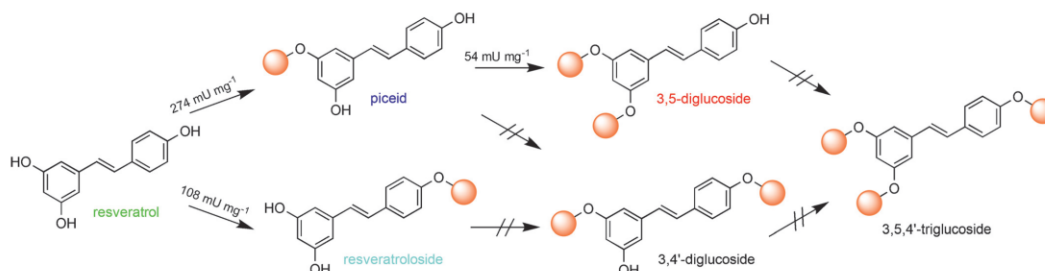
Scheme 2. Three of the five potential glucosides are formed in vitro by β-D-glucosylation of resveratrol with UGT71A15. Arrows indicate transformations confirmed by product identification and activity measurement. Specific activities are based on mass of purified protein. Activities (U = μmol min^{-1}) are initial rates measured at pH 8.0 with 5 mM acceptor and 2 mM UDP-glucose as donor. Orange balls indicate positions of glucosylation.

7.5 mM, whereas resveratrol had maximum activity at 2.5 mM (only 50% at 10 mM). The practical ramification of the elevated K_i values of piceid is that a relatively high concentration can be used without risking strong impairment of enzyme activity. Also, the fact that piceid is somewhat more water-soluble than resveratrol (see below) can be exploited fully in synthetic application.

Figure 2A shows a representative time course of synthesis of resveratrol 3,5-diglucoside from piceid. Using UDP-glucose (20 mM) in excess over piceid (5 mM), near quantitative conversion (96%) of acceptor substrate was achieved, and 3,5-diglucoside was the sole product of the enzymatic glucosylation. With 10 mM UDP-glucose the product yield decreased to around 86% at 23 hours (Figure 2A), likely because of a kinetic rather than a thermodynamic effect as follows. Figure 2A reveals clear slowdown of the reaction at a piceid conversion of about 70% (3.5 mM). The rate-retarding effect, which is explained by the lowered enzyme activity due to drop of the piceid concentration, was more pronounced at the lower UDP-glucose concentration (Figure 2A). When offering resveratrol 3,5-diglucoside and UDP (1 mM each) to UGT71A15 at pH 8.0, no reverse reaction occurred (within the limits of detection of UDP-glucose and piceid, ~0.5% conversion, data not shown). Therefore, this supports the contention that product yield in resveratrol 3,5-diglucoside synthesis was not limited by reaction thermodynamics. Piceid and resveratroloside were also not deglucosylated by the enzyme in the presence of 1 mM UDP at pH 8.0.

We considered that the required fourfold excess of the relatively expensive UDP-glucose might prohibit synthesis of larger amounts of the diglucoside. An enzymatic cascade reaction (Scheme S1) was therefore implemented where the piceid glucosylation was coupled to in situ UDP-glucose production from sucrose and UDP catalyzed by sucrose synthase (SuSy from soybean *Glycine max*, *Gm*SuSy). Sucrose is a more practical glucosyl donor substrate as it is far less expensive than UDP-glucose. Because UDP is recycled in the two-step reaction (Scheme S1), it is used only in sub-stoichiometric amounts. Similar cascades combining the synthetic reaction of a GT with the SuSy reaction were previously shown to offer effective systems for glucosylation of different dihydrochalcone natural
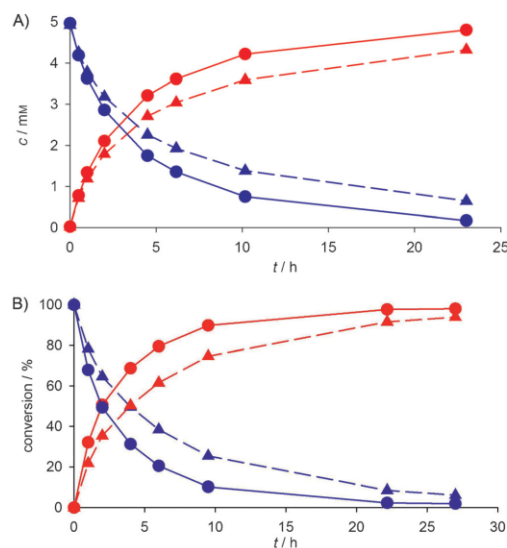
Figure 2. Conversion of piceid (purple) to resveratrol 3,5-diglucoside (red) by UGT71A15. A) 5 mM piceid glucosylated from 10 mM (▲) and 20 mM (●) UDP-glucose, with 470 μg mL^{-1} UGT71A15. B) Glucosylation of 5 (●) and 10 mM (▲) piceid (900 and 1800 μg mL^{-1} UGT71A15, respectively) at pH 7.0 was coupled to UDP-glucose formation from 1 mM UDP and 200 mM sucrose in the presence of *Gm*SuSy (75 μg mL^{-1}).

products.[27] The combination of thermodynamic and kinetic constraints renders UDP-glucose synthesis by *Gm*SuSy unfavorable at high pH.[27a] We were therefore forced to operate the UGT71A15-*Gm*SuSy cascade at pH 7.0, and thus accept a slight decrease in UGT71A15 activity. However, despite replacing 20 mM UDP-glucose with only 1 mM UDP and 200 mM sucrose, conversions of 5 and 10 mM piceid resulted in very high final yields (99 and 97%, respectively; Figure 2B). Although we used relatively high concentrations of UGT71A15 in the reactions (900 and 1800 μg mL^{-1}, for 5 and 10 mM piceid, respectively), the results clearly demonstrate the potential of the

4

two-enzyme cascade for economic synthesis of resveratrol 3,5-diglucoside. Up to $5.4\,g\,L^{-1}$ diglucoside was produced from two inexpensive substrates (piceid and sucrose) with only catalytic amounts of UDP (Figure S7). The kinetic data for UGT71A15 (Figure S6) suggest increasing the piceid concentration (up to ~40 mM), where substrate inhibition of the enzyme becomes effective, as one possible strategy for improving the transformation further. A plausible strategy to cope with the requirement for high amounts of UGT71A15 is the reuse of immobilized enzyme.

About 23 mg of resveratrol 3,5-diglucoside was recovered from conversions of 5 mM piceid on a 10 mL scale by using reversed-phase HPLC, with an isolated yield of 90% based on piceid. Product purity was ~99% (based on HPLC peak area), and product identity was confirmed by comparison with previously reported NMR data.[23h, 28] Resveratroloside was recovered from the more complex product mixture of the resveratrol reaction (Supporting Information).

3,5-Diglucoside was compared with piceid, resveratroloside, and unglucosylated resveratrol (reference) for water solubility and antioxidant activity (Table 1). Attachment of a single glucosyl residue to resveratrol resulted in moderately enhanced solubility, irrespective of the site of glucosylation. Double glucosylation of the resorcinol moiety of resveratrol, by contrast, was extremely effective in increasing the water solubility, by more than 4000-fold in terms of mass concentration. On a molar basis, resveratrol availability in aqueous solution was thus enhanced around 1700-fold or 100-fold when compared to re-

Table 1. Solubility and antioxidative power of resveratrol and its glucosides.

Substrate	Water solubility[a]		DPPH	FRAP
	$g\,L^{-1}$	mM	IC_{50} [μM][b]	FE[c]
resveratrol	0.04	0.18	46.7	1.02
piceid	1.26	3.23	59.3	0.91
resveratroloside	1.42	3.62	>1790	0.16
3,5-diglucoside	172	311	128.5	0.89

[a] After 48 h incubation at 30 °C; [b] Concentrations required for scavenging 50% DPPH after 60 min; [c] Ferrous equivalents (μmol Fe^{2+} produced per μmol of antioxidant after 30 min).

sveratrol or piceid, respectively.

Antioxidant activity was assessed by radical-scavenging and one-electron-reducing abilities, by applying the DPPH (2,2-diphenyl-1-picrylhydrazyl) and FRAP (ferric ion reducing antioxidant power) assays widely used to evaluate the antioxidative power of polyphenols.[29] Both assays gave consistent results, thus showing that glucosylation of the resorcinol moiety of resveratrol caused only a relatively small loss of antioxidant activity compared to the unglucosylated reference resveratrol. Glucosylation of the styryl moiety, by contrast, resulted in a dramatically decreased antioxidant activity, reflected not only in a highly increased IC_{50} value in the DPPH assay and decreased ferrous equivalents (FE) value in the FRAP assay, as shown in Table 1, but also in a greatly decreased reaction rate of resvera-

troloside compared with resveratrol in either antioxidant assay used (Figures S8 and S10). The results clearly underline site selectivity as critical in the enzymatic glucosylation of piceid; this is required for development of a resveratrol derivative exhibiting a desirable combination of high water solubility and similar antioxidant properties.

In conclusion, a highly efficient single-step synthetic route to resveratrol 3,5-diglucoside from piceid is reported here for the first time. Glucosyltransferase UGT71A15 was identified as a perfectly selective catalyst of the reaction. The 3,5-diglucoside was shown to be an interesting water-soluble resveratrol-based antioxidant. It might also constitute a bioavailability-enhanced potential prodrug.[17a] Piceid is more soluble than resveratrol but it is easily hydrolyzed. The 3,5-diglucoside by contrast is more resistant to enzymatic hydrolysis.[17a] 3,5-Diglucoside has been obtained by chemical[17a,b, 28, 30] and enzymatic procedures.[23f, h] However, all these are unselective, thus resulting in complex mixtures of products from which the target diglucoside has to be isolated through elaborate workup protocols. Moreover, chemical synthesis involves a multistep protocol. Clearly, the yield of resveratrol 3,5-diglucoside is seriously limited in synthetic routes and previous biotransformations. The convenient access to the 3,5-diglucoside by simple regio-specific enzymatic production from piceid enables the 3,5-diglucoside to be evaluated more broadly in different applications: its use as a resveratrol prodrug is particularly promising.[17a]

[1] J. M. Harnly, R. F. Doherty, G. R. Beecher, J. M. Holden, D. B. Haytowitz, S. Bhagwat, S. Gebhardt, *J. Agric. Food Chem.* **2006**, *54*, 9966–9977.
[2] a) A. Scalbert, G. Williamson, *J. Nutr.* **2000**, *130*, 2073S–2085S; b) S. Quideau, D. Deffieux, C. Douat-Casassus, L. Pouységu, *Angew. Chem. Int. Ed.* **2011**, *50*, 586–621; *Angew. Chem.* **2011**, *123*, 610–646.
[3] a) P. C.-T. Tang, Y.-F. Ng, S. Ho, M. Gyda, S.-W. Chan, *Pharmacol. Res.* **2014**, *90*, 88–115; b) E.-J. Park, J. M. Pezzuto, *Biochim. Biophys. Acta Mol. Basis Dis.* **2015**, *1852*, 1071–1113.
[4] a) M. Jang, L. Cai, G. O. Udeani, K. V. Slowing, C. F. Thomas, C. W. W. Beecher, H. H. S. Fong, N. R. Farnsworth, A. D. Kinghorn, R. G. Mehta, R. C. Moon, J. M. Pezzuto, *Science* **1997**, *275*, 218–220; b) J. A. Baur, D. A. Sinclair, *Nat. Rev. Drug Discovery* **2006**, *5*, 493–506; c) F. Bollmann, J. Art, J. Henke, K. Schrick, V. Besche, M. Bros, H. Li, D. Siuda, N. Handler, F. Bauer, T. Erker, F. Behnke, B. Mönch, L. Härdle, M. Hoffmann, C.-Y. Chen, U. Förstermann, V. M. Dirsch, O. Werz, H. Kleinert, A. Pautz, *Nucleic Acids Res.* **2014**, *42*, 12555–12569.
[5] B. B. Aggarwal, A. Bhardwaj, R. S. Aggarwal, N. P. Seeram, S. Shishodia, Y. Takada, *Anticancer Res.* **2004**, *24*, 2783–2840.
[6] a) M. Kodali, V. K. Parihar, B. Hattiangady, V. Mishra, B. Shuai, A. K. Shetty, *Sci. Rep.* **2015**, *5*, 8075; b) S. Bastianetto, C. Ménard, R. Quirion, *Biochim. Biophys. Acta* **2015**, *1852*, 1195–1201.
[7] a) K. T. Howitz, K. J. Bitterman, H. Y. Cohen, D. W. Lamming, S. Lavu, J. G. Wood, R. E. Zipkin, P. Chung, A. Kisielewski, L.-L. Zhang, B. Scherer, D. A.

5

Sinclair, *Nature* **2003**, *425*, 191–196; b) J. G. Wood, B. Rogina, S. Lavu, K. Howitz, S. L. Helfand, M. Tatar, D. Sinclair, *Nature* **2004**, *430*, 686–689.

[8] a) J. A. Baur, K. J. Pearson, N. L. Price, H. A. Jamieson, C. Lerin, A. Kalra, V. V. Prabhu, J. S. Allard, G. Lopez-Lluch, K. Lewis, P. J. Pistell, S. Poosala, K. G. Becker, O. Boss, D. Gwinn, M. Wang, S. Ramaswamy, K. W. Fishbein, R. G. Spencer, E. G. Lakatta et al., *Nature* **2006**, *444*, 337–342; b) S. Quideau, D. Deffieux, L. Pouységu, *Angew. Chem. Int. Ed.* **2012**, *51*, 6824–6826; *Angew. Chem.* **2012**, *124*, 6930–6932.

[9] a) R. A. Baxter, *J. Cosmet. Dermatol.* **2008**, *7*, 2–7; b) M. Ndiaye, C. Philippe, H. Mukhtar, N. Ahmad, *Arch. Biochem. Biophys.* **2011**, *508*, 164–170.

[10] a) L. Subramanian, S. Youssef, S. Bhattacharya, J. Kenealey, A. S. Polans, P. R. van Ginkel, *Clin. Cancer Res.* **2010**, *16*, 5942–5948; b) A. Amri, J. C. Chaumeil, S. Sfar, C. Charrueau, *J. Controlled Release* **2012**, *158*, 182–193.

[11] K. R. Patel, C. Andreadi, R. G. Britton, E. Horner-Glister, A. Karmokar, S. Sale, V. A. Brown, D. E. Brenner, R. Singh, W. P. Steward, A. J. Gescher, K. Brown, *Sci. Transl. Med.* **2013**, *5*, 205ra133.

[12] C.-H. Cottart, V. Nivet-Antoine, J.-L. Beaudeux, *Mol. Nutr. Food Res.* **2014**, *58*, 7–21.

[13] N. Summerlin, E. Soo, S. Thakur, Z. Qu, S. Jambhrunkar, A. Popat, *Int. J. Pharm.* **2015**, *479*, 282–290.

[14] a) M. Miksits, K. Wlcek, M. Svoboda, O. Kunert, E. Haslinger, T. Thalhammer, T. Szekeres, W. Jäger, *Planta Med.* **2009**, *75*, 1227–1230; b) J. Hoshino, E.-J. Park, T. P. Kondratyuk, L. Marler, J. M. Pezzuto, R. B. van Breemen, S. Mo, Y. Li, M. Cushman, *J. Med. Chem.* **2010**, *53*, 5033–5043.

[15] a) G. Regev-Shoshani, O. Shoseyov, I. Bilkis, Z. Kerem, *Biochem. J.* **2003**, *374*, 157–163; b) M. Weis, E.-K. Lim, N. Bruce, D. Bowles, *Angew. Chem. Int. Ed.* **2006**, *45*, 3534–3538; *Angew. Chem.* **2006**, *118*, 3614–3618; c) P. Torres, A. Poveda, J. Jimenez-Barbero, J. L. Parra, F. Comelles, A. O. Ballesteros, F. J. Plou, *Adv. Synth. Catal.* **2011**, *353*, 1077–1086; d) K. De Winter, K. Verlinden, V. Křen, L. Weignerová, W. Soetaert, T. Desmet, *Green Chem.* **2013**, *15*, 1949–1955.

[16] a) S. A. Snyder, A. Gollner, M. I. Chiriac, *Nature* **2011**, *474*, 461–466; b) B. S. Matsuura, M. H. Keylor, B. Li, Y. Lin, S. Allison, D. A. Pratt, C. R. J. Stephenson, *Angew. Chem. Int. Ed.* **2015**, *54*, 3754–3757; *Angew. Chem.* **2015**, *127*, 3825–3828.

[17] a) M. Larrosa, J. Tomé-Carneiro, M. J. Yáñez-Gascón, D. Alcántara, M. V. Selma, D. Beltrán, M. T. García-Conesa, R. Lucas, F. Tomás-Barberán, J. C. Morales, J. C. Espín, *J. Med. Chem.* **2010**, *53*, 7365–7376; b) I. Medina, D. Alcántara, M. J. González, P. Torres, R. Lucas, J. Roque, F. J. Plou, J. C. Morales, *J. Agric. Food Chem.* **2010**, *58*, 9778–9786; c) A. Mattarei, M. Azzolini, M. Carraro, N. Sassi, M. Zoratti, C. Paradisi, L. Biasutto, *Mol. Pharmaceutics* **2013**, *10*, 2781–2792.

[18] a) V. Křen, L. Martínková, *Curr. Med. Chem.* **2001**, *8*, 1303–1328; b) D. Bowles, J. Isayenkova, E.-K. Lim, B. Poppenberger, *Curr. Opin. Plant Biol.* **2005**, *8*, 254–263.

[19] a) N. L. Price, A. P. Gomes, A. J. Ling, F. V. Duarte, A. Martin-Montalvo, B. J. North, B. Agarwal, L. Ye, G. Ramadori, J. S. Teodoro, B. P. Hubbard, A. T. Varela, J. G. Davis, B. Varamini, A. Hafner, R. Moaddel, A. P. Rolo, R. Coppari, C. M. Palmeira, R. de Cabo et al., *Cell Metab.* **2012**, *15*, 675–690.

[20] a) A. Matsuoka, K. Takeshita, A. Furuta, M. Ozaki, K. Fukuhara, N. Miyata, *Mutat. Res. Genet. Toxicol. Environ. Mutagen.* **2002**, *521*, 29–35; b) R. Hosoda, A. Kuno, Y. S. Hori, K. Ohtani, N. Wakamiya, A. Oohiro, H. Hamada, Y. Horio, *J. Pharmacol. Exp. Ther.* **2013**, *344*, 124–132; c) Y. Takizawa, R. Nakata, K. Fukuhara, H. Yamashita, H. Kubodera, H. Inoue, *PLoS One* **2015**, *10*, e0120865.

[21] a) S. Mantegna, A. Binello, L. Boffa, M. Giorgis, C. Cena, G. Cravotto, *Food Chem.* **2012**, *130*, 746–750; b) M. Chen, D. Li, Z. Gao, C. Zhang, *Bioprocess Biosyst. Eng.* **2014**, *37*, 1411–1416.

[22] a) H. Zhang, C.-H. Yu, Y.-P. Jiang, C. Peng, K. He, J.-Y. Tang, H.-L. Xin, *PLoS One* **2012**, *7*, e46574; b) H. Park, J. Kim, J.-H. Park, N.-I. Baek, C.-S. Park, H.-S. Lee, J. Cha, *J. Microbiol. Biotechnol.* **2012**, *22*, 1698–1704; c) G. Ravagnan, A. De Filippis, M. Carteni, S. De Maria, V. Cozza, M. Petrazzuolo, M. A. Tufano, G. Donnarumma, *Inflammation* **2013**, *36*, 26–34; d) D. Su, Y. Cheng, M. Liu, D. Liu, H. Cui, B. Zhang, S. Zhou, T. Yang, Q. Mei, *PLoS One* **2013**, *8*, e54505; e) S. De Maria, I. Scognamiglio, A. Lombardi, N. Amodio, M. Caraglia, M. Carteni, G. Ravagnan, P. Stiuso, *J. Transl. Med.* **2013**, *11*, 264.

[23] a) D. Hall, V. De Luca, *Plant J.* **2007**, *49*, 579–591; b) O. F. Iwuchukwu, S. Nagar, *Drug Metabol. Dispos.* **2008**, *36*, 322–330; c) E. H. Hansen, S. A. Osmani, C. Kristensen, B. L. Møller, J. Hansen, *Phytochemistry* **2009**, *70*, 473–482; d) K. Terasaka, Y. Mizutani, A. Nagatsu, H. Mizukami, *FEBS Lett.* **2012**, *586*, 4344–4350; e) S. Ozaki, H. Imai, T. Iwakiri, T. Sato, K. Shimoda, T. Nakayama, H. Hamada, *Biotechnol. Lett.* **2012**, *34*, 475–481; f) M. Zhou, A. Hamza, C.-G. Zhan, J. S. Thorson, *J. Nat. Prod.* **2013**, *76*, 279–286; g) S. Schoffelen, J. Beekwilder, M. F. Debets, D. Bosch, J. C. M. van Hest, *Bioconjugate Chem.* **2013**, *24*, 987–996; h) R. P. Pandey, P. Parajuli, J. Y. Shin, J. Lee, S. Lee, Y.-S. Hong, Y. I. Park, J. S. Kim, J. K. Sohng, *Appl. Environ. Microbiol.* **2014**, *80*, 7235–7243; i) K. Xie, R. Chen, J. Li, R. Wang, D. Chen, X. Dou, J. Dai, *Org. Lett.* **2014**, *16*, 4874–4877.

[24] A. Gutmann, B. Nidetzky, *Angew. Chem. Int. Ed.* **2012**, *51*, 12879–12883; *Angew. Chem.* **2012**, *124*, 13051–13056.

[25] C. Gosch, H. Halbwirth, B. Schneider, D. Hölscher, K. Stich, *Plant Sci.* **2010**, *178*, 299–306.

[26] A. Gutmann, C. Krump, L. Bungaruang, B. Nidetzky, *Chem. Commun.* **2014**, *50*, 5465–5468.

[27] a) L. Bungaruang, A. Gutmann, B. Nidetzky, *Adv. Synth. Catal.* **2013**, *355*, 2757–2763; b) A. Gutmann, L. Bungaruang, H. Weber, M. Leypold, R. Breinbauer, B. Nidetzky, *Green Chem.* **2014**, *16*, 4417–4425.

[28] Z. Zhang, B. Yu, R. R. Schmidt, *Synthesis* **2006**, 1301–1306.

[29] a) D. Huang, B. Ou, R. L. Prior, *J. Agric. Food Chem.* **2005**, *53*, 1841–1856; b) J.-K. Moon, T. Shibamoto, *J. Agric. Food Chem.* **2009**, *57*, 1655–1666; c) G. Clarke, K. N. Ting, C. Wiart, J. Fry, *Antioxidants* **2013**, *2*, 1–10.

[30] V. Jerkovic, F. Nguyen, S. Nizet, S. Collin, *Rapid Commun. Mass Spectrom.* **2007**, *21*, 2456–2466.

6

CHEMBIOCHEM

Supporting Information

Creating a Water-Soluble Resveratrol-Based Antioxidant by Site-Selective Enzymatic Glucosylation

Alexander Lepak,[a] Alexander Gutmann,[a] Sandra T. Kulmer,[a] and Bernd Nidetzky*[a, b]

cbic_201500284_sm_miscellaneous_information.pdf

Contents

1. Methods

1.1. Chemicals and reagents

Unless otherwise indicated, all chemicals were purchased from Sigma-Aldrich in the highest purity available. Resveratrol (99% purity) and piceid (99% purity) were obtained from Carbosynth. *Strep*-Tactin® Sepharose® and desthiobiotin were bought from IBA. BCA assay kit and DNA modifying enzymes were purchased from Thermo Scientific.

1.2. Construction of UGT71A15 and *Gm*SuSy expression strains

The UGT71A15 (GenBank: DQ103712) gene was received from the group of Karl Stich (Institute of Chemical Engineering, Technical University Vienna, Austria) in a pYES 2.1/V5-His-TOPO vector for protein expression in *Saccharomyces cerevisiae*.[1] For expression in *E. coli* the gene was integrated into pET-STRP3, which is a custom made derivate of pET-24d that enables protein expression with an N-terminally fused *Strep*-tag II and was received from the group of Prof. Robert Edwards (Centre for Bioactive Chemistry, Durham University, UK).[2] To clone the UGT71A15 gene into pET-STRP3 the two flanking restriction sites *Nde*I and *Xho*I were introduced by PCR using taaccatatgaagagaccagcacaactagtgttcg and ggtgctcgagttaaattttatcaataaaacgccctaatg as forward and reverse primers, respectively. The resulting fragment was purified on an agarose gel, digested with *Nde*I and *Xho*I fast digest restriction enzymes and cloned into the respective sites of the pET-STRP3 vector. The correct sequence was verified by sequencing the complete gene. The expression strain was created by transformation of electro-competent *E. coli* BL21-Gold (DE3) cells. Construction of the *E. coli* BL21 Gold (DE3) expression strain for *Gm*SuSy (GenBank: AF030231) was described elsewhere.[3]

1.3. Enzyme expression and purification

For expression of UGT71A15 the respective strain was grown at 37°C and 120 rpm in 1 L baffled shake flasks in 300 mL LB-medium containing 50 µg mL^{-1} kanamycin until an optical density at 600 nm of around 0.5 was reached. The temperature was then decreased to 20°C and expression was induced by addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were harvested after 18 h by 30 min centrifugation at 4°C, 5000 rpm, resuspended in ddH$_2$O and frozen at -20°C until cell lysis by sonication (10 s on, 20 s off, 3 min total on-time). The cell lysate was cleared of debris (20,000 x g, 60 min), filtered through a 1.2 µm filter and loaded onto a 3 mL gravity flow *Strep*-Tactin column which was preequilibrated with 3 column volumes (CVs) washing buffer W (100 mM Tris/HCl pH 8.0, 150 mM NaCl, 1 mM EDTA). Purification recommendations given by IBA were closely followed. After washing the column with 5 CVs of buffer W the target protein was eluted with 3 CVs buffer E (100 mM Tris/HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin). The first 0.5 CVs were discarded and the rest was pooled. Subsequently the column was regenerated using 15 CVs of buffer R (100 mM Tris/HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 1 mM hydroxy-azophenyl-benzoic acid) and finally reequilibrated with 10 CVs of buffer W. UGT71A15 was concentrated and buffer exchanged to 25 mM HEPES pH 7.5 using centrifugal concentrators. Expression and purification of *Gm*SuSy was previously reported in detail.[3] Enzyme purity was checked by SDS-polyacrylamide gel electrophoresis (PAGE) and protein concentrations were determined by BCA assay using BSA as standard.

1.4. HPLC-based determination of resveratrol (glucosides) and UDP-(glucose)

Resveratrol and its glucosides were analyzed on a Chromolith® Performance RP-18e, LC 100 x 4.6 mm column at 35°C with a binary gradient using H_2O as mobile phase A and acetonitrile as mobile phase B (0.1% formic acid each). Separation was achieved using following method under UV detection at 320 nm: 5-30% B (7.5 min, 1 mL min^{-1}), 30-100% B (0.01 min, 1 mL min^{-1}), 100% B (2.49 min, 1.5 mL min^{-1}), 100-5% B (0.01 min, 1.5 mL min^{-1}), 5% B (2.49 min, 1.5 mL min^{-1}). Resveratrol and piceid were used as authentic standards and the concentrations of glucosylated products were calculated from peak areas.

To analyze besides resveratrol-glucosides also uridine-derivates (such as UMP, UDP and UDP-glucose) in a single run, we used reversed-phase ion-pairing HPLC with tetra-*n*-butylammonium bromide (TBAB) as ion pairing reagent. HPLC analysis was performed on a reversed phase C18 column (Kinetex^{TM} 5 µm C18 100 Å column 50 x 4.6 mm, Phenomenex) at 35°C. 20 mM potassium phosphate, pH 5.9 containing 40 mM TBAB were used as mobile phase A. It was prepared by dissolving TBAB in a 5-fold diluted 100 mM potassium phosphate buffer (pH 5.9) stock. Acetonitrile was used as mobile phase B. Separation was achieved using following method at a constant flow rate of 2 mL min^{-1}: 5% B (1 min), 5-40% B (3.2 min), 40-90% B (0.01 min), 90% B (0.79 min), 90-5% B (0.01 min), 5% B (0.99 min). Uridine and resveratrol derivatives were monitored at 262 and 320 nm, respectively. UMP, UDP, UDP-glucose, resveratrol and piceid were used as authentic standard and concentrations were calculated from peak areas.

1.5. Determination of initial reaction rates of UGT71A15

Linear initial rates of UGT71A15 conversion were calculated from the amount of resveratrol glucosides formed after various incubation times. The described HPLC assay without ion pairing was used to quantify resveratrol and its glucosides. Reaction mixtures with a typical volume of 200 µL contained besides the respective acceptor and UDP-glucose 13 mM $MgCl_2$, 50 mM KCl, 50 mM HEPES (unless mentioned otherwise), 0.13% (w/v) BSA and 20% (v/v) DMSO. Reactions were started by addition of UGT71A15 at concentrations leading to 5-20% conversion within the observed reaction time of 1h at 30°C without shaking. To stop the reaction samples of typically 20 µL were mixed with an equal volume of acetonitrile and centrifuged for at least 15 min at 13,200 rpm to remove precipitated compounds before HPLC analysis. Typically 4-5 samples were analyzed to determine an initial reaction rate.

The effect of different resveratrol concentrations on UGT71A15 activity was investigated in a range from 10 µM to 10 mM (2 mM UDP-glucose, 12-48 µg mL^{-1} UGT71A15, pH 8.0). In an otherwise identical experiment UDP-glucose was varied between 10 µM and 10 mM (1 mM resveratrol, 20 µg mL^{-1} UGT71A15).

To determine the pH dependency of piceid glucosylation the pH was varied between 5.5 and 10.0 using MES (5.5-6.5), HEPES (6.5-9.0) or CHES (9.0-10.0) as buffer (5 mM piceid, 2 mM UDP-glucose, 100 µg mL^{-1} UGT71A15). Piceid concentrations were varied between 10 µM and 40 mM to evaluate the effect on UGT71A15 activity (2 mM UDP-glucose, 11.6-46.4 µg mL^{-1} UGT71A15, pH 8.0). Furthermore, UDP-glucose concentrations were varied in presence of 5 mM piceid from 10 µM to 10 mM (12 µg mL^{-1} UGT71A15, pH 8.0).

Glucosylation of 1 mM resveratroloside and resveratrol 3,5-diglucoside in presence of 1 mM UDP-glucose and 100 µg mL^{-1} UGT71A15 was tested by incubation at pH 8 for 48 h.

Additionally reversibility of glucosylations was tested by incubation of 1 mM resveratrol glucoside (piceid, resveratroloside or resveratrol 3,5-diglucoside) with 1 mM UDP and 100 µg mL^{-1} UGT71A15 at pH 8.0 for 48 h.

1.6. Resveratrol and piceid conversions

Resveratrol and piceid glucosylations by UGT71A15 were followed by HPLC as described for initial rate measurements to record the respective time courses. Only when *Gm*SuSy was applied for *in situ* regeneration of UDP-glucose the ion pairing HPLC method was used to measure besides the concentrations of resveratrol (glucosides) also those of UDP and UDP-glucose whereas reaction buffers and conditions

were as described for activity measurements. Reaction volumes were adapted to the amount of samples withdrawn in the respective experiment.

Conversion of 1 mM resveratrol with 100 µg mL^{-1} UGT71A15 in presence of 2 mM UDP-glucose was followed for 24 h at pH 7.0. To find the ideal pH for resveratrol glucosylation conversion of 1 mM resveratrol was furthermore followed for 70 h at various pH in the range from 6.0 to 8.0 (1 mM UDP-glucose, 100 µg mL^{-1} UGT71A15, 50 mM HEPES).

Conversion of 5 mM piceid to the 3,5-diglucoside by 470 µg mL^{-1} UGT71A15 was tested with 10 and 20 mM UDP-glucose, respectively (23 h reaction time, pH 8.0). Synthesis of the resveratrol 3,5-diglucoside from 5 and 10 mM piceid with 900 and 1800 µg mL^{-1} UGT71A15, respectively was also coupled to *in situ* UDP-glucose recycling by 75 µg mL^{-1} *Gm*SuSy from 1 mM UDP and 200 mM sucrose.

1.7. Purification of resveratrol glucosides by HPLC

To prepare sufficient amounts of resveratrol glucosides reactions were upscaled. Using 10 mM UDP-glucose either 25 mL of 5 mM resveratrol or 10 mL 5 mM piceid were glucosylated for the synthesis of resveratroloside and resveratrol 3,5-diglucoside, respectively. After removing enzymes with centrifugal concentrators reaction products were separated on a 5 µm Sphereclone ODS(2) (250 x 10.0 mm) column. Water was used as mobile phase A and acetonitrile as mobile phase B (0.1% formic acid each). The flow rate was 3 mL min^{-1} and the percentage of B was increased from 5% by increments of 2.5% until all products were eluted. The purification was done manually and therefore no time for each step is given. After removing acetonitrile under reduced pressure water was removed by freeze drying. 8 mg resveratroloside and 23 mg resveratrol 3,5-diglucoside were recovered.

1.8. Identification of resveratrol glucosides by HPLC and NMR

It was possible to identify resveratrol (glucosides) by comparison of HPLC retention times with authentic standards. For resveratrol and piceid commercial standards were available. Those for resveratroloside and resveratrol 3,5-diglucoside were prepared by ourselves. Besides purified resveratroloside and resveratrol 3,5-diglucoside commercial resveratrol and piceid were analyzed by NMR for unambiguous product identification. ^{1}H, ^{13}C, ^{13}C-HSQC and COSY spectra were recorded on a Varian Unity Inova 500 MHz spectrometer or a Bruker Avance III 300 MHz spectrophotometer and subsequently interpreted in Mnova 9.0.

1.9. Antioxidative Assays

To determine the antioxidative power of resveratrol and its glucosides microplate compatible protocols of DPPH (2,2-diphenyl-1-picrylhydrazyl) and FRAP (ferric ion reducing antioxidant power) assay were adapted from Clarke et al.[5] According to previously described methods for DPPH assays dilutions of the antioxidant, dissolved in either DMSO or methanol, were prepared. 20 µL of the dilutions were then mixed with 160 µL DPPH in methanol (40 µg mL^{-1}) and 20 µL of methanol or DMSO to reach a final volumetric ratio of 1:9 DMSO:methanol. This was necessary as resveratrol is not properly soluble in methanol and the reference α-tocopherol is not soluble in DMSO. The absorbance was measured in 90 s intervals at 517 nm on a platereader. Blanks without antioxidant and α-tocopherol standards were run simultaneously. The scavenging was calculated by following equation:

$$\text{DPPH radical scavenging (\%)} = \frac{A_{blank} - A_{sample}}{A_{blank}} * 100\%$$

The concentration which scavenged 50% of DPPH (IC$_{50}$) was calculated by linear interpolation between the closest points above and below 50% scavenging.

For FRAP measurements 20 µL of resveratrol (glucosides) in DMSO were mixed with 180 µL FRAP-buffer consisting of 300 mM acetate buffer pH 3.6, 10 mM 2,4,6-tripyridyl-s-triazine (TPTZ) in 40 mM HCl and 20 mM FeCl$_3$ x 6 H$_2$0 mixed in a volumetric ratio of 10:1:1. The absorbance was measured in 90 s intervals at 595

nm on a platereader to monitor Fe^{2+} formation by comparison with a serial dilution of $FeSO_4$. DMSO as blank and $FeSO_4$ as control were run simultaneously. The FRAP results, given in Ferrous Equivalents (FE), correspond to the amount of electrons donated per molecule of resveratrol (glucoside) to reduce Fe^{3+} to Fe^{2+}. It is calculated from the slope of a linear regression of the concentration of released Fe^{2+} against that of applied resveratrol (glucoside).

1.10. Solubility measurements

Suspensions of resveratrol, piceid, resveratroloside and resveratrol 3,5-diglucoside were prepared by adding excess amounts of the respective powders in ddH_2O. After incubation for 48 h at 2000 rpm on a thermomixer at 30°C the suspensions were centrifuged at 15,200 rpm at 30°C for 15 min. Afterwards 80% of the aqueous phase between pellet and surface was taken and again centrifuged. This was repeated 3 times until no pellet or solid particles on the surface were visible. From the final aqueous solution samples were taken and the concentration of resveratrol, piceid, resveratroloside and 3,5-diglucoside determined with the described HPLC-assay. All steps were carried out with preheated pipette tips to avoid precipitation.

2. Results

2.1. Protein expression
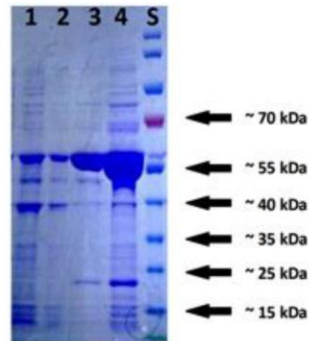
Figure S1: SDS-PAGE of *Strep*-tagged UGT71A15 (53.86 kDa): Lane 1: insoluble fraction after cell disruption; Lane 2-4: UGT71A15 purified by *Strep*-tag affinity chromatography (8, 17 and 50 µg respectively); Lane S: Page Ruler Prestained™ Protein Ladder.

2.2. Resveratrol and piceid conversions
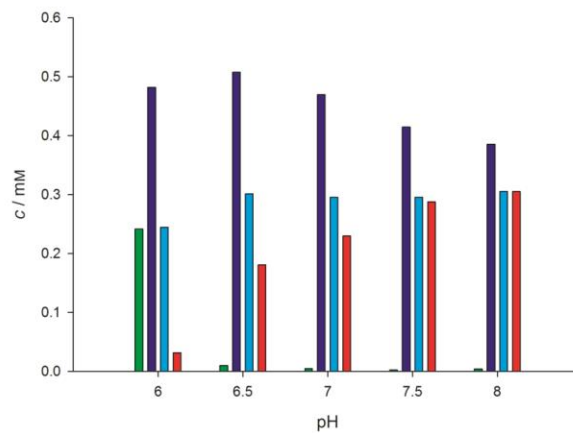
Figure S2: Concentrations of resveratrol and its glucosides after 70 h of glucosylation are pH dependent (1 mM resveratrol, 1 mM UDP-glucose, 100 µg mL^{-1} UGT71A15). Due to increased activity resveratrol was fully converted at high pH values and higher concentrations of the 3,5-diglucoside were obtained. Resveratrol (green), piceid (purple), resveratroloside (blue) and 3,5-diglucoside (red).
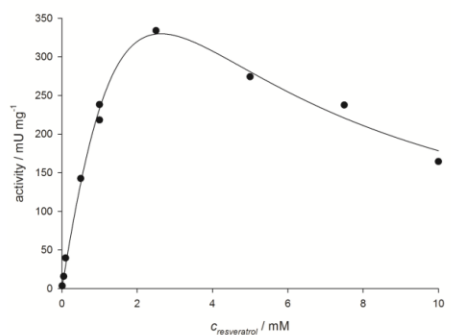
Figure S3: Influence of resveratrol concentration on UGT71A15 activity with 2 mM UDP-glucose at pH 8.0. Resveratrol concentrations higher than 2.5 mM showed a severe inhibitory effect lowering the initial rates at 10 mM to about 50% of maximum activity.
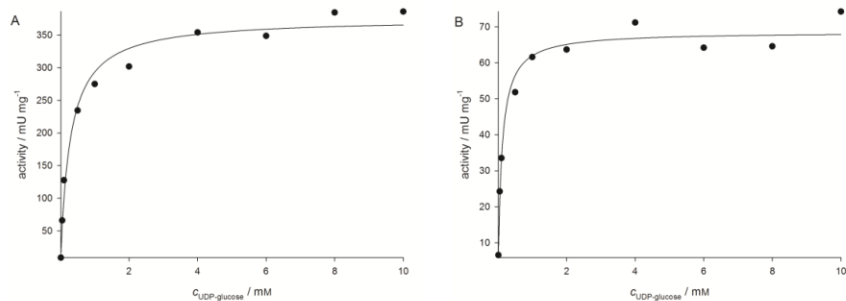
Figure S4: Variation of UDP-glucose concentration in presence of A) 1 mM resveratrol (20 μg mL^{-1} UGT71A15) or B) 5 mM piceid (12 μg mL^{-1} UGT71A15) revealed no substrate inhibition (pH 8). The half-saturation constants for UDP-glucose were determined to be 0.25 mM and 0.1 mM when glycosylating resveratrol or piceid, respectively.
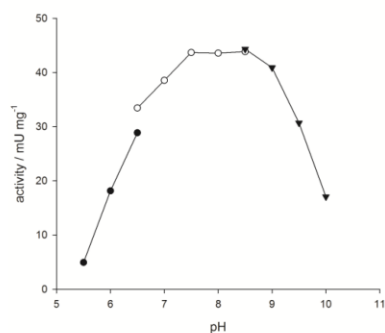
Figure S5: pH profile of piceid (5 mM) glucosylation by UGT71A15 (100 µg mL^{-1}) with 2 mM UDP-glucose. A broad optimum was observed between pH 7.5 and 8.5. MES (filled circle), HEPES (hollow circle) and CHES buffer (triangle) were used in this study and showed no interference with the activity.
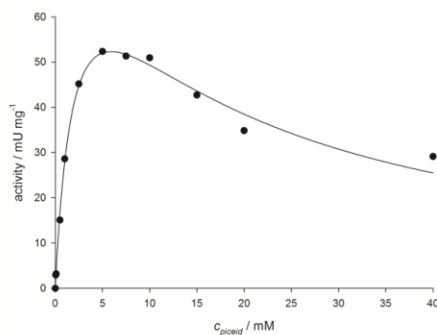
Figure S6: Influence of piceid concentration on UGT71A15 activity with 2 mM UDP-glucose at pH 8.0. Piceid concentrations higher than 7.5 mM showed inhibitory effects that were not as pronounced as with resveratrol as acceptor.
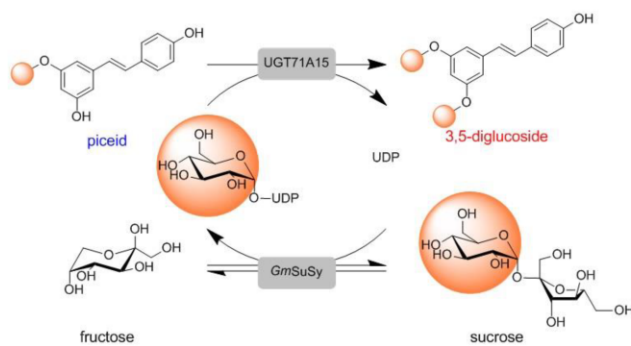
Scheme S1: GT-catalyzed cascade reaction for synthesis of resveratrol 3,5-diglucoside. Glucosylation mediated by UGT71A15 is coupled with *in situ* regeneration of UDP-glucose from UDP and sucrose by *Gm*SuSy.
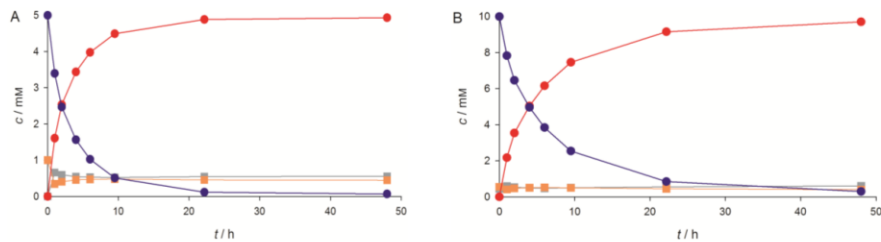
Figure S7: Conversion of piceid to the resveratrol 3,5-diglucoside by UGT71A15 is coupled to *in situ* UDP-glucose formation by *Gm*SuSy (75 µg mL^{-1}) from 1 mM UDP and 200 mM sucrose: A) 5 mM piceid (900 µg mL^{-1} UGT71A15); B) 10 mM piceid (1800 µg mL^{-1} UGT71A15); A constant UDP-glucose concentration of ~0.5 mM is quickly established and maintained throughout the reaction. Although 5 mM piceid (A) are already almost converted after 22 h the UDP-glucose concentration is only slightly increased within the remaining 26 h. This indicates that *Gm*SuSy is not limiting and that the concentrations of UDP and UDP-glucose are mainly determined by the equilibrium of the *Gm*SuSy reaction. Piceid (purple), 3,5-diglucoside (red), UDP (orange) and UDP-glucose (grey).

2.3. Antioxidative assays
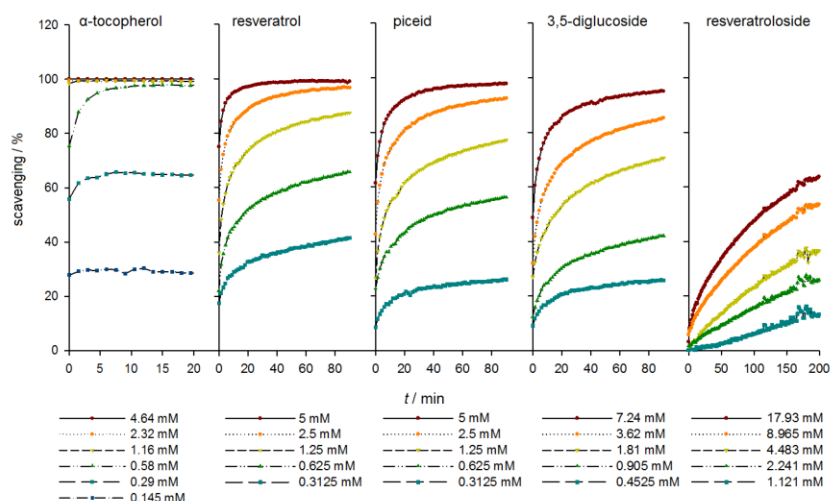
Figure S8: Time courses of DPPH radical scavenging by resveratrol and its glucosides compared to α-tocopherol as positive control (40 µg mL^{-1} DPPH). Despite using higher concentrations of resveratroloside the response time was compared to the other resveratrol derivatives much prolonged.
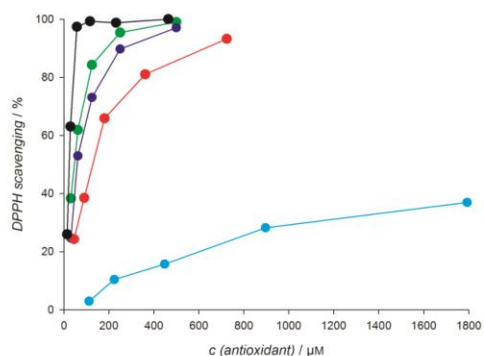
Figure S9: Comparison of concentration dependency of DPPH radical scavenging by resveratrol (glucosides) and α-tocopherol (40 μg mL^{-1} DPPH, after 60 min, data extracted from Figure S8). Resveratroloside shows largely reduced DPPH scavenging. Resveratrol (green), piceid (purple), resveratroloside (blue), 3,5-diglucoside (red), α-tocopherol (black)
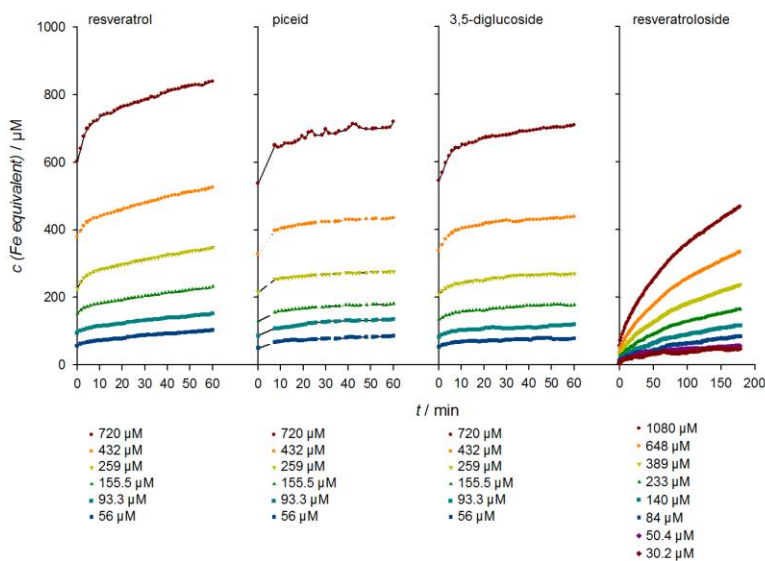
Figure S10: Time resolved production of Fe^{2+} equivalents by resveratrol and its glucosides in FRAP measurements. Although the concentration of resveratroloside was slightly increased the reaction time was drastically prolonged.

2.4. NMR identification of purified resveratrol glucosides

Table S1: ^{1}H and ^{13}C NMR data of resveratrol and its glucosides obtained by preparative HPLC.

resveratrol ($R_{1,3}$=H)[a]			piceid (R_1=glucose, $R_{2,3}$=H)[b]			3,5-diglucoside ($R_{1,2}$=glucose, R_3=H)[c]			resveratroloside ($R_{1,2}$=H, R_3=glucose)[c]		
nr	δC	δH	nr	δC	δH	nr	δC	δH	nr	δC	δH
1	139.76		1	139.62		1	139.78		1	139.44	
2/6	104.25	6.29 (d, J = 2.2 Hz, 2H)	2	105.06	6.55 (1H, unresolved)	2/6	108.03	6.68 (2H, unresolved)	2/6	104.91	6.34 (d, J = 2.2 Hz, 2H)
3/5	156.71		3	158.62		3/5	158.95		3/5	158.97	
4	101.09	6.00 (t, J = 2.2 Hz, 1H)	4	103.06	6.16 (t, J = 2.2 Hz, 1H)	4	103.38	6.39 (1H, unresolved)	4	102.49	6.07 (1H, unresolved)
			5	157.57							
			6	107.47	6.38 (1H, unresolved)						
H-1	125.43	6.78 (d, J = 16.3 Hz, 1H)	H-1	125.51	6.68 (d, J = 16.4 Hz, 1H)	H-1	125.22	6.75 (d, J = 16.4 Hz, 1H)	H-1	127.61	6.85 (1H, unresolved)
H-2	127.85	6.64 (1H, unresolved)	H-2	128.26	6.85 (d, J = 16.4 Hz, 1H)	H-2	128.31	6.94 (d, J = 16.5 Hz, 1H)	H-2	127.81	6.89 (1H, unresolved)
1'	128.86		1'	128.82		1'	129.55		1'	131.26	
2/6'	127.24	7.17 (d, J = 8.5 Hz, 2H)	2/6'	128.20	7.22 (d, J = 8.3 Hz, 2H)	2/6'	128.44	7.22 (d, J = 8.2 Hz, 2H)	2/6'	128.04	7.44 (d, J = 8.3 Hz, 2H)
3/5'	114.92	6.59 (d, J = 8.6 Hz, 2H)	3/5'	115.80	6.58 (d, J = 8.5 Hz, 2H)	3/5'	115.97	6.58 (d, J = 8.1 Hz, 2H)	3/5'	116.84	6.95 (d, J = 8.2 Hz, 2H)
4'	158.02		4'	159.15		4'	157.85		4'	157.43	
			G-1	100.97	4.62 (d, J = 7.2 Hz, 1H)	G-1	100.76	4.85 (2H, unresolved)	G-1	100.73	4.81 (d, J = 7.1 Hz, 1H)

a) ^{1}H: 300.36 MHz; ^{13}C: 75.53 MHz; (CD$_3$OD, δ in ppm)
b) ^{1}H: 300.36 MHz; ^{13}C: 75.53 MHz; (DMSO-d$_6$, δ in ppm)
c) ^{1}H: 300.36 MHz; ^{13}C: 125.70 MHz; (DMSO-d$_6$, δ in ppm)
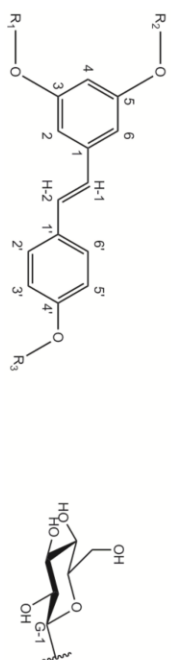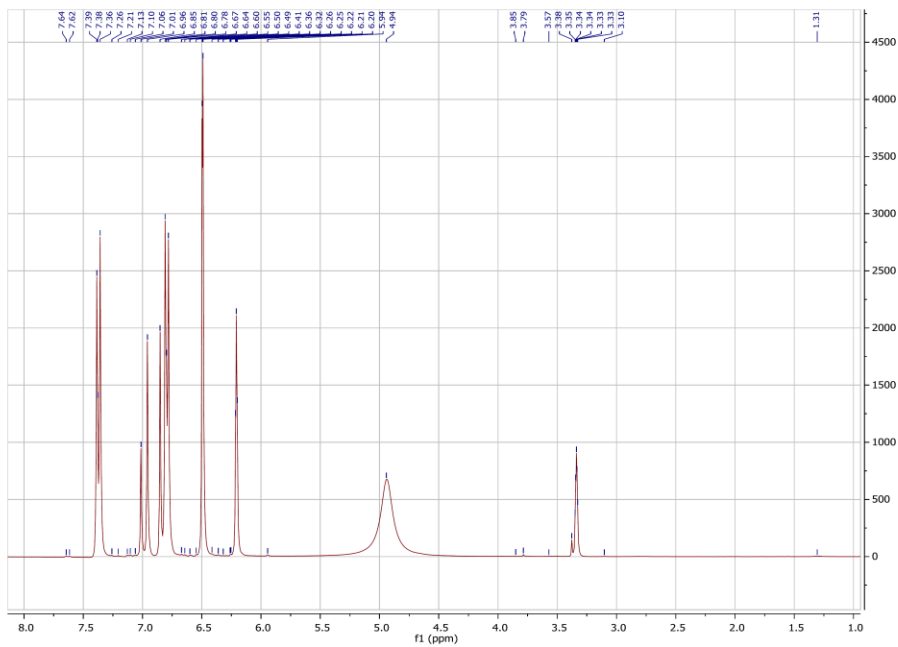
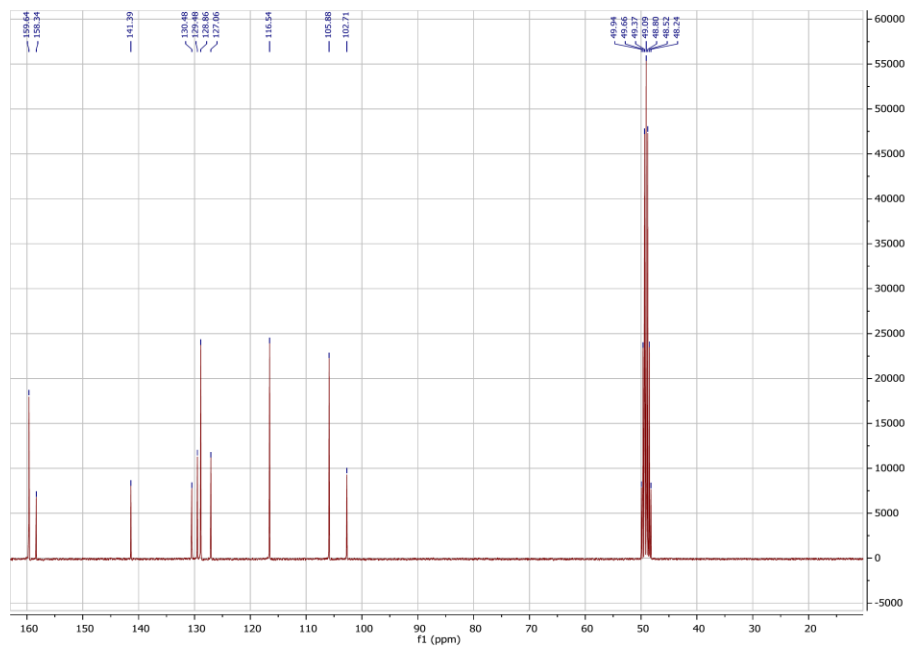Figure S11: ^{1}H-NMR of commercial resveratrol
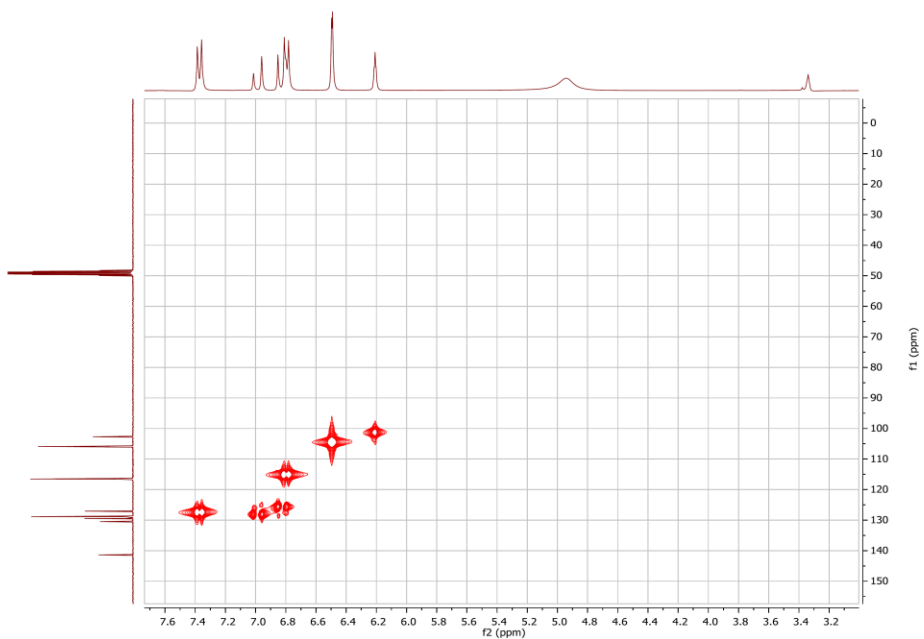

Figure S12: ^{13}C-NMR of commercial resveratrol
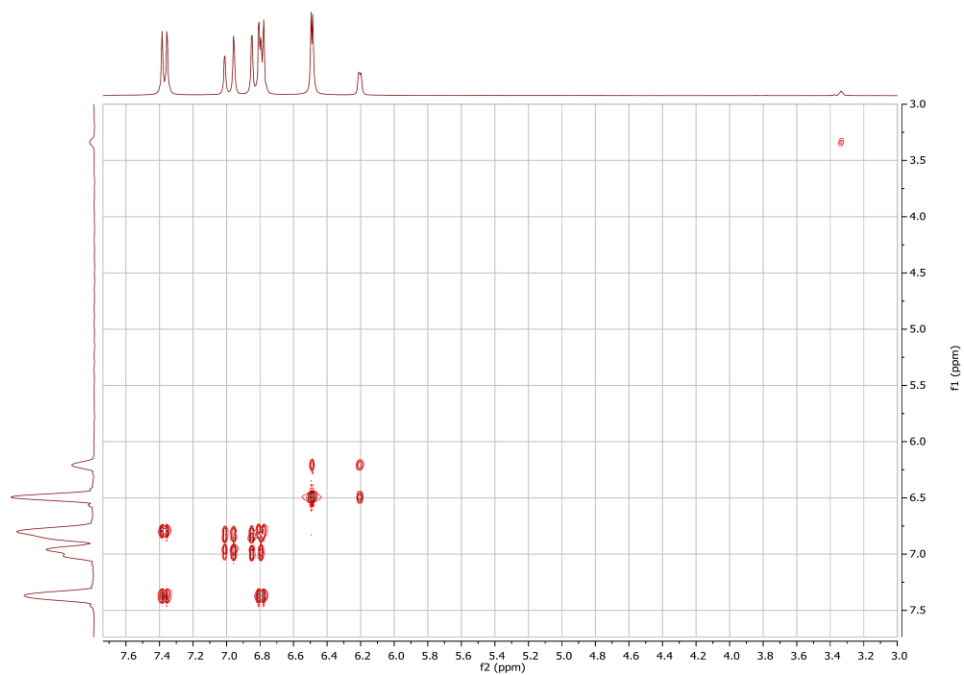
Figure S13: 2D HSQC-NMR of commercial resveratrol

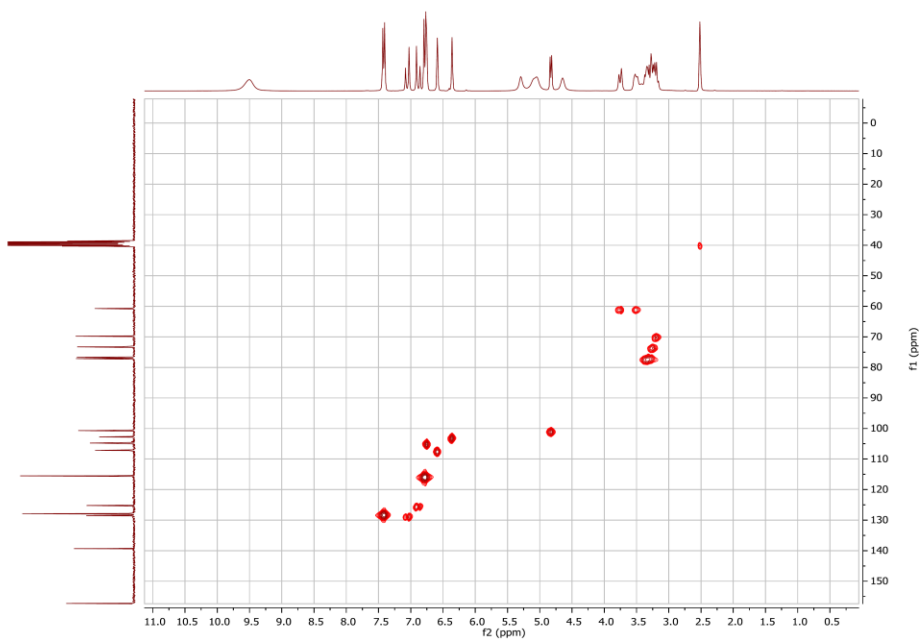Figure S14: 2D COSY-NMR of commercial resveratrol

Figure S15: ¹H-NMR of commercial piceid

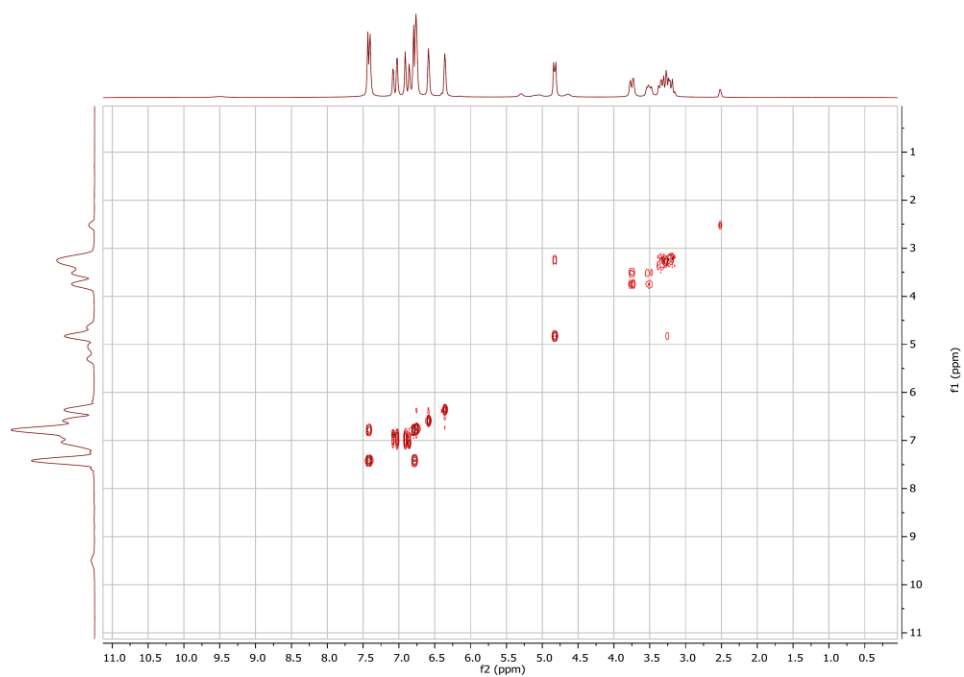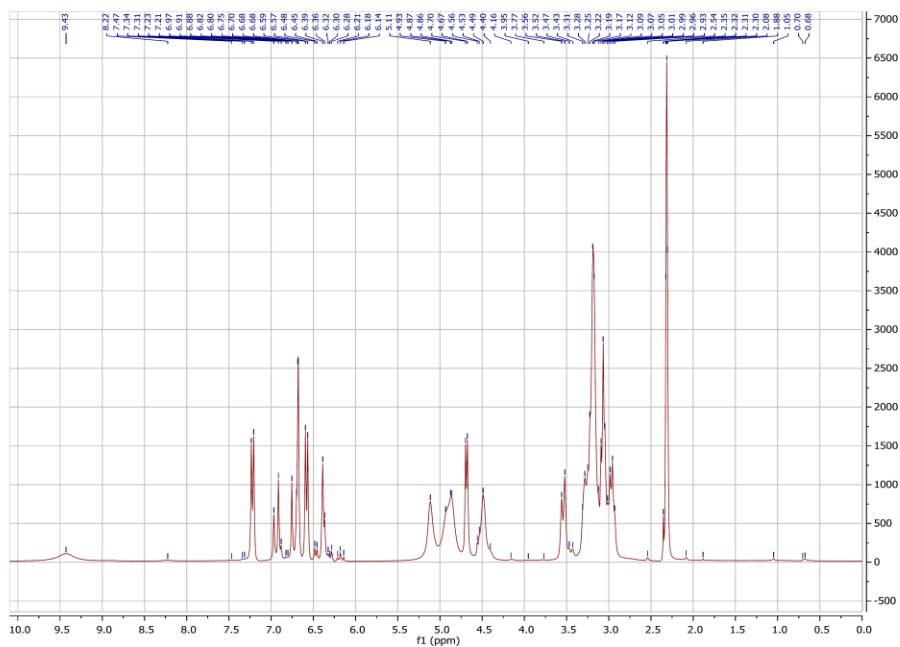Figure S16: ¹³C-NMR of commercial piceid

Figure S17: 2D HSQC-NMR of commercial piceid

Figure S18: 2D COSY-NMR of commercial piceid

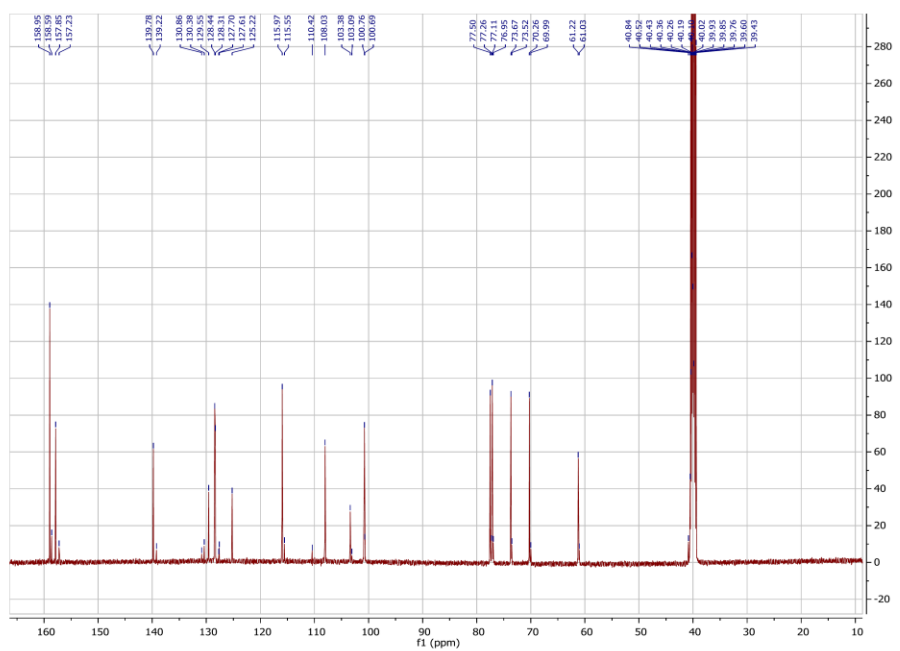Figure S19: ^{1}H-NMR of HPLC purified resveratrol 3,5-diglucoside

Figure S20: ^{13}C-NMR of HPLC purified resveratrol 3,5-diglucoside
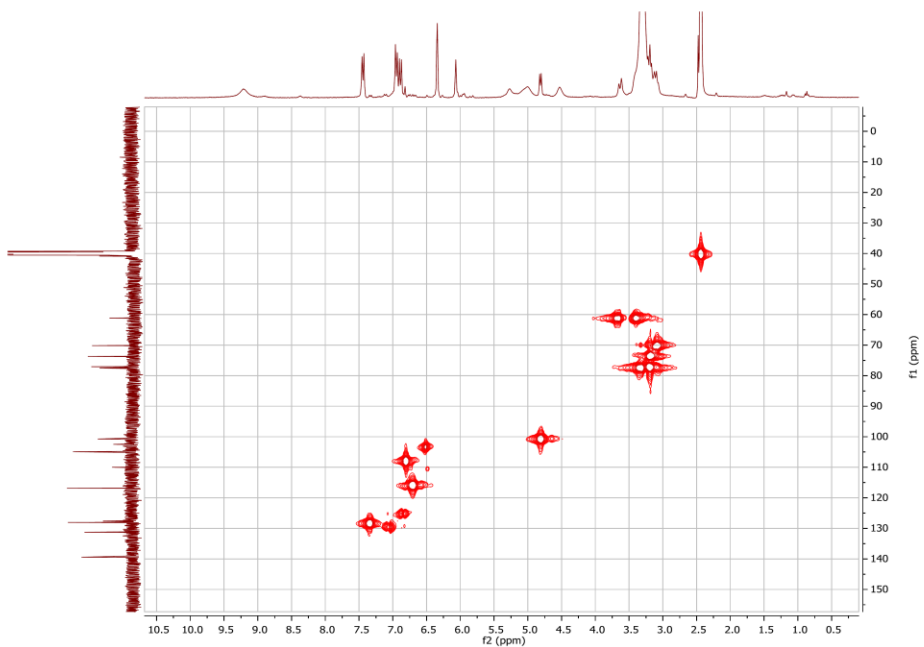
Figure S21: 2D HSQC-NMR of HPLC purified resveratrol 3,5-diglucoside
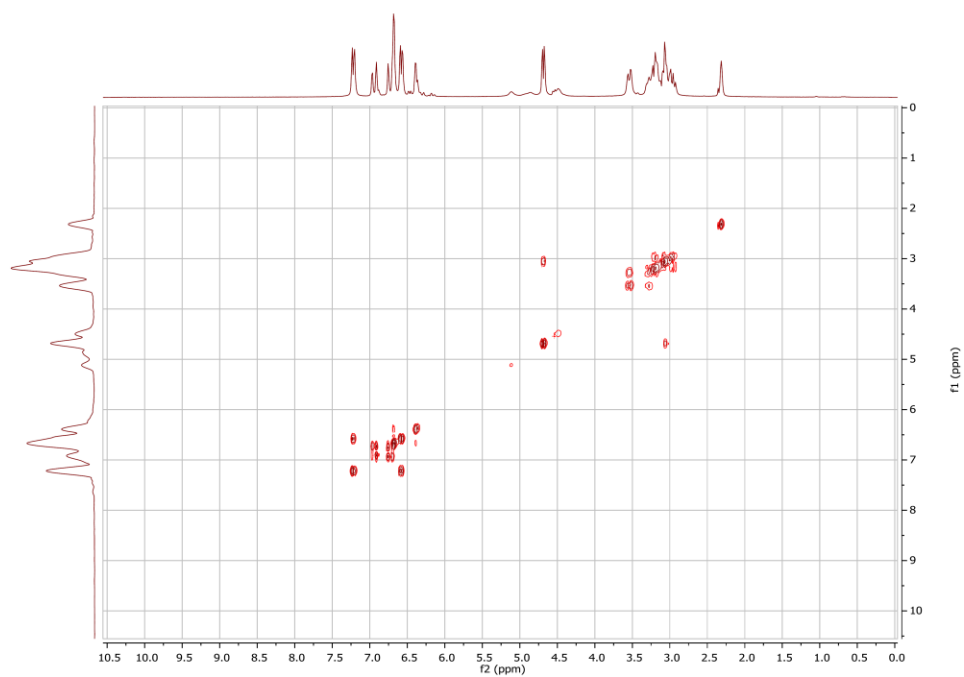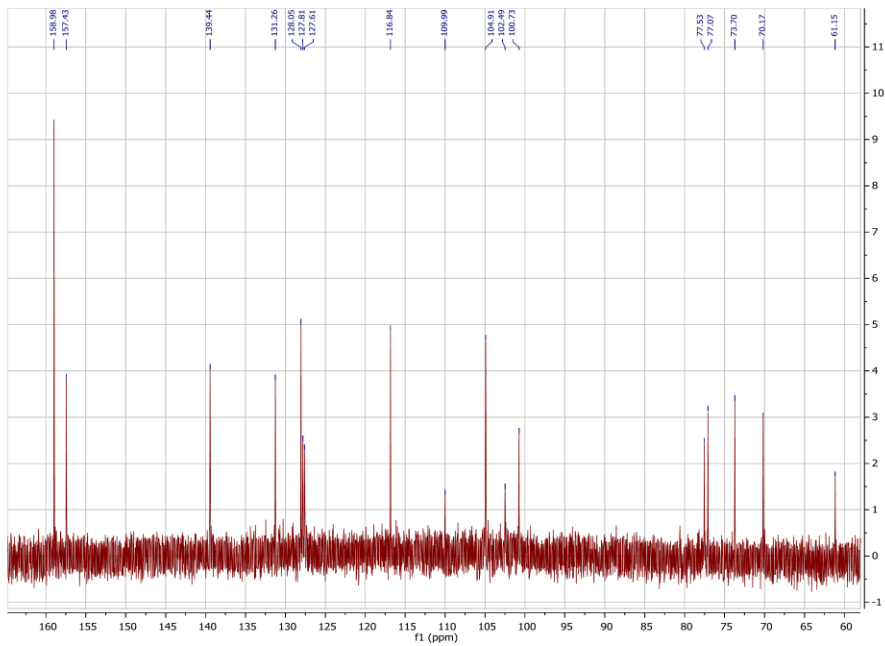

Figure S22: 2D COSY-NMR of HPLC purified resveratrol 3,5-diglucoside
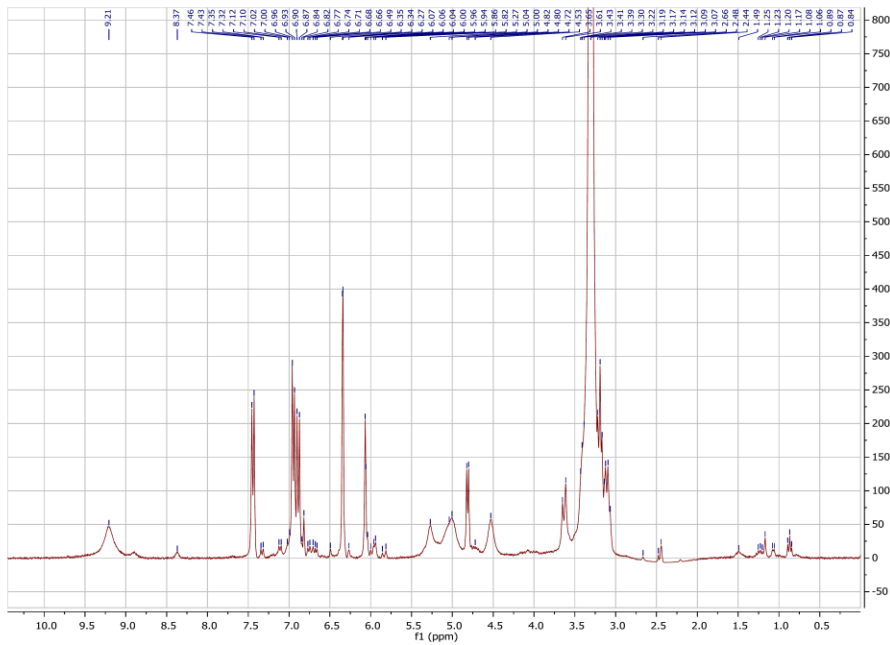
Figure S23: ^{1}H-NMR of HPLC purified resveratroloside

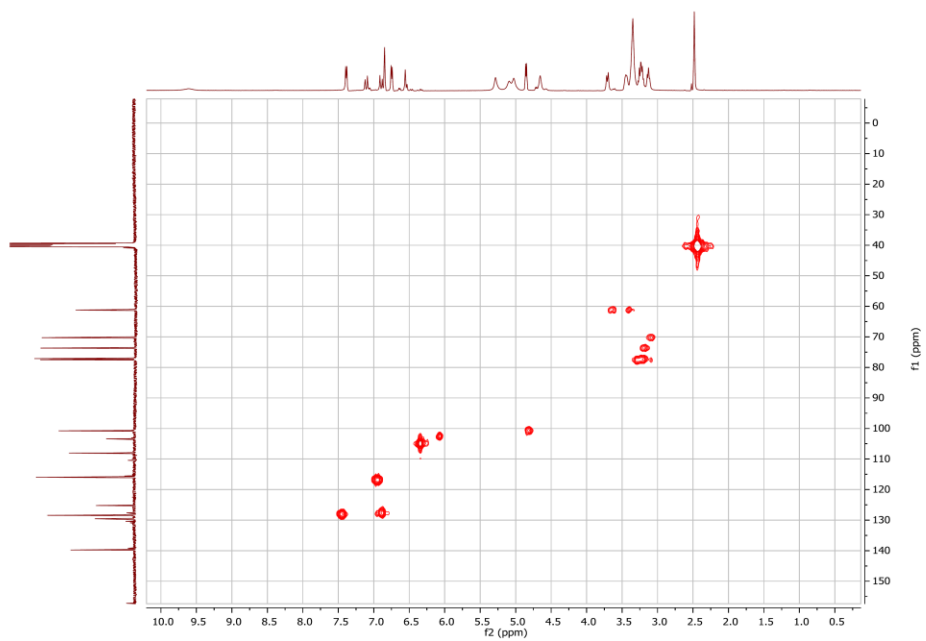
Figure S24: ^{13}C-NMR of HPLC purified resveratroloside

Figure S25: 2D HSQC-NMR of HPLC purified resveratroloside
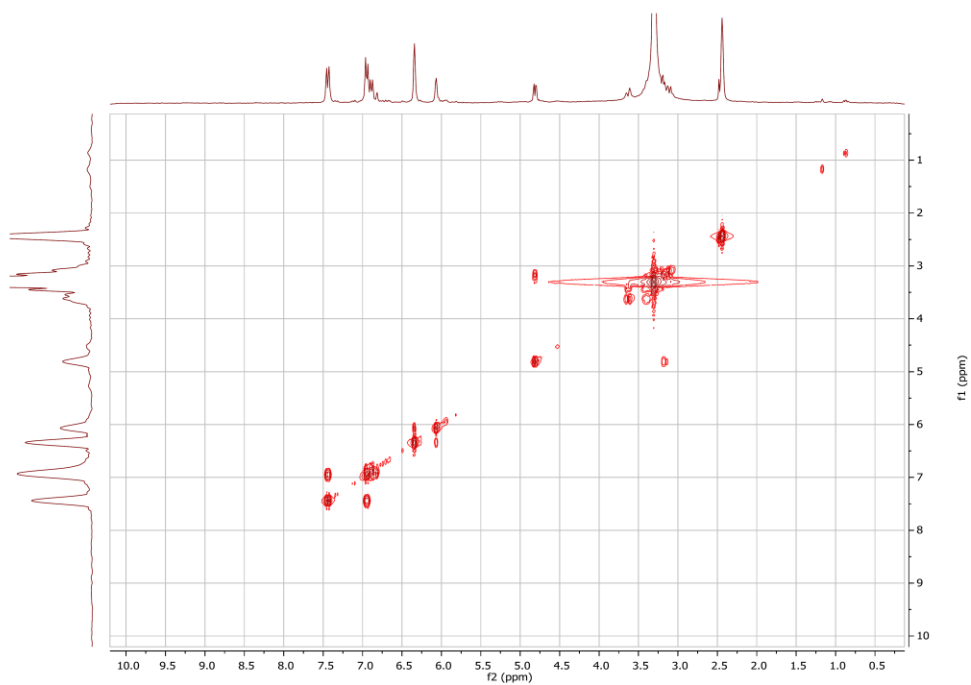

Figure S26: 2D COSY-NMR of HPLC purified resveratroloside

3. References

[1] C. Gosch, H. Halbwirth, B. Schneider, D. Hölscher, K. Stich, *Plant Sci.* **2010**, *178*, 299–306.
[2] D. P. Dixon, A. Lapthorn, P. Madesis, E. A. Mudd, A. Day, R. Edwards, *J. Biol. Chem.* **2008**, *283*, 20268–20276.
[3] L. Bungaruang, A. Gutmann, B. Nidetzky, *Adv. Synth. Catal.* **2013**, *355*, 2757–2763.
[4] A. Gutmann, L. Bungaruang, H. Weber, M. Leypold, R. Breinbauer, B. Nidetzky, *Green Chem.* **2014**, *16*, 4417–4425.
[5] G. Clarke, K. N. Ting, C. Wiart, J. Fry, *Antioxidants* **2013**, *2*, 1–10.

Chapter 2:
Screening of Recombinant Glycosyltransferases Reveals the
Broad Acceptor Specificity of Stevia UGT76G1

Screening of recombinant glycosyltransferases reveals the broad acceptor specificity of stevia UGT-76G1

Griet Dewitte[a], Maarten Walmagh[a], Margo Diricks[a], Alexander Lepak[b], Alexander Gutmann[b], Bernd Nidetzky[b,c], Tom Desmet[a,*]

[a] Centre for Industrial Biotechnology and Biocatalysis, Department of Biochemical and Microbial Technology, Ghent University, Coupure Links 653, B-9000 Ghent, Belgium
[b] Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, NAWI Graz, Petersgasse 12, A-8010 Graz, Austria
[c] Austrian Centre of Industrial Biotechnology, Petersgasse 14, 8010 Graz, Austria

ARTICLE INFO

ABSTRACT

UDP-glycosyltransferases (UGTs) are a promising class of biocatalysts that offer a sustainable alternative for chemical glycosylation of natural products. In this study, we aimed to characterize plant-derived UGTs from the GT-1 family with an emphasis on their acceptor promiscuity and their potential application in glycosylation processes. Recombinant expression in *E. coli* provided sufficient amounts of enzyme for the in-depth characterization of the salicylic acid UGT from *Capsella rubella* (UGT-SACr) and the stevia UGT from *Stevia rebaudiana* (UGT-76G1Sr). The latter was found to have a remarkably broad specificity with activities on a wide diversity of structures, from aliphatic and branched alcohols, over small phenolics to larger flavonoids, terpenoids and even higher glycoside compounds. As an example for its industrial potential, the glycosylation of curcumin was thoroughly evaluated. Under optimized conditions, 96% of curcumin was converted within 24 h into the corresponding curcumin β-glycosides. In addition, the reaction was performed in a coupled system with sucrose synthase from *Glycine max*, to enable the cost-efficient (re)generation of UDP-Glc from sucrose as abundant and renewable resource.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Glycosylation is an efficient tool to increase the solubility of hydrophobic compounds (De Winter et al., 2015; Dirks-Hofmeister et al., 2015), extend the stability of labile compounds (Kwon et al., 2007) or interfere with the release of fragrances and flavors (Schwab et al., 2015). Also, glycosylated compounds are applied in the pharmaceutical industry both as an active moiety for therapeutic activity (Kren and Rezanka, 2008; Oh et al., 2011) and as an innovative approach for targeted drug delivery (Jain et al., 2012). Although many chemical glycosylation methods exist (Muthana et al., 2009; Mydock et al., 2011), the necessary tedious (de)protection and activation steps often lead to low yields and generate a large amount of waste. In that respect, enzymatic reactions offer a promising alternative with a tremendous improvement in eco-efficiency (Desmet et al., 2012). Biocatalyst technology has become appealing in this field as it can easily replace multistep chemical processes. In that respect, glycosyltransferases (GTs) offer a promising alternative as they enable single-step glycosylation at specific positions (Palcic, 2011).

Glycosyltransferases (EC 2.4) are a large family of carbohydrate-active enzymes able to catalyze sugar transfer to an acceptor molecule. The latter could be a small molecule, such as a saccharide, or a macromolecule, such as a polypeptide or a lipid. Based on their preferred donor substrate, GTs can be divided into two groups: the Leloir and the non-Leloir GTs (Lairson et al., 2008). The donor substrates for members of the Leloir group are nucleotide-activated sugars such as UDG-Glc. In contrast, the donor substrates for the non-Leloir GTs are sugar phosphates such as glucose 1-phosphate, or can even be non-activated sugars such as sucrose or starch-derived oligosaccharides. Only about a dozen or so of non-Leloir enzymes have so far been reported while a much larger diversity (>200 EC-entries) is available for the Leloir GTs.

To date, more than 230,000 putative GT sequences have been deposited, but less than 1% of those have been characterized (Schmid et al., 2016). Furthermore, few studies have thoroughly evaluated the acceptor promiscuity of GTs as potential biocatalysts for the glycosylation of a wide range of chemical structures. In that

respect, the most promising enzymes can be found in family GT-1, which contains a large number of plant GTs that use UDP-sugars as donor substrate (so-called UGTs) and use a variety of natural products as acceptor substrate (Cantarel et al., 2009). In this work, a total of five different plant UGTs from GT-1 were recombinantly expressed in *E. coli* to enable their biochemical characterization, with particular emphasis on their acceptor promiscuity and applicability as glycosylating biocatalysts.

2. Materials and methods

2.1. DNA techniques

The genes of the selected GTs were chemically synthesized (GenScript) with a codon usage that is optimal for expression in *E. coli* (Supplementary materials, S1). These sequences were inserted in the constitutive pCXP34h (in-house, Aerts et al., 2011) and inducible pET21a expression vectors (Novagen) via Gibson Assembly (Gibson et al., 2009), using primers listed in Supplementary S2. As a result, a C-terminal His$_6$-tag was inserted in all final constructs. The plasmids were used to transform (electrocompetent) *E. coli* BL21 (DE3) cells (Novagen). Ampicillin (100 μg/mL) was used for proper selection of clones and correct inserts were revealed via sequencing (LCG Genomics).

2.2. Enzyme production and purification

Inoculum was routinely grown at 37 °C in 5 mL LB medium supplemented with 100 μg/mL ampicillin. After overnight incubation with continuous shaking at 200 rpm, 1% (v/v) of inoculum was added to 1 L shake flasks containing 250 mL LB medium, supplemented with ampicillin (100 μg/mL). For constitutive expression, the cells were collected after 6 h (at 37 °C) or 18 h (at 16 °C) of growth by centrifugation in a Sorvall RC-6+ rotator. For inducible expression, after approximately 2 h incubation at 37 °C (when the optical density at 600 nm reached a value of 0.6) expression was started by adding 1 mM IPTG to the inoculum. After 6 h incubation at 37 °C or 18 h incubation at 16 °C, cultures were centrifuged and obtained cell pellets were stored at −20 °C for at least 1 h. Cell pellets from 250 mL cultures were then dissolved in 8 mL lysis buffer (50 mM NaPB pH 7.4 and 500 mM NaCl (PBS), 10 mM imidazole, 100 μM PMSF and 1 mg/mL lysozyme) and exposed to 2 times 3 min of sonication (Branson 250 Sonifier, level 3, 50% duty cycle). Then, cell debris was removed by centrifugation, yielding crude cell extract containing the soluble protein fraction. The His$_6$-tagged proteins were purified out of these cell extracts by Ni-NTA affinity chromatography, using gravity-flow columns as described by the supplier (MC Lab). Finally, buffer was exchanged to 100 mM MOPS pH 7.0 in 30 K Amicon Ultra centrifugal filters (Merck). Production of sucrose synthase from *Glycine max* (*Gm*SuSy) has been described previously (Bungaruang et al., 2013).

2.3. Enzyme characterization

Protein concentration was measured according to the BCA^{TM} Protein Assay kit (Pierce) with bovine serum albumin (BSA) as standard. Enzymatic activity was measured continuously using the colorimetric PK/LDH assay, which is based on NADH detection through the direct coupling of the UDP production with pyruvate kinase (PK) and lactate dehydrogenase (LDH) (Gosselin et al., 1994). The assay mix consisted of 100 times diluted PK/LDH enzymes from rabbit muscle (Sigma-Aldrich), in buffered aqueous glycerol solution, (900–1400 U/mL LDH, 600–1000 U/mL PK), 5 mM MgCl$_2$, 1.6 mg/mL BSA, 0.3 mM NADH and 0.3 mM PEP. Aliquots of assay mix were added to reaction mixtures containing appropriate concentrations of UDP-Glc and acceptor, i.e. stevioside and salicylic

acid for UGT-76G1*Sr* and UGT-SA*Cr*, respectively. In this way, the amount of UDP produced by GT activity was indirectly linked to a drop in absorbance at 340 nm. Temperature profiles were determined using 5 mM UDP-Glc and 5 mM acceptor in 100 mM MOPS buffer at pH 7.0, supplied with 0.6 mg/mL of UGT-76G1*Sr* or UGT-SA*Cr*. A universal Britton-Robinson (BR) buffer system, consisting of 25 mM H$_3$BO$_3$, H$_3$PO$_4$ and CH$_3$COOH was used to determine pH profiles of UGT-76G1*Sr* or UGT-SA*Cr*. Kinetic parameters for UDP-Glc were determined in 100 mM MOPS pH 7.0 at 40 °C, using 5 mM stevioside and 5 mM salicylic acid for UGT-76G1*Sr* and UGT-SA*Cr*, respectively, as co-substrate whereas kinetics for acceptor molecules were determined with 5 mM UDP-Glc. Values were calculated by non-linear regression of the Michaelis-Menten equation using Sigma Plot 11.0. One unit of GT activity was defined as the activity that corresponds to the release of 1 μmol UDP per minute from 5 mM UDP-Glc and 5 mM acceptor in 100 mM MOPS buffer at pH 7.0 and 40 °C. Reaction mixtures for screening the catalytic promiscuity of UGT-76G1*Sr* and UGT-SA*Cr* consisted of 0.5 mM acceptor, 0.5 mM UDP-Glc and 0.25 mg/mL purified enzyme in 100 mM MOPS buffer at pH 7.0.

2.4. Chromatographic analysis

Thin Layer Chromatography (TLC) was applied as a fast method to evaluate the activity of GTs towards different acceptor molecules. To that end, 1 μL samples of reaction mixtures was spotted on TLC silica gel 60 F254 plates (Merck). TLC eluent consisted of EtOAc-MeOH-H$_2$O in a ratio of 30-5-4. Detection was achieved by UV absorption (254 nm) or oxidation with 10% (v/v) H2SO4 solution. HPLC measurements were performed on a reversed phase column (Kinetex^{TM} 5 μm C18 100 å Column 50 × 4.6 mm, Phenomenex) at 35 °C. Mobile phase A contained 20 mM KPO$_4$, pH 5.9, 40 mM TBAB, mobile phase B contained pure acetonitrile. For nucleotide analysis, analysis was performed with 10% B and isocratic flow at a constant flow rate of 2 mL/min for 2 min. Signals were detected at 262 nm. For measurements in which uridine-derivates (such as UMP, UDP and UDP-Glc) and curcumin derivates were analyzed, the time program was altered as followed: 0–1 min 10% B, 1–4.20 min to 40%, 4.20–4.21 min to 90% B, hold until 5 min, 5–5.01 min to 10% B, hold until 6.00 min at a constant flow rate of 2 mL/min. Signals were detected at 262 nm (uridine derivates) and 370 nm (curcumin derivates). UDP, UDP-Glc, curcumin and curcumin derivatives were used as authentic standards, and the product concentration was calculated from peak area.

2.5. Production and purification of curcumin glycosides

As a representative example, the synthesis of curcumin glycosides is described. 2 mg/mL purified UGT-76G1Sr and 0.075 mg/mL GmSuSy was added to reagents UDP (1 mM), sucrose (100 mM) and curcumin (2.5 mM) in a total volume of 25 mL. The solution was buffered in 100 mM MOPS at pH 7.0, supplemented with 10% (v/v) DMSO. Subsequently, the resulting mixture was stirred under reflux for 24 h and temperature was controlled at 30 °C. Proteins were deactivated by a heat treatment of 5 min at 95 °C. After removing precipitated enzyme with centrifugal concentrators, reaction products were separated by preparative HPLC on a 5 μm Spheroclone ODS (250 × 10.0 mm) column. Water was used as mobile phase A and acetonitrile as mobile phase B (0.1% formic acid each). The flow rate was 3 mL/min and the percentage of B was increased from 20% until 90% so that all products were eluted. This preparative HPLC purification was done manually, and therefore no time for each step is given. The solvent acetonitrile was evaporated under reduced pressure and residual water was removed by freeze drying. Preparatory yields were typically more than 50% of the initially applied curcumin (>10 mg product recovery). ^{1}H NMR spectra

Table 1
Overview of the GTs selected in this study.

	EC	Plant source	Uniprot	Protein ID
Monoterpenol β-glucosyltransferase (UGT-MT*Vv*)	2.4.1.127	Vitis vinifera	F6H6Q5	XP_003634093
Gallate 1-O-β-glucosyltransferase (UGT-GA*Qr*)	2.4.1.135	Quercus robus	V5LLZ9	AHA54051
Salicylic acid β-glucosyltransferase (UGT-SA*Cr*)	2.4.1.-	Capsella rubella	R0HCD5	XP_006294206
Stevioside glucosyltransferase (UGT-76G1*Sr*)	2.4.1.-	Stevia rebaudiana	Q6VAB4	AAR06912
Kaempferol 3-O-β-galactosyltransferase (GalT-F3*Ph*)	2.4.1.234	Petunia hybrida	Q9SBQ8	AAD55985

of the isolated products were recorded on a Varian Unity Inova 500 MHz spectrometer and subsequently interpreted in MestReNova.

3. Results and discussion

3.1. Selection of glycosyltransferases

Glycosyltransferases (GTs) that are useful for the glycosylation of natural products are mainly found in family GT-1 (Cantarel et al., 2009; Osmani et al., 2009). Therefore, a representative set of enzymes was selected from this family to cover a wide diversity in acceptor structures, such as phenolics, flavonoids, terpenoids and glycosides (Table 1). Furthermore, a galactosyltransferase (GalT-F3Ph) was included to complement the glucosylation reactions catalyzed by the other enzymes. Indeed, flavonol 3-O-β-galactosyltransferases are able to galactosylate acceptors like kaempferol and quercetin selectively at the 3-position (Miller et al., 1999), with their sugar preference sometimes being determined by a single residue in the C-terminal domain (Sharma et al., 2014).

Monoterpenol β-glucosyltransferases (UGT-MTVv) are known to act on a broad range of substrates, i.e. not only terpenoids such as geraniol and menthol but also aromatic and aliphatic alcohols (Martinkus and Croteau, 1981; Ohgami et al., 2015). In that respect, the isoforms present in grape (Vitis vinifera) seem especially interesting (Bönisch et al., 2014a, 2014b). In turn, salicylic acid β-glucosyltransferases are able to process various hydroxybenzoic acids, forming ether as well as ester linkages (Lim et al., 2002). Here, an enzyme from the pink shepherd's purse (Capsella rubella) was selected for further characterization and acceptor screening (UGT-SACr). Next, gallate 1-O-β-glucosyltransferases (UGT-GAQr) caught our interest because of their activity towards trihydroxybenzoic acid as well as the phenolic aldehyde vanillin (Gross, 1983, 1982). Finally, we also included stevioside glucosyltransferase (UGT-76G1Sr) because of the unusual structure of its acceptor and its obvious commercial interest (Richman et al., 2005).

3.2. Recombinant expression

Glycosyltransferases, especially from plant origin, are generally known for their poor expression yields in bacterial expression systems (Desmet et al., 2012). As full characterization of the selected GTs demands sufficient amount of enzymes, recombinant expression in *E. coli* needed to be optimized. The production yields of these GT enzymes were stepwise improved by evaluating two different promotor strengths (strong T7 vs intermediate P34), inducible versus constitutive expression systems (pET21 vs pCXP34 h) and two strains of *E. coli* as expression host (BL21(DE3) vs Origami2(DE3)). The combination of an inducible expression system with a strong T7 promoter, provided by the pET21a-vector, gave the highest total expression yields. Indeed, in comparison

with the constitutive P34 promoter system, a 3- to 12-fold increase for UGT-SACr and UGT-76G1Sr, respectively, was found (Supplementary S3). For GalT-F3Ph switching the expression host from BL21(DE3) to Origami2(DE3) was helpful, most likely because it enabled the proper formation of disulfide bridges. Unfortunately, for UGT-MTVv and UGT-GAQr, none of these optimization steps resulted in higher yields, and both enzymes were therefore discarded from further characterization. Overall, a generalized approach for efficient plant GT-expression could not be found, as the procedure seems to depend on the GT enzyme under investigation.

3.3. Temperature profile, pH optimum and kinetic properties of UGT-SACr and UGT-76G1Sr

The two enzymes with the highest expression yield, UGT-SACr and UGT-76G1Sr, were selected for further characterization. UGT-76G1Sr, firstly, originates from the plant *Stevia rebaudiana*, a perennial herb of great commercial interest (Ceunen and Geuns, 2013). The enzyme is involved in the biosynthetic pathway for the natural sweetener stevia, where it converts stevioside into rebaudioside A or steviolbioside into rebaudioside B through the addition of a β-1,3-linked glucose moiety (Mohamed et al., 2011; Richman et al., 2005). The second enzyme, UGT-SACr shares 92% sequence identity with UGT-74F2 from *Arabidopsis thaliana* (Lee et al., 1999) but other salicylic acid β-glucosyltransferases have also been reported (Lim et al., 2002). However, none of these enzymes have been thoroughly evaluated for their potential to glycosylate different acceptors.

Both UGT-76G1Sr and UGT-SACr exhibited a temperature optimum of 40 °C. Further, the highest enzyme activities in the direction of glycoside synthesis were found from pH 7.0–8.5 and around 7.0 for UGT-76G1Sr and UGT-SACr, respectively (Fig. 1). Whereas the observed pH optimum is quite typical for GT enzymes (Gachon et al., 2005), the temperature optimum of 40 °C is rather high, especially since the source organisms *Stevia rebaudiana* and *Capsella rubella* have optimal growth temperatures around $20 - 24$ °C (Sicard et al., 2011; Singh and Rao, 2005). In addition, the kinetic stability of UGT-76G1Sr and UGT-SACr was assessed by measuring their half-life time ($t_{1/2}$) at different temperatures. At 37 °C, the $t_{1/2}$ were 35.1 h and 21.4 h, respectively, while at 50 °C those values dropped to 52.7 min and 6.1 min. When stored at 4 °C, $t_{1/2}$ raised to 15.5 days and 21.5 days, respectively.

Apparent kinetic parameters were determined at 40 °C for both enzymes (Table 2). The K_m of UGT-76G1Sr and UGT-SACr for UDP-Glc were 0.45 and 0.66 mM, respectively. K_m values lower than 1 mM (0.24 and 0.83 mM) were also found for their *in vivo* acceptor substrates stevioside and salicylic acid. An important difference between both enzymes was detected when analysing the catalytic efficiencies. Indeed, k_{cat}/K_m values for donor and acceptor substrates were about a factor 10 higher for UGT-76G1Sr when compared to UGT-SACr.
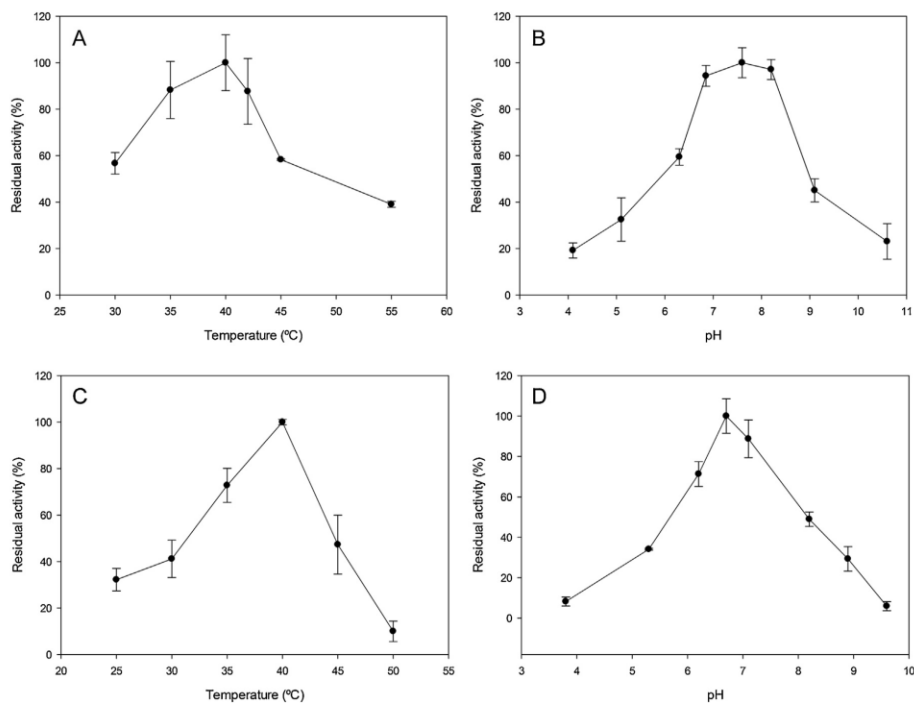
31

Fig. 1. Effect of temperature and pH on the activity of UGT-76G1*Sr* (A–B) and UGT-SA*Cr* (C-D). Reaction conditions: (A) 5 mM stevioside, 5 mM UDP-Glc and 0.6 mg/mL UGT-76G1*Sr* in 100 mM MOPS buffer at pH 7.0 and varying temperature. (B) 5 mM stevioside, 5 mM UDP-Glc and 0.6 mg/mL UGT-76G1*Sr* in 25 mM universal BR buffer at 40 °C and varying pH. (C) 5 mM salicylic acid, 5 mM UDP-Glc and 0.6 mg/mL UGT-SA*Cr* in 100 mM MOPS buffer at pH 7.0 and varying temperature. (D) 5 mM salicylic acid, 5 mM UDP-Glc and 0.6 mg/mL UGT-SA*Cr* in 25 mM universal BR buffer at 40 °C and varying pH.

3.4. Substrate specificity of UGT-SACr and UGT-76G1Sr

Unlike microbial GTs such as Yjic from *Bacillus licheniformis* (Pandey et al., 2014), OleD and variants thereof from *Streptomyces* (Gantt et al., 2008), plant GTs often possess high substrate specificity *in vivo* (Bowles et al., 2005). However, for *in vitro* glycosylation processes, somewhat contradictory findings have been reported. Some authors state that plant GTs are rather promiscuous, being limited mainly by their regioselectivity (Hansen et al., 2003). On the contrary, evidence for highly specific GTs has also been presented (Fukuchi-Mizutani et al., 2003).

Here, the substrate promiscuity of both enzymes was evaluated towards a diverse set of 58 acceptors using TLC analysis as a quick qualitative screening method (Supplementary S5). A total of 78% and 44% of the tested substrates acted as acceptors of UGT-76G1*Sr* and UGT-SA*Cr* respectively (Fig. 2). From this screening, some general features regarding substrate specificity of both UGTs emerged. Firstly, for UGT-76G1*Sr*, acceptors included aro-

matic as well as aliphatic compounds. Interestingly, the enzymatic activity is not limited to linear alcohols, also branched alcohols such as cinnamyl alcohol or geraniol were glucosylated by UGT-76G1*Sr*. This in contrast to UGT-SA*Cr*, which could not glycosylate any aliphatic alcohols. The latter seems similar to the promiscuity of microbial *Bc*GT1, where acceptors needed to contain an aromatic ring system, hypothesizing a π-π interaction between an amino acid in the active site and the aromatic ring (Chiu et al., 2016). Further, *ortho* positioned substituents seemed to be unfavorable for UGT-76G1*Sr*'s glycosylation potential. Examples from the screening include (acceptor vs. non-acceptor): orcinol vs. thymol and hydroquinone vs. pyrogallol. Likewise, *trans*-cyclohexanediol was an acceptor, whereas UGT-76G1*Sr* did not show activity towards *cis*-cyclohexanediol. Further, esters of gallic acid could not be glycosylated by UGT-SA*Cr*. In contrast, UGT-76G1*Sr* showed activity towards these compounds but the glycosylation potential decreased with increasing length of the ester chain. Indeed, methyl-, ethyl- and propylgallate were efficiently glycosylated by UGT-76G1*Sr*, whereas in the case of octyl- and laurylgallate only small product spots were detected. Further, both enzymes enabled glucosylation of flavonoids such as catechin, quercetin and curcumin. Lastly, elongation of the sugar chain moiety is achievable with UGT-76G1*Sr*, not with UGT-SA*Cr*. Indeed, the former enzyme catalyzed a reaction with not only the glucoside of *p*-nitrophenol, but also the galactoside and even the fucoside derivative. It can be stated that both enzymes are promiscuous, with UGT-76G1*Sr* clearly being superior over UGT-SA*Cr*. These results illustrate the versatile applicability of UGT-76G1*Sr* for the production of glyco-

Table 2
Kinetic parameters of UGT-76G1*Sr* and UGT-SA*Cr*.

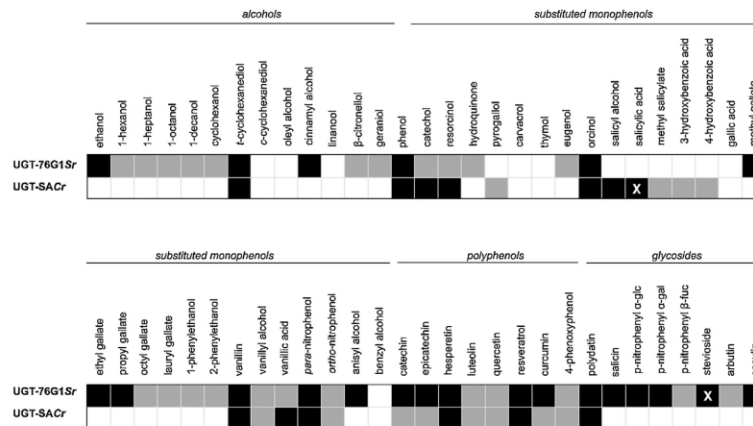|	K_m (mM)	V_{max} (U/mg)	k_{cat}/K_m (M^{-1} s^{-1})
UGT-76G1*Sr*			
UDP-glucose	0.44 ± 0.06	11.06 ± 0.46	2.27 × 10^5
stevioside	0.23 ± 0.05	9.30 ± 0.57	1.80 × 10^5
UGT-SA*Cr*			
UDP-glucose	0.66 ± 0.08	1.58 ± 0.05	1.03 × 10^4
salicylic acid	0.83 ± 0.19	1.92 ± 0.16	1.00 × 10^4

Fig. 2. Acceptor promiscuity of UGT-76G1*Sr* compared to UGT-SA*Cr*. Qualitative survey of promiscuity on 58 diverse acceptors. The semi-quantitative amount of formed product after 24 h incubation of the acceptors with UDP-Glc and the GTs is represented in three categories, namely high (40 − 100%), medium (5 − 40%) and low (< 5%/undetectable), which are represented by black, grey and white squares, respectively. The reaction with the natural acceptor is marked with 'X' and is taken as reference (100%). Reaction conditions: 0.5 mM acceptor, 0.5 mM UDP-glucose and 0.25 mg/mL purified UGT-76G1*Sr* or UGT-SA*Cr* in 100 mM MOPS buffer at pH 7.0.

sides. GTs with broad substrate specificity are generally recognized as powerful tools in the glycodiversification of natural products (Thibodeaux et al., 2008).

3.5. Glycosylation of curcumin

UGT-76G1*Sr* was found here to have a remarkably broad specificity, which can be exploited for the glucosylation of a wide range of acceptor molecules. As an example, the glucosylation of curcumin was examined in more detail. Curcumin is a bright yellow substance commonly used as food colorant and also known to have pharmaceutical properties such as antioxidant, anti-inflammatory, antimicrobial and anti-carcinogenic activities (Anand et al., 2007). Curcumin has also been suggested as treatment for Alzheimer's disease, since it suppresses amyloid β-protein, the causative agent of the disease (Hamaguchi et al., 2010). The biggest problem of curcumin is its very poor water solubility. Nevertheless, glycosylation can enhance water solubility of such hydrophobic compounds significantly (Kaminaga et al., 2003).

A number of parameters, including the presence of cosolvent and additives as well as the concentration of donor and acceptor, are known to influence the glycosylation efficiency. The enzyme concentration was fixed at 2 mg/mL, as this conveniently yielded 10% conversion within 1 h of reaction time. A ten-fold excess of donor substrate UDP-Glc (5 mM) was used to push the reaction towards conversion of 0.5 mM curcumin. DMSO was used as a cosolvent to solubilize the hydrophobic substrate curcumin, with 10% offering the best balance between acceptor solubility and enzyme stability (Fig. 3). Indeed, using 2% DMSO resulted in a large portion of curcumin to remain precipitated, whereas in 20% DMSO enzymatic activity was lost after only 4 h. Interestingly, the acceptor conversion could be further increased up to 93% by adding 500 mM trehalose to the reaction, most likely because it stabilizes the enzyme against the presence of cosolvent. Trehalose is a well-known protein protectant and is not an acceptor of UGT-76G1*Sr*, indicating that the stabilization originates from non-specific interactions at the enzyme surface (De Winter et al., 2013; Szabó et al., 2009). Similarly, the addition of sucrose also led to increased yields, although 80% seemed to be the maximum in this case (Fig. 4). Extending the incubation time from 8 h to 24 h yielded a small

increase in final conversion rates (96% and 85% for 500 mM trehalose and 500 mM sucrose, respectively).

The major drawback of applying GTs in large-scale glucosylation processes, is the fact that they require expensive UDP-glucose as a sugar donor. In this respect, sucrose synthase (SuSy) is an interesting biocatalyst as it can be used for the production of activated sugars starting from the abundant and cheap donor substrate sucrose (Diricks et al., 2015; Elling and Kula, 1993; Elling et al., 1993; Figueroa et al., 2013; Schmölzer et al., 2015). Even more, when SuSy and a GT are combined in a one-pot reaction, the UDP generated by the latter enzyme can instantly be regenerated to UDP-glucose by the former, preventing possible enzyme inhibition by high concentrations of UDP (Bungaruang et al., 2013; Gutmann et al., 2014; Huang et al., 2016; Lepak et al., 2015; Terasaka et al., 2012). Interestingly, as discussed in this work, the presence of sucrose as starting material should have a favorable effect on the stability of UGT-76G1*Sr* in a coupled reaction. The (re)generation system with the SuSy from *Glycine max* (*Gm*SuSy) was tested for the case of curcumin glycosylation (Fig. 5). If UGT-76G1*Sr* (2 mg/mL)
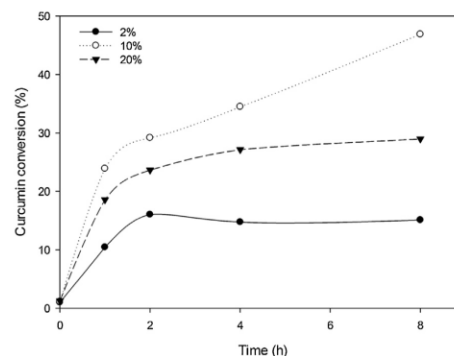

Fig. 3. Effect of varying the amount of cosolvent DMSO (% v/v) on the conversion of curcumin into glucosides. The production of glucosides is presented in function of time. Reaction conditions: 0.5 mM curcumin, 5 mM UDP-Glc, 2 mg/mL UGT-76G1*Sr* in 100 mM MOPS buffer pH 7.0 at 30 °C.
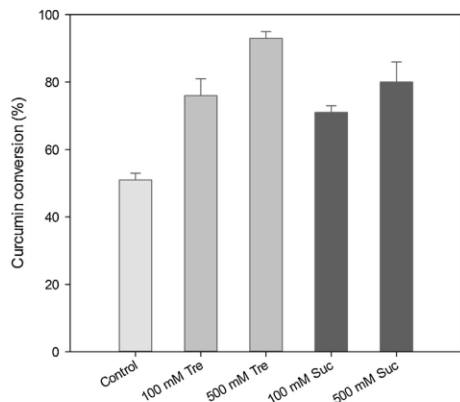
Fig. 4. Effect of varying the amount of trehalose and sucrose added to the reaction mix on the conversion of curcumin into glucosides after 8 h incubation. Reaction conditions: 0.5 mM curcumin, 5 mM UDP-Glc, 2 mg/mL UGT-76G1*Sr*, 10% (v/v) DMSO and a varying concentration of Tre or Suc, in 100 mM MOPS buffer pH 7.0 at 30 °C. Tre and Suc were omitted from the reaction mixture in the control. Tre: trehalose; Suc: sucrose.

was used without *Gm*SuSy, the glucosides were formed at a rate of 99 µM/h, whereas a rate of 135 µM/h could be obtained using the coupled system (2 mg/mL UGT-76G1*Sr* and 0.075 mg/mL *Gm*SuSy). After 24 h of reaction, the coupled system reached 95% conversion with respect to the initial acceptor concentration. Interestingly, the product mainly consisted of curcumin diglucoside (69%) and only contained minor amounts of the monoglucoside, which is similar to what has been reported for the CaUGT2 from *Catharanthus roseus* (Masada et al., 2007). The kinetic profile of the production of mono- and diglucoside suggested both products are formed consecutively (Supplementary S6). The products were purified from the reaction mixture and NMR analysis was in agreement with the expected positioning of glucose groups, being beta attached to one or both free OH-groups of curcumin (Supplementary S7).
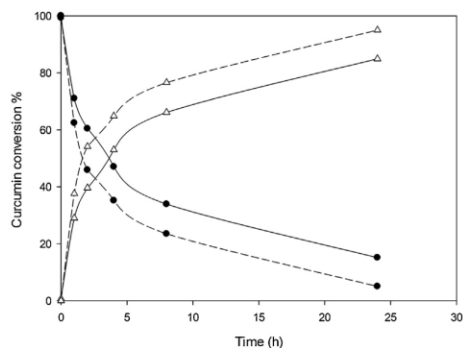

Fig. 5. Glycosylation of curcumin by direct action of UGT-76G1*Sr* (full line) compared to the coupled system with UDP-Glc (re)generation using *Gm*SuSy (dashed line). The conversion of curcumin and the production of its glucosides is presented in function of time. Reaction conditions: Full lines: 0.5 mM curcumin, 5 mM UDP-Glc, 100 mM sucrose, 2 mg/mL UGT-76G1*Sr*; Dashed lines: 0.5 mM curcumin, 1 mM UDP, 100 mM sucrose, 2 mg/mL UGT-76G1*Sr*, 0.075 mg/mL *Gm*SuSy; both in 10% DMSO and 100 mM MOPS buffer pH 7.0 at 30 °C. Closed circles: curcumin; Open triangles: total glucoside product.

4. Conclusions

In this work, screening of a diversity of GT enzymes revealed that some of them are more promiscuous than could be expected from literature. Indeed, recombinantly expressed UGT-76G1 from *Stevia rebaudiana* showed glycosylating activity not only towards aliphatic and branched alcohols, but also towards (substituted) phenols, flavonoids, gallates and glycosides. Careful optimization of reaction conditions allowed the glucosylation of curcumin up to a yield of 96%. Interestingly, the production of curcumin β-glucosides was also achieved by coupling UGT-76G1*Sr* to *Gm*SuSy in order to (re)generate expensive UDP-glucose from sucrose as substrate. In conclusion, the broad acceptor specificity discovered here for UGT-76G1*Sr* offers valuable potential for the glycosylation of a myriad of natural products. In addition, the enzyme's flexibility to fit in one-pot synthesis setups makes it an attractive biocatalyst for cost-effective glycosylation. Enzymatic structure-function relationship studies might give us more insight in the underlying reason of this remarkable promiscuity, information that might prove useful to engineer new glycosylation biocatalysts with broad specificity.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jbiotec.2016.06.034.

References

Aerts, D., Verhaeghe, T., De Mey, M., Desmet, T., Soetaert, W., 2011. A constitutive expression system for high-throughput screening. Eng. Life Sci. 11, 10–19, http://dx.doi.org/10.1002/elsc.201000065.
Anand, P., Kunnumakkara, A.B., Newman, R.A., Aggarwal, B.B., 2007. Bioavailability of curcumin: problems and promises. Mol. Pharm. 4, 807–818, http://dx.doi.org/10.1021/mp700113r.
Bönisch, F., Frotscher, J., Stanitzek, S., Rühl, E., Wüst, M., Bitz, O., Schwab, W., 2014a. A UDP-glucose: monoterpenol glucosyltransferase adds to the chemical diversity of the grapevine metabolome. Plant Physiol. 165, 561–581, http://dx.doi.org/10.1104/;pp.113.232470.
Bönisch, F., Frotscher, J., Stanitzek, S., Rühl, E., Wüst, M., Bitz, O., Schwab, W., 2014b. Activity based profiling of a physiologic aglycone library reveals sugar acceptor promiscuity of family 1 UDP-glucosyltransferases from vitis vinifera. Plant Physiol. 166, 23–39, http://dx.doi.org/10.1104/pp.114.242578.
Bowles, D., Isayenkova, J., Lim, E.K., Poppenberger, B., 2005. Glycosyltransferases: managers of small molecules. Curr. Opin. Plant Biol. 8, 254–263, http://dx.doi.org/10.1016/j.pbi.2005.03.007.
Bungaruang, L., Gutmann, A., Nidetzky, B., 2013. Leloir glycosyltransferases and natural product glycosylation: biocatalytic synthesis of the C-glucoside nothofagin, a major antioxidant of redbush herbal tea. Adv. Synth. Catal. 355, 2757–2763, http://dx.doi.org/10.1002/adsc.201300251.
Cantarel, B.I., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., Henrissat, B., 2009. The carbohydrate-active EnZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res. 37, 233–238, http://dx.doi.org/10.1093/nar/gkn663.
Ceunen, S., Geuns, J.M.C., 2013. Steviol glycosides: chemical diversity, metabolism, and function. J. Nat. Prod. 76, 1201–1228, http://dx.doi.org/10.1021/np400203b.
Chiu, H.-H., Shen, M.-Y., Liu, Y.-T., Fu, Y.-L., Chiu, Y.-A., Chen, Y.-H., Huang, C.-P., Li, Y.-K., 2016. Diversity of sugar acceptor of glycosyltransferase 1 from Bacillus cereus and its application for glucoside synthesis. Appl. Microbiol. Biotechnol., http://dx.doi.org/10.1007/s00253-015-7270-1.
De Winter, K., Verlinden, K., Křen, V., Weignerová, L., Soetaert, W., Desmet, T., 2013. Ionic liquids as cosolvents for glycosylation by sucrose phosphorylase: balancing acceptor solubility and enzyme stability. Green Chem. 15, 1949–1955, http://dx.doi.org/10.1039/c3gc40449h.
De Winter, K., Dewitte, G., Dirks-Hofmeister, M.E., De Laet, S., Pelantová, H., Křen, V., Desmet, T., 2015. Enzymatic glycosylation of phenolic antioxidants:

phosphorylase-mediated synthesis and characterization. J. Agric. Food Chem. 63, 10131–10139, http://dx.doi.org/10.1021/acs.jafc.5b04380.

Desmet, T., Soetaert, W., Bojarová, P., Křen, V., Dijkhuizen, L., Eastwick-Field, V., Schiller, A., 2012. Enzymatic glycosylation of small molecules: challenging substrates require tailored catalysts. Chem. A Eur. J. 18, 10786–10801, http://dx.doi.org/10.1002/chem.201103069.

Diricks, M., De Bruyn, F., Van Daele, P., Walmagh, M., Desmet, T., 2015. Identification of sucrose synthase in nonphotosynthetic bacteria and characterization of the recombinant enzymes. Appl. Microbiol. Biotechnol. 99, 8465–8474, http://dx.doi.org/10.1007/s00253-015-6548-7.

Dirks-Hofmeister, M.E., Verhaeghe, T., De Winter, K., Desmet, T., 2015. Creating space for large acceptors: rational biocatalyst design for resveratrol glycosylation in an aqueous system. Angew. Chemie – Int. Ed. 54, 9289–9292, http://dx.doi.org/10.1002/anie.201503605.

Elling, L., Kula, M.R., 1993. Purification of sucrose synthase from rice and its protein-chemical characterization. J. Biotechnol. 29, 277–286, http://dx.doi.org/10.1016/0168-1656(93)90059-V.

Elling, L., Grothus, M., Kula, M.R., 1993. Investigation of sucrose synthase from rice for the synthesis of various nucleotide sugars and saccharides. Glycobiology 3, 349–355, http://dx.doi.org/10.1093/glycob/3.4.349.

Figueroa, C.M., Asención Diez, M.D., Kuhn, M.L., McEwen, S., Salerno, G.L., Iglesias, A. a., Ballicora, M. a., 2013. The unique nucleotide specificity of the sucrose synthase from Thermosynechococcus elongatus. FEBS Lett. 587, 165–169, http://dx.doi.org/10.1016/j.febslet.2012.11.011.

Fukuchi-Mizutani, M., Okuhara, H., Fukui, Y., Nakao, M., Katsumoto, Y., Yonekura-sakakibara, K., Kusumi, T., Hase, T., Tanaka, Y., 2003. Biochemical and molecular characterization of a novel UDP-Glucose: anthocyanin 3′-O-Glucosyltransferase, a key enzyme for blue anthocyanin biosynthesis, from gentian. Plant Physiol. 132, 1652–1663, http://dx.doi.org/10.1104/pp.102.018242.

Gachon, C.M.M., Langlois-Meurinne, M., Saindrenan, P., 2005. Plant secondary metabolism glycosyltransferases: the emerging functional analysis. Trends Plant Sci. 10, 542–549, http://dx.doi.org/10.1016/j.tplants.2005.09.007.

Gantt, R.W., Goff, R.D., Williams, G.J., Thorson, J.S., 2008. Probing the aglycon promiscuity of an engineered glycosyltransferase. Angew. Chemie – Int. Ed. 47, 8889–8892, http://dx.doi.org/10.1002/anie.200803508.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.a., Smith, H.O., 2009. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods 6, 343–345, http://dx.doi.org/10.1038/NMETH.1318.

Gosselin, S., Alhussaini, M., Streiff, M.B., Takabayashi, K., Palcic, M.M., 1994. A continuous spectrophotometric assay for glycosyltransferases. Anal. Biochem. 220, 92–97, http://dx.doi.org/10.1006/abio.1994.1303.

Gross, G.G., 1982. Synthesis of b-glucogallin from UDP-glucose and gallic acid by an enzyme preparation from oak leaves. FEBS Lett. 148, 67–70, http://dx.doi.org/10.1016/0014-5793(82)81244-1.

Gross, G.G., 1983. Partial purification and properties of UDP-glucose: vanillate 1-O-glucosyl transferase from oak leaves. Phytochemistry 22, 2179–2182, http://dx.doi.org/10.1016/S0031-9422(00)80141-7.

Gutmann, A., Bungaruang, L., Weber, H., Leypold, M., Breinbauer, R., Nidetzky, B., 2014. Towards the synthesis of glycosylated dihydrochalcone natural products using glycosyltransferase-catalysed cascade reactions. Green Chem. 16, 4417–4425, http://dx.doi.org/10.1039/c4gc00960f.

Hamaguchi, T., Ono, K., Yamada, M., 2010. Curcumin and Alzheimer's disease. CNS Neurosci. Ther. 16, 285–297, http://dx.doi.org/10.1111/j.1755-5949.2010.00147.x.

Hansen, K.S., Kristensen, C., Tattersall, D.B., Jones, P.R., Olsen, C.E., Bak, S., Møller, B.L., 2003. The in vitro substrate regiospecificity of recombinant UGT85B1, the cyanohydrin glucosyltransferase from Sorghum bicolor. Phytochemistry 64, 143–151, http://dx.doi.org/10.1016/S0031-9422(03)00261-9.

Huang, F.-C., Hinkelmann, J., Hermenau, A., Schwab, W., 2016. Enhanced production of β-glucosides by in-situ UDP-glucose regeneration. J. Biotechnol. 224, 35–44, http://dx.doi.org/10.1016/j.jbiotec.2016.02.022.

Jain, K., Kesharwani, P., Gupta, U., Jain, N.K., 2012. A review of glycosylated carriers for drug delivery. Biomaterials 33, 4166–4186, http://dx.doi.org/10.1016/j.biomaterials.2012.02.033.

Kaminaga, Y., Nagatsu, A., Akiyama, T., Sugimoto, N., Yamazaki, T., Maitani, T., Mizukami, H., 2003. Production of unnatural glucosides of curcumin with drastically enhanced water solubility by cell suspension cultures of Catharanthus roseus. FEBS Lett. 555, 311–316, http://dx.doi.org/10.1016/S0014-5793(03)01265-1.

Kren, V., Rezanka, T., 2008. Sweet antibiotics – the role of glycosidic residues in antibiotic and antitumor activity and their randomization. FEMS Microbiol. Rev. 32, 858–889, http://dx.doi.org/10.1111/j.1574-6976.2008.00124.x.

Kwon, T., Kim, C.T., Lee, J.H., 2007. Transglucosylation of ascorbic acid to ascorbic acid 2-glucoside by a recombinant sucrose phosphorylase from Bifidobacterium longum. Biotechnol. Lett. 29, 611–615, http://dx.doi.org/10.1007/s10529-006-9285-2.

Lairson, L.L., Henrissat, B., Davies, G.J., Withers, S.G., 2008. Glycosyltransferases: structures, functions, and mechanisms. Annu. Rev. Biochem. 77, 521–555, http://dx.doi.org/10.1146/annurev.biochem.76.061005.092322.

Lee Il, H., Raskin, I., 1999. Purification, cloning, and expression of a pathogen inducible UDP- glucose:salicylic acid glucosyltransferase from tobacco. J. Biol. Chem. 274, 36637–36642, http://dx.doi.org/10.1074/jbc.274.51.36637.

Lepak, A., Gutmann, A., Kulmer, S.T., Nidetzky, B., 2015. Creating a water-Soluble resveratrol-based antioxidant by site-Selective enzymatic glucosylation. Chembiochem 16, 1870–1874, http://dx.doi.org/10.1002/cbic.201500284.

Lim, E.K., Doucet, C.J., Li, Y., Elias, L., Worrall, D., Spencer, S.P., Ross, J., Bowles, D.J., 2002. The activity of Arabidopsis glycosyltransferases toward salicylic acid, 4-hydroxybenzoic acid, and other benzoates. J. Biol. Chem. 277, 586–592, http://dx.doi.org/10.1074/jbc.M109287200.

Martinkus, C., Croteau, R., 1981. Metabolism of monoterpenes: evidence for compartmentation of L-menthone metabolism in peppermint (Mentha piperita) leaves. Plant Physiol. 68, 99–106, http://dx.doi.org/10.1104/pp.68.1.99.

Masada, S., Kawase, Y., Nagatoshi, M., Oguchi, Y., Terasaka, K., Mizukami, H., 2007. An efficient chemoenzymatic production of small molecule glucosides with in situ UDP-glucose recycling. FEBS Lett. 581, 2562–2566, http://dx.doi.org/10.1016/j.febslet.2007.04.074.

Miller, K.D., Guyon, V., Evans, J.N.S., Shuttleworth, W. a., Taylor, L.P., 1999. Purification, cloning, and heterologous expression of a catalytically efficient flavonol 3-O-galactosyltransferase expressed in the male gametophyte of Petunia hybrida. J. Biol. Chem. 274, 34011–34019, http://dx.doi.org/10.1074/jbc.274.48.34011.

Mohamed, A.A.A., Ceunen, S., Geuns, J.M.C., Van den Ende, W., De Ley, M., 2011. UDP-dependent glycosyltransferases involved in the biosynthesis of steviol glycosides. J. Plant Physiol. 168, 1136–1141, http://dx.doi.org/10.1016/j.jplph.2011.01.030.

Muthana, S., Cao, H., Chen, X., 2009. Recent progress in chemical and chemoenzymatic synthesis of carbohydrates. Curr. Opin. Chem. Biol. 13, 573–581, http://dx.doi.org/10.1016/j.cbpa.2009.09.013.

Mydock, L.K., Kamat, M.N., Demchenko, A.V., 2011. Direct synthesis of diastereomerically pure glycosyl sulfonium salts. Org. Lett. 13, 2928–2931, http://dx.doi.org/10.1021/ol2009818.

Oh, T.-J., Kim, D.H., Kang, S.Y., Yamaguchi, T., Sohng, J.K., 2011. Enzymatic synthesis of vancomycin derivatives using galactosyltransferase and sialyltransferase. J. Antibiot. (Tokyo). 64, 103–109, http://dx.doi.org/10.1038/ja.2010.131.

Ohgami, S., Ono, E., Horikawa, M., Murata, J., Totsuka, K., Toyonaga, H., Ohba, Y., Dohra, H., Asai, T., Matsui, K., Mizutani, M., Watanabe, N., Ohnishi, T., 2015. Volatile glycosylation in tea plants: sequential glycosylations for the biosynthesis of aroma β-primeverosides are catalyzed by two Camellia sinensis glycosyltransferases. Plant Physiol. 168, 464–477, http://dx.doi.org/10.1104/pp.15.00403.

Osmani, S. a., Bak, S., Møller, B.L., 2009. Substrate specificity of plant UDP-dependent glycosyltransferases predicted from crystal structures and homology modeling. Phytochemistry 70, 325–347, http://dx.doi.org/10.1016/j.phytochem.2008.12.009.

Palcic, M.M., 2011. Glycosyltransferases as biocatalysts. Curr. Opin. Chem. Biol. 15, 226–233, http://dx.doi.org/10.1016/j.cbpa.2010.11.022.

Pandey, R.P., Gurung, R.B., Parajuli, P., Koirala, N., Tuoi, L.T., Sohng, J.K., 2014. Assessing acceptor substrate promiscuity of YjiC-mediated glycosylation toward flavonoids. Carbohydr. Res. 393, 26–31, http://dx.doi.org/10.1016/j.carres.2014.03.011.

Richman, A., Swanson, A., Humphrey, T., Chapman, R., McGarvey, B., Pocs, R., Brandle, J., 2005. Functional genomics uncovers three glucosyltransferases involved in the synthesis of the major sweet glucosides of Stevia rebaudiana. Plant J. 41, 56–67, http://dx.doi.org/10.1111/j.1365-313X.2004.02275.x.

Schmölzer, K., Gutmann, A., Diricks, M., Desmet, T., Nidetzky, B., 2015. Sucrose synthase: a unique glycosyltransferase for biocatalytic glycosylation process development. Biotechnol. Adv. 34, 88–111, http://dx.doi.org/10.1016/j.biotechadv.2015.11.003.

Schmid, J., Heider, D., Wendel, N.J., Sperl, N., Sieber, V., 2016. Bacterial glycosyltransferases: challenges and opportunities of a highly diverse enzyme class toward tailoring natural products. Front. Microbiol. 7, 1–7, http://dx.doi.org/10.3389/fmicb.2016.00182.

Schwab, W., Fischer, T., Wüst, M., 2015. Terpene glucoside production: improved biocatalytic processes using glycosyltransferases. Eng. Life Sci. 15, 376–386, http://dx.doi.org/10.1002/elsc.201400156.

Sharma, R., Panigrahi, P., Suresh, C.G., 2014. In-Silico analysis of binding site features and substrate selectivity in plant flavonoid-3-O glycosyltransferases (F3GT) through molecular modeling, docking and dynamics simulation studies. PLoS One 9, e92636, http://dx.doi.org/10.1371/journal.pone.0092636.

Sicard, A., Stacey, N., Hermann, K., Dessoly, J., Neuffer, B., Bäurle, I., Lenhard, M., 2011. Genetics, evolution, and adaptive significance of the selfing syndrome in the genus Capsella. Plant Cell 23, 3156–3171, http://dx.doi.org/10.1105/tpc.111.088237.

Singh, S., Rao, G., 2005. Stevia: the herbal sugar of 21 st century. Sugar Tech. 71, 17–24, http://dx.doi.org/10.1007/BF02942413.

Szabó, A., Kotormán, M., Laczkó, I., Simon, L.M., 2009. Influence of carbohydrates on stability of papain in aqueous tetrahydrofuran mixture. J. Chem. Technol. Biotechnol. 84, 133–138, http://dx.doi.org/10.1002/jctb.2018.

Terasaka, K., Mizutani, Y., Nagatsu, A., Mizukami, H., 2012. In situ UDP-glucose regeneration unravels diverse functions of plant secondary product glycosyltransferases. FEBS Lett. 586, 4344–4350, http://dx.doi.org/10.1016/j.febslet.2012.10.045.

Thibodeaux, C.J., Melançon, C.E., Liu, H.W., 2008. Natural-product sugar biosynthesis and enzymatic glycodiversification. Angew. Chemie – Int. Ed. 47, 9814–9859, http://dx.doi.org/10.1002/anie.200801204.

1   **Supporting information**

2

3   **Screening of Recombinant Glycosyltransferases Reveals the Broad Acceptor Specificity of Stevia**

4   **UGT-76G1**

5

6   Griet Dewitte[a], Maarten Walmagh[a], Margo Diricks[a], Alexander Lepak[b], Alexander Gutmann[b], Bernd

7   Nidetzky[b, c] and Tom Desmet[a,*]

8

9   [a] Centre for Industrial Biotechnology and Biocatalysis, Department of Biochemical and Microbial

10   Technology, Ghent University, Coupure Links 653, B-9000 Ghent, Belgium

11   [b] Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, NAWI Graz,

12   Petersgasse 12, A-8010 Graz, Austria

13   [c] Austrian Centre of Industrial Biotechnology, Petersgasse 14, 8010 Graz, Austria

14

15   * Corresponding author:

16   Tel.: +32 9 264 99 20; Fax: +32 9 264 60 32; E-mail address: tom.desmet@ugent.be

17

18

S1: Sequences of selected GTs

20 **<UGT-76G1*Sr*:** gb|AGL95113.1| (Glycosyltransferase_GTB_type superfamily)

21 *DNA-sequence*

22 ATGGAAAATAAAACCGAAACCACCGTCCGTCGCCGTCGTCGTATCATTCTGTTCCCGGTCCCGTTCCAAGGTCACATCAACCCGATTCTGC
23 AGCTGGCCAACGTGCTGTATAGCAAAGGTTTCTCTATCACCATCTTCCATACGAACTTCAACAAACCGAAAACCTCTAACTACCCGCACTT
24 TACGTTCCGTTTTATTCTGGATAACGACCCGCAGGATGAACGCATCAGTAATCTGCCGACCCATGGTCCGCTGGCGGGTATGCGTATTCC
25 GATTATCAACGAACACGGCGCAGATGAACTGCGTCGCGAACTGGAACTGCTGATGCTGGCCTCTGAAGAAGATGAAGAAGTTAGTTGCC
26 TGATCACCGACGCACTGTGGTATTTTGCCCAGAGTGTTGCAGATTCCCTGAACCTGCGTCGCCTGGTCCTGATGACGAGCTCTCTGTTCAA
27 TTTTCATGCCCACGTTTCCCTGCCGCAGTTCGATGAACTGGGTTATCTGGACCCGGATGACAAAACCCGCCTGGAAGAACAAGCTTCAGG
28 CTTTCCGATGCTGAAAGTCAAAGATATTAAAAGTGCGTACTCCAACTGGCAGATTCTGAAAGAAATCCTGGGTAAAATGATCAAACAAAC
29 CCGTGCAAGTTCCGGCGTCATCTGGAATTCCTTCAAAGAACTGGAAGAATCAGAACTGGAAACGGTGATTCGCGAAATCCCGGCTCCGT
30 CTTTTCTGATTCCGCTGCCGAAACATCTGACCGCGTCATCGAGCTCTCTGCTGGATCACGACCGTACGGTGTTTCAGTGGCTGGATCAGCA
31 ACCGCCGAGTTCCGTGCTGTACGTTAGCTTCGGTAGCACCTCTGAAGTGGATGAAAAAGACTTTCTGGAAATCGCTCGTGGCCTGGTTGA
32 TTCAAAACAATCGTTCCTGTGGGTGGTTCGCCCGGGTTTTGTGAAAGGCAGCACGTGGGTTGAACCGCTGCCGGATGGCTTCCTGGGTG
33 AACGTGGTCGCATTGTCAAATGGGTGCCGCAGCAAGAAGTGCTGGCACATGGTGCTATCGGCGCGTTTTGGACCCACTCAGGTTGGAAC
34 TCGACGCTGGAAAGCGTTTGTGAAGGTGTCCCGATGATTTTCTCGGATTTTGGCCTGGACCAGCCGCTGAATGCACGTTATATGAGCGAT
35 GTTCTGAAAGTCGGTGTGTACCTGGAAAACGGTTGGGAACGCGGCGAAATTGCGAATGCCATCCGTCGCGTTATGGTCGATGAAGAAG
36 GCGAATATATCCGTCAGAATGCTCGCGTCCTGAAACAAAAAGCGGACGTTAGTCTGATGAAAGGCGGTTCATCGTACGAATCCCTGGAA
37 TCACTGGTCTCCTACATTTCTTCTCTGGGCTCGCATCATCATCATCATCATTA

38 *AA-sequence*

39 MENKTETTVRRRRIILFPVPFQGHINPILQLANVLYSKGFSITIFHTNFNKPKTSNYPHFTFRFILDNDPQDERISNLPTHGPLAGMRIPIINEHGA
40 DELRRELELLMLASEEDEEVSCLITDALWYFAQSVADSLNLRRLVLMTSSLFNFHAHVSLPQFDELGYLDPDDKTRLEEQASGFPMLKVKDIKSA
41 YSNWQILKEILGKMIKQTRASSGVIWNSFKELEESELETVIREIPAPSFLIPLPKHLTASSSSLLDHDRTVFQWLDQQPPSSVLYVSFGSTSEVDEKD
42 FLEIARGLVDSKQSFLWVVRPGFVKGSTWVEPLPDGFLGERGRIVKWVPQQEVLAHGAIGAFWTHSGWNSTLESVCEGVPMIFSDFGLDQPL
43 NARYMSDVLKVGVYLENGWERGEIANAIRRVMVDEEGEYIRQNARVLKQKADVSLMKGGSSYESLESLVSYISSLGSHHHHHH

44

45 **<UGT-SA*Cr*:** ref|XP_006294206.1| (Glycosyltransferase_GTB_type superfamily)

46 *DNA-sequence*

47 ATGGAGAAGAAACGTGGTCATGTCCTGGCAGTGCCGTATCCGACCCAGGGCCACATTACGCCGATCCGTCAATTTTGCAAACGCCTGCAT
48 TACAAAGGCCTGAAAACCACGCTGGCTCTGACCACGTTTATTTTCAACTCTATCAAACCGGATCTGTCGGGCCCGATTAGCATCGCGACC
49 ATTTCGGACGGTTATGATCATGGCGGTTTTGACACCGCCGGCAGCATCCACGATTACCTGCAGAACTTCAAAACCAGTGGTTCCAAAACG
50 ATTGCGGACATTATCCGTAAACACCAAACCAGTGATAATCCGGTTACGTGCATCGTCTATGACGCCTTTATGCCGTGGGCGCTGGATGTC
51 GCACGTGAATTCGGTCTGGTGGCGACCCCGTTTTTCACGCAGCCGTGTGCCGTTAACTACGTCTACTACCTGTCGTACATCAACAATGGCA
52 GCCTGGAACTGCCGATCGAAGAACTGCCGTTTCTGGAACTGCAGGACCTGCCGAGCTTTTTCTCAGTCTCGGGTAGCTATCCGGCATACT
53 TCGAAATGGTTCTGCAGCAATTCATCAACTTCGAAAAAGCTGATTTCGTTCTGGTCAATAGCTTCCAGGAACTGGAACTGCATGAAAAAG
54 AACTGTGGTCTAAAAGTTGCCCGGTGCTGACCATTGGTCCGACGATTCCGTCTATCTATCTGGATCAGCGTATCAAAAGTGATACCGACT
55 ACGATCTGAACCTGTTTGGCCTCTACGGACGCGAGTTATTGTACCAATTGGCTGGATACGCGTCCGCAGGGCTCGGTGGTTTACGTGGCAT
56 TCGGTAGCATGGCTCAGCTGAACAATGTGCAAATGGAAGAACTGGCGCTGGCCGTTTCCAAATTTTCATTCCTGTGGGTCGTGCGTAGCT
57 CTGAAGAGCAAAACTGCCGAGCGGCTTTCTGGAAACCGTTGATAAAGAAAAATCTATGGTCCTGAAATGGAGTCCGCAGCTGCAAGTG
58 CTGTCTAACAAAGCCATTGGCTGCTTTCTGACCCATTGTGGTTGGAATAGTACCATGGAAGCACTGACGTTCGGTGTGCCGATGGTTGCT
59 ATGCCGCAGTGGACCGACCAACCGATGAACTCCAAATATATCCAGGATGTTTGGAAAGCGGGTGTGCGTGTTAAAACCGATAAAGAATC
60 AGGCATTGCAAAACGCGCTGAAATTGAATTTTCCATCCAAGAAGTGATGGAAGGCGAAAAATCAGTGGAAATGAAGAAAAACGTTAAA
61 AAATGGCGCGATCTGGCCGTTAAATCCCTGAACGAAGGCGGTTCAACCGACATTAATATCGATACGTTTGTCTCCAAAGTGGAATCAAAA
62 GGCGGTCATCACCATCACCATCACTAA

63 *AA-sequence*

64 MEKKRGHVLAVPYPTQGHITPIRQFCKRLHYKGLKTTLALTTFIFNSIKPDLSGPISIATISDGYDHGGFDTAGSIHDYLQNFKTSGSKTIADIIRKH
65 QTSDNPVTCIVYDAFMPWALDVAREFGLVATPFFTQPCAVNYVYYLSYINNGSLELPIEELPFLELQDLPSFFSVSGSYPAYFEMVLQQFINFEKA

66 DFVLVNSFQELELHEKELWSKSCPVLTIGPTIPSIYLDQRIKSDTDYDLNLFGSTDASYCTNWLDTRPQGSVVYVAFGSMAQLNNVQMEELALA
67 VSKFSFLWVVRSSEEAKLPSGFLETVDKEKSMVLKWSPQLQVLSNKAIGCFLTHCGWNSTMEALTFGVPMVAMPQWTDQPMNSKYIQDVW
68 KAGVRVKTDKESGIAKRAEIEFSIQEVMEGEKSVEMKKNVKKWRDLAVKSLNEGGSTDINIDTFVSKVESKGGHHHHHH

69

70 **<GalT-F3*Ph***: sp|Q9SBQ8.1|KGLT_PETHY (Glycosyltransferase_GTB_type superfamily)

71 *DNA-sequence*

72 ATGTCTAACTATCATGTTGCGGTCCTGGCCTTTCCGTTCGCAACCCACGCTGGTCTGCTGCTGGGTCTGGTCCAGCGTCTGGCTAACGCGC
73 TGCCGAATGTGACCTTTACGTTTTTCAACACGTCAAAATCGAATAGCTCTCTGTTCACCACGCCGCATGATAACAACATCAAACCGTTCAA
74 CATCAGCGACGGCGTGCCGGAAGGTTACGTGGTTGGCAAAGGCGGTATTGAAGCACTGATCGGTCTGTTTTTCAAATCTGCGAAAGAAA
75 ACATTCAGAATGCCATGGCGGCCGCAGTTGAAGAAAGTGGCAAGAAAATTACCTGCGTCATGGCCGATGCATTTATGTGGTTCTCCGGT
76 GAAATTGCCGAAGAACTGTCAGTGGGCTGGATCCCGCTGTGGACCTCTGCTGCGGGTAGCCTGTCTGTGCATGTTTATACGGATCTGATT
77 CGTGAAAACGTTGAAGCACAAGGCATCGCTGGTCGCGAAGACGAAATTCTGACCTTTATCCCGGGTTTCGCGGAACTGCGTCTGGGCAG
78 TCTGCCGTCCGGTGTCGTGAGTGGCGATCTGGAATCACCGTTCTCGGTGATGCTGCACAAAATGGGTAAAACCATTGGCAAAGCTACGG
79 CGCTGCCGGTGAATTCGTTTGAAGAACTGGATCCGCCGATCGTTGAAGACCTGAAAAGCAAATTCAACAATTTCCTGAACGTTGGTCCGT
80 TCAATCTGACCACCCCGCCGCCGTCGGCAAACATTACCGATGAATATGGTTGTATCGCATGGCTGGACAAACAGGAACCGGGCAGCGTG
81 GCATACATTGGCTTTGGTACCGTTGCTACCCCGCCGCCGAATGAACTGAAAGCCATGGCAGAAGCTCTGGAAGAATCAAAAACCCCGTTT
82 CTGTGGTCGCTGAAAGACCTGTTTAAAAGCTTTTTCCCGGAAGGTTTTCTGGAACGCACGTCTGAATATGGCAAAATTGTGAGTTGGGCG
83 CCGCAGGTCCAAGTGCTGAGCCATGGCTCTGTTGGTGTCTTTATCAACCACTGCGGTTGGAATAGTGTTCTGGAATCCATTGCCGCAGGT
84 GTGCCGGTTATCTGTCGTCCGTTTTTCGGCGATCATCAACTGAACGCGTGGATGGTCGAAAAAGTGTGGAAAATTGGTGTTAAAATCGAA
85 GGCGGTGTCTTCACCAAAGACGGCACGATGCTGGCCCTGGATCTGGTGCTGAGCAAAGACAAACGCAATACCGAACTGAAACAGCAAA
86 TTGGCATGTACAAAGAACTGGCACTGAACGCAGTCGGTCCGTCCGGTAGTTCCGCAGAAAACTTCAAAAAACTGGTTGATATCATCACCT
87 CATGCAACGGCGGTCATCACCATCACCATCACTAA

88 *AA-sequence*

89 MSNYHVAVLAFPFATHAGLLLGLVQRLANALPNVTFTFFNTSKSNSSLFTTPHDNNIKPFNISDGVPEGYVVGKGGIEALIGLFFKSAKENIQNA
90 MAAAVEESGKKITCVMADAFMWFSGEIAEELSVGWIPLWTSAAGSLSVHVYTDLIRENVEAQGIAGREDEILTFIPGFAELRLGSLPSGVVSGD
91 LESPFSVMLHKMGKTIGKATALPVNSFEELDPPIVEDLKSKFNNFLNVGPFNLTTPPPSANITDEYGCIAWLDKQEPGSVAYIGFGTVATPPPNEL
92 KAMAEALEESKTPFLWSLKDLFKSFFPEGFLERTSEYGKIVSWAPQVQVLSHGSVGVFINHCGWNSVLESIAAGVPVICRPFFGDHQLNAWMV
93 EKVWKIGVKIEGGVFTKDGTMLALDLVLSKDKRNTELKQQIGMYKELALNAVGPSGSSAENFKKLVDIITSCNGGHHHHHH

94

95 **<UGT-MT*Vv***: ref|XP_003634093.1| (Glycosyltransferase_GTB_type superfamily)

96 *DNA-sequence*

97   ATGTTTATGCTATAATTTTTCGGTATCCATGGCGGCCGCTCTAGAAGAAGCTTGGGATCCGTCGACCTCGAATTCGGAGGAAACAAAGAT
98   GGAAAGTGTGGTTCTGTATCCGTCCCCGGGCATGGGTCATCTGATCAGCATGGTGGAACTGGGCAAACTGATTCTGAAACATCACCCGA
99   GTTTTTCCATTACCATCTTCATTGTTACGCCGCCGTATAACACCGGTTCTACGGCACCGTACCTGGCTCGTGTCAGCTCTACCATCCCGAGT
100 TTTACGTTCCATCACCTGCCGACCATTAGTCTGCCGCTGGATTCCTTTAGTTCCCGAACCATGAAACGCTGGCGTTCGAACTGCTGCGTCT
101 GAACAATCCGAATGTTCACCAGGCCCTGGTCTCAATCTCGAACAATTCATCGGTGCGCGCGCTGATTGTTGATTCATTTTGCACCACGGCA
102 CTGTCGGTTGCAGCACAACTGAACATCCCGTGTTATTACTTTTTCACCAGCGGCGGTGCTGCTGTCTGGCCTCTTATCTGTACCTGCCGTTTAT
103 TCATCAGCAAACCACGAAATCTTTCAAAGATCTGAATACCCATCTGCACATCCCGGGTCTGCCGCCGGTGCCGGCAAGTGACATGGCTAA
104 ACCGATTCTGGATCGTGAAGACAAAGCATACGAATTTTTCCTGAACATGTTTATCCACCTGCCGCGTTCTGCTGGCATTATCGTTAATACC
105 TTCGAAGCACTGGAACCGCGCGCTGTCAAAACGATTCTGGATGGTCTGTGTGCGTGCTGGACAGGTCCGACCAGCCCGATCTTCTGTATTGGT
106 CCGCTGATCGCAACGGATGACCGTTCAGGCGGTGGCGGTGGCGGTGGCGGTGGCATTCCGGAATGCCTGACCTGGCTGGAATCACAGC
107 CGAAACGCTCGGTGCTGTTTCTGTGTTTCGGCAGCCTGGGTCTGTTTTCTGAAGAACAGCTGAAAGAAATTGCCGTTGGCCTGGAACGTA
108 GTGGTCAACGCTTTCTGTGGGTCGTGCGTAGCCCGCCGTCTAAAGATCCGTCGCGTCGCTTCCTGGCACCGCCGGATCGGGACCTGAACA
109 GCCTGCTGCCGGATGGCTTTCTGGACCGTACCAAAGAACGCGGTCTGGTTGTCAAAAGTTGGGCACCGCAGGTCGCAGTGCTGAACCAT
110 GCGTCCGTGGGTGGCTTTGTTACCCACTGCGGCTGGAATTCAGTCCTGGAAGCGGTGTGTGCCGGTGTTCCGATGGTCGCATGGCCGCT
111 GTATGCCGAACAACGTTTCAATCGCGTGGTTATGGTCGAAGAACTGAAAGTGCCGCTGCCGATGGAAGAATCCGAAGAAGGCTTTATCA
112 CCGCTACGGAAGTTGAAAAACGTGGCCGCGAACTGATGGAAAGCGAAGAAGGTAATACCCTGCGTCTGCGCATTATGGCAATGAAAAA
113 AGCAGCTGAAACCGCTATGTCGGATGGTGGCAGCTCTCGTAATGCCCTGACGAAACTGGTGGAAAGCTGGCGCCTGGGTGGCGGTCAT
114 CACCATCACCATCACTAA

115 *AA-sequence*

116 MESVVLYPSPGMGHLISMVELGKLILKHHPSFSITIFIVTPPYNTGSTAPYLARVSSTIPSFTFHHLPTISLPLDSFSSPNHETLAFELLRLNNPNVHQ
117 ALVSISNNSSVRALIVDSFCTTALSVAAQLNIPCYYFFTSGACCLASYLYLPFIHQQTTKSFKDLNTHLHIPGLPPVPASDMAKPILDREDKAYEFFL
118 NMFIHLPRSAGIIVNTFEALEPRAVKTILDGLCVLDGPTSPIFCIGPLIATDDRSGGGGGGGGGGIPECLTWLESQPKRSVLFLCFGSLGLFSEEQLKE
119 IAVGLERSGQRFLWVVRSPPSKDPSRRFLAPPDPDLNSLLPDGFLDRTKERGLVVKSWAPQVAVLNHASVGGFVTHCGWNSVLEAVCAGVP
120 MVAWPLYAEQRFNRVVMVEELKLALPMEESEEGFITATEVEKRGRELMESEEGNTLRLRIMAMKKAAETAMSDGGSSRNALTKLVESWRLG
121 GGHHHHHH

122

123 **<UGT-GA*Qr***: gb|AHA54051.1| (Glycosyltransferase_GTB_type superfamily)

124 *DNA-sequence*

125 ATGGGCTCGGAAGCCCTGGTCCATGTGTTTCTGGTTAGCTTCCCGGGCCAGGGTCACGTTAACCCGCTGCTGCGTCTGGGTAAACGTCTG
126 GCGGCCAAAGGTCTGCTGGTTACCTTTAGTACGCCGGAATCCATTGGCAAACAGATGCGTAAAGCTTCCAATATCACCGATGAACCGGCA
127 CCGGTGGGTGAAGGTTTTATTCGCTTCGAATTTTTCGAAGACGGTTGGGATGAAGACGAACCGCGTCGCCAGGATCTGGACCAATATCT
128 GCCGCAGCTGGAACTGATTGGCAAAGATATTATCCCGAAAATGATCCGTAAAAACGCGGAAATGGGTCGCCCTGTGAGCTGTCTGATTA
129 ACAATCCGTTTATCCCGTGGGTGTCAGATGTTGCCGAATCGCTGGGTCTGCCGAGCGCAATGCTGTGGGTTCAGTCGTGCGCTTGTTTCT
130 GCGCGTATTACCATTATTACCACGGTCTGGTGCCGTTTCCGAGCGAAGCTGAACCGTTCATTGATATCCAGCTGCCGTGCATGCCGCTGCT
131 GAAATATGACGAAACCCCGTCTTTTCTGTATCCGACCACGCCGTACCCGTTCCTGCGTCGCGCAATTCTGGGCCAATACGGTAATCTGGAT
132 AAACCGTTTTGTATCCTGATGGACACCTTCCAGGAACTGGAACATGAAGTTATCGAATTCATGTCTAAAATCTGCCCGATCAAAACGGTC
133 GGCCCGCTGTTTAAAAACCCGAAAGCTCCGAATAGTGTGCGTGGTGATTTCATGAAAGCGGATGACTGTCTGGAATGGCTGGACTCTAA
134 ACCGCCGCAAAGTGTGGTTTATATCAGTTTTGGTTCCGTCGTGTACCTGACCCAGAAACAAGTTGATGAAATTGCGTTTGGCCTGCTGCA
135 GTCAGGTGTGTCGTTCCTGTGGGTTATGAAACCGCCGCATAAAGATGCTGGCCTGGAACTGCTGGTTCTGCCGGATGGTTTTCTGGAAAA
136 AGCGGGCGACAACGGTCGCGTTGTCCAGTGGTCTCCGCAGGAACAAGTGCTGGCTCATCCGAGTGTCGCGTGTTTCGTGACCCACTGCG
137 GCTGGAATTCAACCATGGAAAGCCTGACGTCTGGCATGCCGGTGGTTGCCTTTCCGCAATGGGGTGATCAGGTTACCGACGCAGTCTAT
138 CTGGTTGATGTCTTCAAAACGGGCGTGCGTATGTGTCGCGGTGAAGCCGAAAACCGTGTTATCACCCGCGATGAAGTCGAAAAATGCCT
139 GCTGGAAGCCACGGTCGGTCCGAAAGCAGTGGAAATGAAACAAAACGCCTCCAAATGGAAAGCAGCTGCGGAAGCCGCATTTTCAGAA
140 GGCGGTAGCTCTGATCGTAATATCCAGGCATTCGTCGACGAAGTGCGTGCACGCTCCGTCGCAATTACCGGCAAATCAACGGCTAATGG
141 TGTGGCCGCTGGATCTGGCCGAAAAAAGTGCAGAAATTAACGGCAAAGTTGACCTGGTCGAAACCAAAACGAATGGTAAAGTGGAACTG
142 GTTGAAGCGGGCGGTCATCACCATCACCATCACTAA

143 *AA-sequence*

144 MGSEALVHVFLVSFPGQGHVNPLLRLGKRLAAKGLLVTFSTPESIGKQMRKASNITDEPAPVGEGFIRFEFFEDGWDEDEPRRQDLDQYLPQLE
145 LIGKDIIPKMIRKNAEMGRPVSCLINNPFIPWVSDVAESLGLPSAMLWVQSCACFCAYYHYYHGLVPFPSEAEPFIDIQLPCMPLLKYDETPSFLY
146 PTTPYPFLRRAILGQYGNLDKPFCILMDTFQELEHEVIEFMSKICPIKTVGPLFKNPKAPNSVRGDFMKADDCLEWLDSKPPQSVVYISFGSVVYL
147 TQKQVDEIAFGLLQSGVSFLWVMKPPHKDAGLELLVLPDGFLEKAGDNGRVVQWSPQEQVLAHPSVACFVTHCGWNSTMESLTSGMPVVA
148 FPQWGDQVTDAVYLVDVFKTGVRMCRGEAENRVITRDEVEKCLLEATVGPKAVEMKQNASKWKAAAEAAFSEGGSSDRNIQAFVDEVRAR
149 SVAITGKSTANGVALDLAEKSAEINGKVDLVETKTNGKVELVEAGGHHHHHH

150

151

152

153

154

155 **S2: Primers used for GT subcloning**

ORF	Vector		Sequence (5' → 3')
GT fragments			
UGT-76G1*Sr*	pCXP34h	Forw	CCGTCGACCTCGAATTCGGAGGAAACAAAGATGGAAAATAAAACCGAAACCAC
		Rev	CATATGGTCGACCTGCAGTTAATGATGATGATGATGATGCGAGCCCAG
	pET21a	Forw	GAAGGAGATATACGGGATGGAAAATAAAACCGAAACCAC
		Rev	GTTAGCAGCCGGATCTTAATGATGATGATGATGATGCG
UGT-SA*Cr*	pCXP34h	Forw	CGAATTCGGAGGAAACAAAGATGGAGAAGAAACGTGGTCATG
		Rev	CCCATATGGTCGACCTGCAGTTAGTGATGGTGATGGTGATGACC
	pET21a	Forw	AAGAAGGAGATATACGGGATGGAGAAGAAACGTGGTC
		Rev	GTTAGCAGCCGGATCTTAGTGATGGTGATGGTGATG
GalT-F3*Ph*	pCXP34h	Forw	CCTCGAATTCGGAGGAAACAAAGATGTCTAACTATCATGTTGCG
		Rev	CCCATATGGTCGACCTGCAGTTAGTGATGGTGATGGTGATGACC
	pET21a	Forw	GAAGGAGATATACGGGATGTCTAACTATCATGTTGCG
		Rev	GTTAGCAGCCGGATCTTAGTGATGGTGATGGTGATG
UGT-MT*Vv*	pCXP34h	Forw	CGAATTCGGAGGAAACAAAGATGGAAAGTGTGGTTCTGTATC
		Rev	CCCATATGGTCGACCTGCAGTTAGTGATGGTGATGGTGATGACC
	pET21a	Forw	AAGAAGGAGATATACGGGATGGAAAGTGTGGTTCTGTATC
		Rev	GTTAGCAGCCGGATCTTAGTGATGGTGATGGTGATG
UGT-GA*Qr*	pCXP34h	Forw	CGAATTCGGAGGAAACAAAGATGGGCTCGGAAGCCCTGGTCC
		Rev	CCCATATGGTCGACCTGCAGTTAGTGATGGTGATGGTGATGACC
	pET21a	Forw	AAGAAGGAGATATACGGGATGGGCTCGGAAGCCCTG
		Rev	GTTAGCAGCCGGATCTTAGTGATGGTGATGGTGATG
Vector fragments			
pCXP34h_backbone	pCXP34h	Forw	CTGCAGGTCGACCATATGGG
		Rev	CTTTGTTTCCTCCGAATTCGAGGTC
pET21a_backbone	pET21a	Forw	CCCGTATATCTCCTTCTTAAAG
		Rev	GATCCGGCTGCTAACAAAG

156

157 **S3: Recombinant expression of GTs: yields in different expression systems**

Constitutive expression (pCXP34h)		
Enzyme	T (°C)	Yield (mg/L culture)
UGT-MT*Vv*	16	0.6
	37	0.6
UGT-GA*Qr*	16	0.3
	37	0.4
UGT-SA*Cr*	16	0.6
	37	1.0
UGT-76G1*Sr*	16	0.8
	37	1.0
GalT-F3*Ph*	16	0.2
	37	< 0.1

158

Inducible expression (pET21a)			
Enzyme	T (°C)	IPTG (mM)	Yield (mg/L culture)
UGT-MT*Vv*	16	0.2	< 0.1
	16	1.0	< 0.1
	37	0.2	< 0.1
	37	1.0	< 0.1
UGT-GA*Qr*	16	0.2	< 0.1
	16	1.0	< 0.1
	37	0.2	< 0.1
	37	1.0	< 0.1
UGT-SA*Cr*	16	0.2	0.6
	16	1.0	0.3
	37	0.2	3.0
	37	1.0	2.2
UGT-76G1*Sr*	16	0.2	7.1
	16	1.0	17.8 - 30.5
	37	0.2	3.2
	37	1.0	3.4
GalT-F3*Ph*	16	0.2	< 0.1
	16	1.0	0.2
	37	0.2	0.2
	37	1.0	0.3

159

6

S4: Kinetic characterization of UGT-76G1*Sr* (A-B) and UGT-SA*Cr* (C-D)
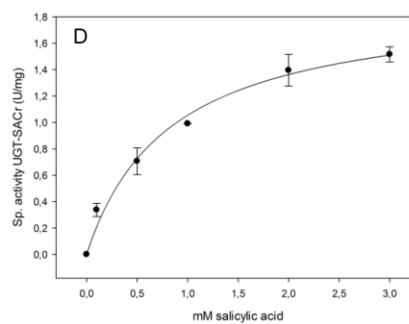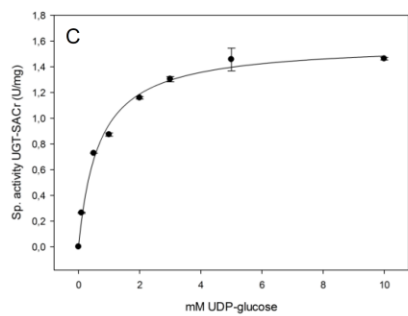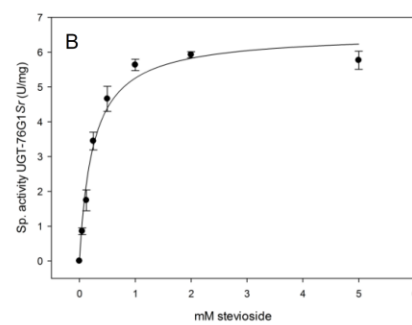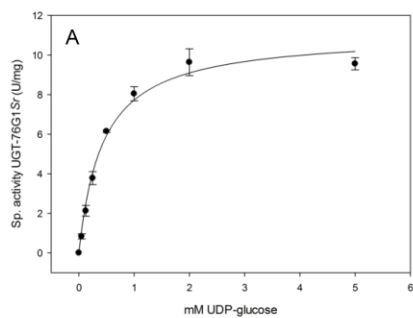
162   **S5: Chemical structure of acceptors used for screening enzyme promiscuity**

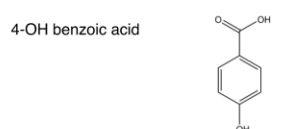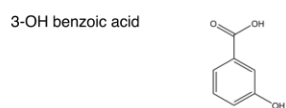|			
Ethanol	H_3C—OH	Phenol	
Hexanol		Catechol	
Heptanol			
Octanol		Resorcinol	
Decanol		Hydroquinone	
Cyclohexanol		Pyrogallol	
t-cyclohexanediol			
c-cyclohexanediol		Carvacrol	
Oleyl alcohol		Thymol	
Cinnamyl alcohol			
Linalool		Eugenol	
β-Citronellol		Orcinol	
Geraniol		Salicyl alcohol	

163

8

43

Salicylic acid		Octyl gallate	
Methyl salicylate		Lauryl gallate	
3-OH benzoic acid		1-phenylethanol	
4-OH benzoic acid		2-phenylethanol	
Gallic acid		Vanillin	
Methyl gallate		Vanillyl alcohol	
Ethyl gallate		Vanillic acid	
Propyl gallate		*para*-nitrophenol	
		ortho-nitrophenol	

164
165

Anisyl alcohol		4-phenoxyphenol	
Benzyl alcohol		Polydatin	
Catechin		Salicin	
Epicatechin		p-nitrophenyl α-glucopyranoside	
Hesperetin		p-nitrophenyl α-galactopyranoside	
Luteolin		p-nitrophenyl β-fucopyranoside	
Quercetin		Stevioside	
Resveratrol		Arbutin	
Curcumin		Esculin	

166
167
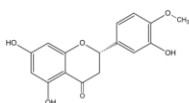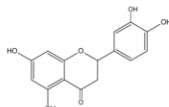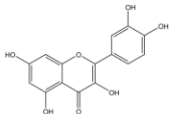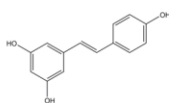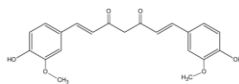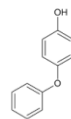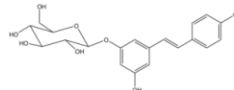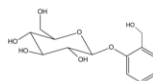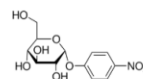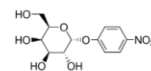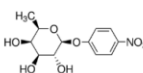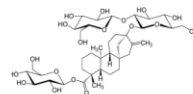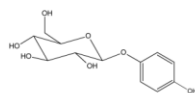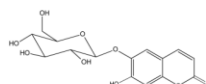
S6: Kinetics of the production of curcumin mono- and diglucoside

169 **(A)** Schematic representation of curcumin conversion into consequently mono- and diglucoside, as
170 well as the UDP-Glc regeneration system. **(B)** The conversion of curcumin and the production of its
171 glucosides is presented in function of time. Reaction conditions: 0.5 mM curcumin, 5 mM UDP-Glc,
172 100 mM sucrose, 2 mg/mL UGT-76G1*Sr* in 10% DMSO and 100 mM MOPS buffer pH 7.0 at 30 °C.
173 Closed circles: curcumin; Open circles: curcumin b-monoglucoside; Closed triangles: curcumin b-
174 diglucoside

175 **A**

176

B

11
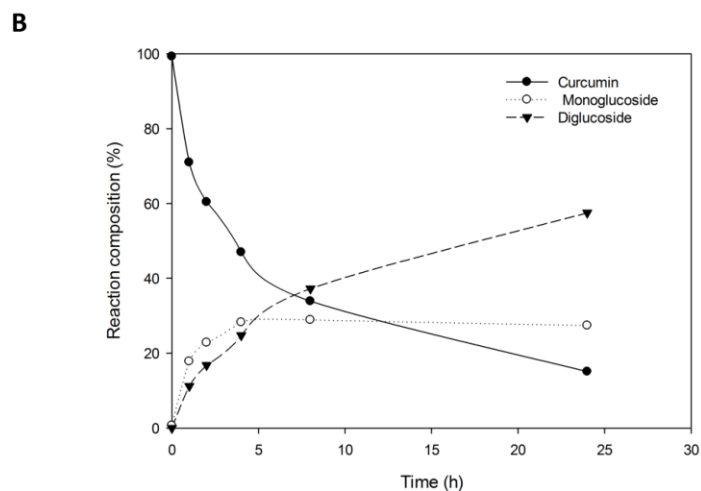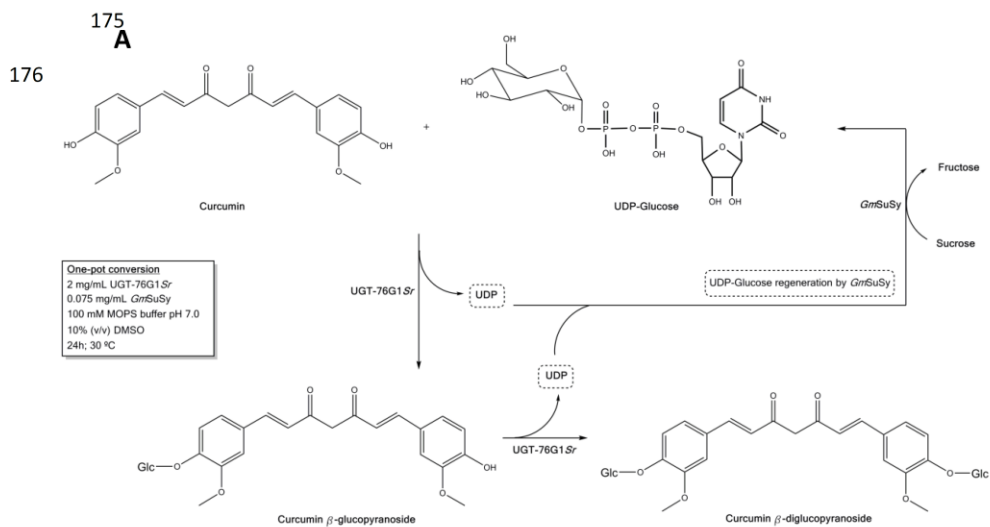
S7: ^{1}H-NMR analysis of curcumin mono- and diglucoside

178

179

12
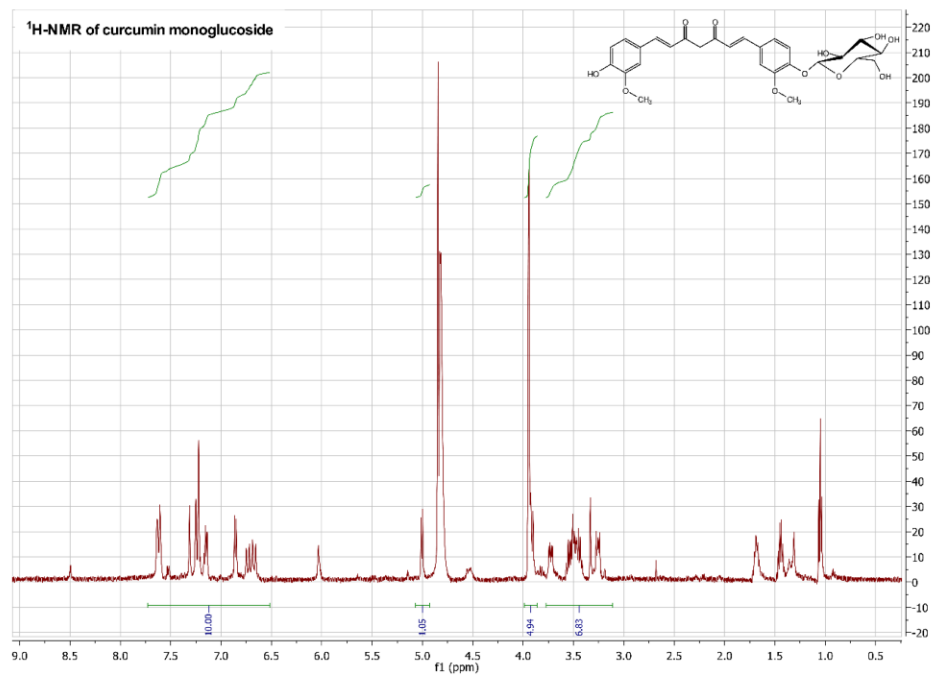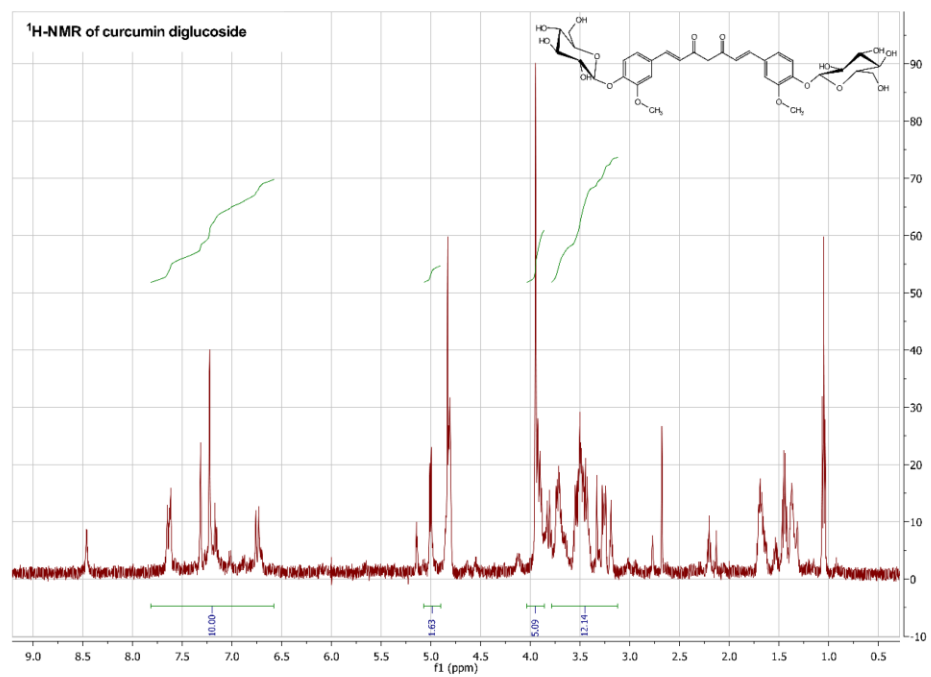
Chapter 3:
Glycosyltransferase cascades for natural product glycosylation: Use of plant instead of bacterial sucrose synthases improves the UDP-glucose recycling from sucrose and UDP

Research Article

Glycosyltransferase cascades for natural product glycosylation: Use of plant instead of bacterial sucrose synthases improves the UDP-glucose recycling from sucrose and UDP

Alexander Gutmann[1], Alexander Lepak[1], Margo Diricks[2], Tom Desmet[2] and Bernd Nidetzky[1,3]

[1] Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, NAWI Graz, Graz, Austria
[2] Centre for Industrial Biotechnology and Biocatalysis, Department of Biochemical and Microbial Technology, Ghent University, Ghent, Belgium
[3] Austrian Centre of Industrial Biotechnology, Graz, Austria

Natural product glycosylations by Leloir glycosyltransferases (GTs) require expensive nucleotide-activated sugars as substrates. Sucrose synthase (SuSy) converts sucrose and uridine 5'-diphosphate (UDP) into UDP-glucose. Coupling of SuSy and GT reactions in one-pot cascade transformations creates a UDP cycle, which regenerates the UDP-glucose continuously and so makes it an expedient donor for glucoside production. Here we compare SuSys with divergent kinetic characteristics for UDP-glucose recycling in the synthesis of the natural *C*-glucoside nothofagin. Development of a fast reversed-phase ion-pairing HPLC method, quantifying all relevant reactants from the coupled conversion in a single run, was key to dissect the main factors of recycling efficiency. Limitations due to high K_M, both for UDP and sucrose, were revealed for the bacterial SuSy from *Acidithiobacillus caldus*. The L637M-T640V double mutant of this SuSy with a 60-fold reduced K_M for UDP substantially improved UDP-glucose recycling. The SuSy from *Glycine max* (soybean) was nevertheless the most active enzyme at the UDP (≤0.5 mM) and sucrose (≤1 M) concentrations used. It was also unexpectedly stable at up to 50°C where spontaneous decomposition of UDP-glucose started to become problematic. The herein gained in-depth understanding of requirements for UDP-glucose regeneration supports development of efficient GT-SuSy cascades.

Supporting information
available online

1 Introduction

Because of their unmatched regio- and stereoselectivity, nucleoside diphosphate (NDP)-sugar-dependent Leloir

Correspondence: Prof. Bernd Nidetzky, Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Petersgasse 12, 8010 Graz, Austria
E-mail: bernd.nidetzky@tugraz.at

Abbreviations: *AcSuSy,* sucrose synthase from *Acidithiobacillus caldus*; ***GmSuSy,*** sucrose synthase from *Glycine max*; **GT,** Leloir glycosyltransferase; K_M, Michaelis-Menten constant; **NDP,** nucleoside diphosphate; ***OsCGT,*** *C*-glucosyltransferase from *Oryza sativa*; **SuSy,** sucrose synthase; **TBAB,** tetra-*n*-butylammonium bromide; **UDP,** uridine 5'-diphosphate; **UMP,** uridine 5'-monophosphate

glycosyltransferases (GTs) are the biocatalysts of choice for high-precision synthesis of complex oligosaccharide structures and glycosylated natural products [1–3]. Common problem of GT-catalyzed glycosylations is that the NDP-sugar substrates are usually too costly to be used as reagents in synthetic applications [4, 5]. Efficient in situ regeneration of NDP-sugars is therefore a clear requirement for the economic feasibility of glycosylation processes at production scale.

A widely used approach of NDP-sugar recycling is based on the use of sugar-1-phosphate nucleotidyltransferases in a multi-step enzyme cascade transformation, as shown in Scheme 1A [4, 6, 7]. While generally efficient and also relatively flexible in its application to different NDP-sugars, the cascade requires three or even more enzymes. Besides costs, therefore, reaction optimization can

Scheme 1. Glycosyltransferase cascades applied to biocatalytic synthesis are shown. Glycosylations can be combined with NDP-sugar regeneration by (**A**) sugar-1-phosphate nucleotidyltransferases or (**B**) the reverse reaction of a GT. In (**A**) the sugar-1-phosphate is usually prepared in situ by phosphorylation of the sugar from NTP using a suitable kinase. (**C**) 3'-C-β-D-glucosylation of phloretin was coupled with in situ UDP-glucose recycling by SuSy.

become a complex task. A conceptually simpler, single-step approach of NDP-sugar regeneration is to use the reverse reaction of a GT: a glycoside substrate is reacted with NDP to give the desired NDP-sugar (Scheme 1B) [8–10]. The "GT in reverse" approach is somewhat limited by the availability of suitable glycoside donors. Reverse glycosylations are generally disfavored thermodynamically unless driven to NDP-sugar formation by a high degree of donor glycoside activation [10, 11]. Both routes of NDP-sugar regeneration are justified and even have complementary use. However, in cases where a NDP-glucose substrate is required, sucrose and sucrose synthase (EC 2.4.1.13; SuSy) are particularly useful [12, 13]. Among the naturally abundant disaccharides, sucrose is exceptional in that it possesses a glycosidic linkage unusually rich in energy [8, 13]. Sucrose therefore represents an ideal donor substrate to promote the NDP-dependent conversion into NDP-glucose.

SuSy catalyzes the reversible glucosyl transfer reaction, sucrose + NDP ↔ NDP-glucose + fructose [12, 13]. The specificity of SuSy for NDP is relatively broad, thus enabling synthesis of the corresponding NDP-glucoses [12, 14]. However, because the majority of the known glucosyltransferases requires UDP-glucose, this was also the clear focus of past efforts of using SuSy for donor substrate recycling [15–20]. To expand the scope of SuSy-GT cascade reactions beyond glucosyl transfer, additional NDP-glucose-modifying enzymes were added to enable regeneration of UDP-galactose [21–25], UDP-glucuronic acid [26] and dTDP-rhamnose [27].

Considering the importance of UDP-glucose recycling, detailed evidence on the SuSy reaction in the context of GT-SuSy cascade transformations is required. Systematic analysis is necessary to make sure that the potential of SuSy is properly utilized in these transformations. Up to now, the influence of single parameters like

the UDP concentration [17], the pH [18] or the ratio of SuSy to GT [15] were investigated in certain case studies. However, comprehensive study of possible bottlenecks in GT-SuSy cascade reactions is still needed. We envisioned that not only the optimization of reaction conditions but also the enzyme selection would be crucial. Despite large differences in substrate affinity, specific activity and thermostability of the SuSys reported so far, direct comparison of enzymes under conditions of sugar nucleotide regeneration is lacking. A general constraint of GT-SuSy cascade reactions for glucosylation by UDP-glucose is that, in order to reduce the overall process cost, the number of UDP-glucose regeneration cycles (transferred glucose moieties/UDP applied) should be maximized [28]. This is achieved by reducing the concentration of UDP, which is by far the most expensive substance added to the reaction mixture. Additionally, lowering the UDP input into the reaction also prevents pronounced product inhibition by UDP, which is frequently observed for GTs [16, 29]. The sucrose concentration is generally less important, for sucrose is expedient and usually shows no inhibition. Based on these considerations, therefore, one would suggest that a SuSy with a low Michaelis-Menten constant (K_M) for UDP should be applied for UDP-glucose recycling.

The known SuSys belong to two major evolutionary lineages, one in plants and the other in bacteria [13, 30]. Bacterial SuSys appear to prefer ADP and show low affinity for UDP. SuSys from plants have high affinity for UDP. Previously, with some notable exceptions [24, 25, 31], only SuSys from plant origin were applied in GT-SuSy cascade reactions. Choice of the plant SuSys was probably not so much inspired by considerations of kinetic suitability but by enzyme availability at the time of study. The bacterial SuSys represent a discovery of a more recent date [30, 32]. Thermostability, specific activity and recombinant ex-

pression levels appear to be significantly higher in some of the bacterial SuSys as compared to the plant enzymes [30, 32, 33]. Therefore, this raised considerable interest in the industrial application of the bacterial SuSys. An ideal SuSy would however combine the robustness of bacterial enzymes with the high UDP affinity of plant enzymes.

We have recently shown that point mutations in the nucleotide binding pocket of a bacterial SuSy from *Acidithiobacillus caldus* (*Ac*SuSy) could strongly increase the affinity for UDP without negative effects on enzyme activity or expression [31]. The L637M-T640V double mutant, having L<u>M</u>DK<u>V</u>VA as nucleotide binding motif, exhibited a 60-fold decreased K_M for UDP which in preliminary investigations resulted in significantly improved UDP-glucose recycling [31]. In this study we present an in-depth comparison of UDP-glucose regeneration by SuSy from soybean (*Glycine max*, *Gm*SuSy), *Ac*SuSy wild-type and the *Ac*SuSy variant [17, 30, 31]. Reaction of each SuSy was coupled to the well-established and synthetically useful 3'-*C*-β-D-glucosylation of the dihydrochalcone phloretin by a *C*-glucosyltransferase (CGT) from rice (*Oryza sativa*, *Os*CGT) [17, 34] (Scheme 1C) [17, 34]. The resulting product nothofagin is a major health promoting antioxidant in rooibos tea, for which various interesting applications as nutraceutical exist [35–37]. To fully understand the performance of GT-SuSy cascade conversions, it was important to analyze not only substrate and product (here: phloretin and nothofagin) but also the components of the regeneration cycle, UDP and UDP-glucose. A reversed-phase ion-pairing HPLC assay was established to monitor these compounds efficiently in a single run. Insights thus obtained were not only vital for optimization of the CGT-SuSy cascade but also revealed purely chemical limitations due to UDP-glucose instability that could otherwise have gone unnoticed.

2 Materials and methods

2.1 Materials

Unless otherwise indicated, all chemicals were from Sigma-Aldrich (Vienna, Austria) in the highest purity available. Phloretin (98%), UDP (≥97%) and UDP-glucose (>90%) were purchased from Carbosynth (Berkshire, UK) and tetra-*n*-butylammonium bromide (TBAB) was obtained from Carl Roth (Karlsruhe, Germany).

2.2 HPLC assays

Reversed-phase HPLC assays were used to follow enzymatic conversions. Phloretin and nothofagin as well as UMP, UDP and UDP-glucose were quantified. Aliquots of 25 µL were withdrawn from reactions and enzymes were inactivated by mixing with 25 µL water and 50 µL acetonitrile. By centrifugation at 13 200 rpm for 15 min pre-

cipitated proteins were removed and 5 µL of the supernatant were applied to HPLC analysis. If the concentration of a nucleotide in the reaction mixture was below 5 µM, the injection volume was increased to 20 µL to improve quantification. All assays were performed on an Agilent 1200 HPLC system equipped with a Kinetex™ C18 column (5 µm, 100 Å, 50 x 4.6 mm; Phenomenex, Aschaffenburg, Germany). A constant flow rate of 2 mL/min was applied at 35°C. Using an Agilent G1365D multiple-wavelength detector dihydrochalcones and (sugar) nucleotides were quantified by UV-detection at 288 and 262 nm, respectively. For quantification of phloretin and nothofagin water and acetonitrile, each containing 0.1% formic acid, were used as mobile phase A and B, respectively. Separation was achieved by following gradient: 4 min 5–35% B; 0.01 min 35–80% B; 0.99 min 80% B; 0.01 min 80–5% B; 1.49 min 5% B. UMP, UDP and UDP-glucose were analyzed by reversed-phase ion-pairing HPLC during 1.5 min long isocratic runs. 87.5% mobile phase A (20 mM potassium phosphate buffer, pH 5.9 containing 40 mM TBAB) and 12.5% mobile phase B (acetonitrile) were used. To analyze besides (sugar) nucleotides also phloretin and nothofagin in a single run, a gradient was added to the ion-pairing protocol: 1 min 10% B; 4 min 10–50% B; 0.01 min 50–80% B; 0.99 min 80% B; 0.01 min 80–10% B; 1.49 min 10% B. To determine linear dynamic range and limits of quantitation and detection, serial dilutions of mixtures containing either UMP, UDP and UDP-glucose or nothofagin and phloretin were analyzed. Nucleotide concentrations ranged from 10 nM to 10 mM and dihydrochalcones were applied from 50 nM to 50 mM.

2.3 Enzyme preparation

E. coli BL21-Gold (DE3) strains for expression of *Os*CGT (GenBank: CAQ77160.1) [38, 39], *Gm*SuSy (GenBank: AAC39323.1) [17], *Ac*SuSy (GenBank: AIA55343.1) [30] and the *Ac*SuSy L637M-T640V double mutant (alternatively termed L<u>M</u>DK<u>V</u>VA) [31] were described elsewhere. Enzyme expression and purification by Strep- (*Os*CGT, *Gm*SuSy) or His-tag (*Ac*SuSy, *Ac*SuSy variant) affinity chromatography were performed as previously reported [14, 39].

2.4 Enzymatic conversions

All enzymatic conversions were performed in 1.5 mL reaction tubes without agitation. Reactions were buffered to pH 7.5 by 50 mM HEPES containing 50 mM KCl, 13 mM MgCl$_2$ and 1.3 mg/mL BSA. Because pH changes during conversions were below 0.1 pH units no pH correction was required. The influence of the sucrose concentration (100–1000 mM) on SuSy activity was studied by converting 1 mM UDP with 1 µg/mL SuSy at 30°C. All *Os*CGT-SuSy cascade conversions contained 17.5 mM β-cyclodextrin (β-CD) to avoid precipitation of 5 mM phloretin.

As glucose source 1 M sucrose was added. Unless mentioned otherwise, 0.5 mM UDP was used and the reaction temperature was 30°C. Reaction mixtures were preheated before starting conversions by addition of 30 µg/mL *Os*CGT and 10 µg/mL SuSy. The effect of the UDP concentration on the coupled conversion was studied in a range of 0.1–5 mM and the temperature dependence was studied between 30 and 60°C. Reaction rates and analyte concentrations varied in multiple repeats of enzymatic conversions by less than 5%.

3 Results and discussion

3.1 Fast comprehensive reaction analysis for facilitated optimization

Synthesis of nothofagin according to the cascade reaction in Scheme 1C was analyzed. Although conversion of phloretin into nothofagin was the clear target, the cascade's performance could not be fully understood unless the dynamics of the regeneration cycle was also resolved. Recent work of Elling and coworkers shows the importance of detailed reactant fingerprinting, via monitoring substrates, intermediates and products of the reaction simultaneously, for optimization of multi-enzyme cascades in NDP-sugar synthesis [40]. Ideally, the obtained information should also allow for identification and quantification of side-reactions. Here, changes in the concentrations of UDP and UDP-glucose therefore had to be measured in dependence of conversion or time. Reversed-phase HPLC with a water/acetonitrile gradient was previously used for the quantification of phloretin and nothofagin (Supporting information, Fig. S1A) [17]. Although no enzymatic side reactions were expected, decomposition of UDP-glucose to UMP and glucose-1,2-cyclic phosphate was foreseen as potentially problematic (Supporting information, Scheme S1). Therefore, UMP besides UDP and UDP-glucose had to be quantified in addition to phloretin and nothofagin. We noted that (sugar) nucleotides like UMP, UDP and UDP-glucose could also be analyzed by reversed-phase HPLC if an TBAB based ion-pairing protocol was used [11]. By applying an isocratic flow at 12.5% acetonitrile (sugar) nucleotides were separated in just 1.5 min (Supporting information, Fig. S1B).

However, the use of two separate HPLC assays with different buffers – one for (sugar) nucleotide analysis, the other for dihydrochalcone analysis – was labor-intensive. More importantly even, under the conditions used for nucleotide analysis, nothofagin and phloretin did not elute and would eventually clog the column. To solve the problem we developed a novel reversed-phase ion-pairing HPLC method to analyze the relevant compounds in a single run (Fig. 1). In essence, the isocratic ion-pairing protocol had to be extended by an optimized acetonitrile gradient. Using 40 mM TBAB, nucleotides were resolved
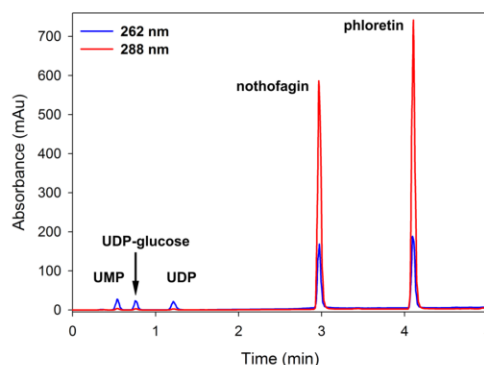
Figure 1. Reversed-phase ion-pairing HPLC with an acetonitrile gradient was used to quantify nucleotides and dihydrochalcones in a single HPLC run. Compounds were monitored by UV detection.

within 1 min at 10% acetonitrile before starting a 4-min long gradient to 50% acetonitrile to analyze the dihydrochalcones. Including a wash at 80% acetonitrile and re-equilibration of the column, the overall run time was only 7.5 min. The protocol enabled a simultaneous quantification of UMP, UDP, UDP-glucose, nothofagin and phloretin, as shown in Fig. 1. For all compounds a linear dynamic range was obtained over three orders of magnitude from around 2 µM to 3 mM (Supporting information, Fig. S2A and S2B). The relative standard deviations from multiple injections were typically below 0.3%. Because reaction mixtures contained up to 5 mM phloretin and nothofagin, samples were four-fold diluted before HPLC analysis. It was critical to be nevertheless able to quantify small amounts of nucleotides, which were present in only low catalytic concentrations. Without dilution, the limit of quantitation of nucleotides was slightly below 1 µM, and the corresponding limit of detection was around 0.2 µM (Supporting information, Fig. S2C). If required, the same sensitivity was also obtainable with diluted reaction mixtures upon increasing the injection volume from 5 to 20 µL. The high throughput of 192 samples per day provided convenient access to all required data for reaction analysis and optimization.

We envisioned that the herein developed assay would not only be applicable to optimize the *Os*CGT-SuSy cascade reaction, but could be a generally useful tool to study conversion of hydrophobic acceptors by GTs.

3.2 Establishing the cascade of SuSy and CGT

To facilitate reaction control and to preclude any side-reactions from enzyme impurities, all conversions were performed with affinity purified enzymes from *E. coli* overexpression cultures (Supporting information, Fig. S3). Before comparing UDP-glucose recycling by the different

Biotechnology
Journal
www.biotechnology-journal.com

Biotechnol. J. 2017, 12, 1600557

ADVANCED
SCIENCE NEWS
www.advancedsciencenews.com
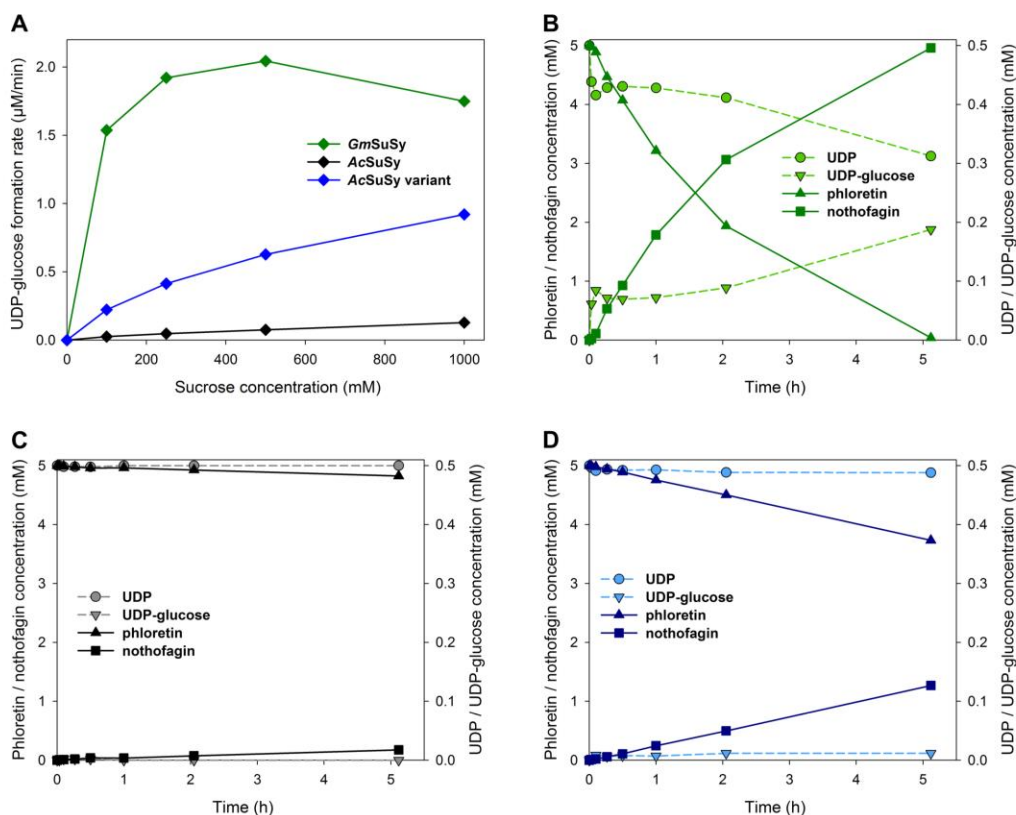
Figure 2. Kinetic evaluation of the different SuSys used is shown. (**A**) The sucrose dependence of the initial UDP-glucose formation rate in SuSy reactions is shown. Reaction conditions: 1 mM UDP, 1 µg/mL SuSy, 30°C, pH 7.5; standard deviation <0.1 µM/min. (**B–D**) Time courses of phloretin conversion by *Os*CGT-SuSy cascade reactions are shown. *Gm*SuSy (**B**), *Ac*SuSy (**C**) or *Ac*SuSy variant (**D**) was used. Reaction conditions: 5 mM phloretin, 0.5 mM UDP, 1 M sucrose, 30 µg/mL *Os*CGT, 10 µg/mL SuSy, 30°C, pH 7.5; standard deviation <5%.

SuSys we had to define useful reaction conditions for the cascade conversions. Unlike nothofagin synthesis by *Os*CGT (pH optimum ≈8.5), UDP-glucose formation by SuSy was favored under acidic conditions (≈pH 5–6) [11, 17]. We previously recognized that both enzymatic conversions were well combinable when buffered to pH 7.5 with 50 mM HEPES containing 50 mM KCl and 13 mM $MgCl_2$ [17]. Because of the limited thermostability of *Gm*SuSy, a reaction temperature of 30°C was selected [11]. The concentrations of phloretin and UDP were set to 5 and 0.5 mM, respectively, so that 10 regeneration cycles would be required for full conversion of phloretin. To compensate for the limited water solubility of phloretin, 17.5 mM β-CD were added to form a soluble phloretin/β-CD inclusion complex [34].

To identify the ideal sucrose concentration for the cascade reaction we measured the rates of UDP-glucose formation by *Gm*SuSy, *Ac*SuSy and *Ac*SuSy variant at various sucrose concentrations (Fig. 2A). The rates with *Ac*SuSy wildtype and variant gradually increased at higher sucrose concentration without reaching substrate saturation within the tested range of 0.1–1 M sucrose. However, the *Ac*SuSy variant was roughly eight times more active than the wildtype enzyme. The improved conversion of 1 mM UDP by the SuSy variant was caused by a 60-fold reduced K_M for UDP [31]. SuSy reaction kinetics involve a ternary complex between enzyme, sucrose and NDP which must be formed prior to any product release [13, 41]. This implies that the K_M-value for sucrose typically depends also on the binding affinity for the NDP substrate. Interestingly, therefore, despite changes in the

K_M for UDP caused by AcSuSy engineering, the K_M for sucrose was not affected in the enzyme variant compared to AcSuSy wild-type. It applies also for other bacterial SuSys from *Anabaena* sp. and *Nitrosomonas europaea* that unusually high K_M-values for sucrose (≈300 mM) were measured in presence of UDP [30, 42]. Wang and co-workers recognized that the lower K_M-values of plant SuSys were advantageous for UDP-glucose recycling [43]. It was therefore not a complete surprise that GmSuSy handled low sucrose concentrations much better than the bacterial enzymes. Using GmSuSy, 75% of the maximum rate was already attained at 0.1 M sucrose and GmSuSy was 59 and 6.9 times more active than AcSuSy wildtype and variant, respectively. At higher sucrose concentrations the advantage of GmSuSy diminished gradually and at 1 M sucrose GmSuSy was only 1.9-fold more active than the AcSuSy variant. To offer conditions where all SuSys showed appreciable activities we used 1 M sucrose in the cascade conversions reported later.

After having defined key operational parameters we compared cascade reactions with the different SuSys. To emphasize on the effect of UDP-glucose regeneration, OsCGT (30 µg/mL) was applied in three-fold excess over SuSy (10 µg/mL). The resulting time courses demonstrated that the efficiency of UDP-glucose recycling controlled the nothofagin formation rate (Fig. 2B–D). Only with GmSuSy a sufficient amount of UDP-glucose was formed so that the UDP-glucose consumption by OsCGT was fueled properly (Fig. 2B). Within 2 min a steady-state concentration of 60–90 µM UDP-glucose was reached. Only after 5 h when phloretin was fully converted to nothofagin and UDP-glucose consumption ceased, the UDP-glucose concentration increased further to around 200 µM. Using AcSuSy, in contrast, the nothofagin synthesis was severely limited by the lack of UDP-glucose formed (Fig. 2C). Throughout the course of the reaction less than 2 µM UDP-glucose were present. Although less pronounced, the same restriction was observed with the AcSuSy variant (Fig. 2D). As a consequence of these limitations, the final nothofagin concentration after 5 h was 4.96 mM (99.2% conversion) with GmSuSy but only 0.18 mM (3.5% conversion) and 1.27 mM (25.4% conversion) with AcSuSy wildtype and variant, respectively. Also the initial nothofagin formation rates were affected in a similar manner. Although the replacement of AcSuSy wild-type with the variant accelerated the nothofagin formation almost by an order of magnitude from 0.5 to 4.1 µM/min, conversion with GmSuSy (34 µM/min) was still nearly 10-fold faster. Therefore, only GmSuSy was able to recycle UDP-glucose efficiently under the applied conditions. However, the performance of the bacterial SuSys should be boosted at elevated UDP concentrations and high reaction temperatures. To provide a comprehensive comparison, therefore, we also studied time courses of biocatalytic reactions under these conditions.

3.3 *Gm*SuSy is most efficient in regenerating UDP-glucose at low UDP concentrations

In view of the high K_M of AcSuSy for UDP (7.8 mM) [30], we expected a performance boost for this enzyme at elevated UDP concentrations. Therefore, we recorded time courses of OsCGT-SuSy cascade reactions at 0.1 to 5 mM UDP and compared the initial nothofagin formation rates under these conditions (Fig. 3). Using AcSuSy, as anticipated, the phloretin conversion increased almost linearly with the UDP concentration. A strong influence of UDP was also observed with the AcSuSy variant. However, because of the decreased K_M clear signs of saturation were recognized above 2 mM. Only with GmSuSy very little dependence on the UDP concentration was seen and already at 1 mM UDP the maximum conversion rate was reached. Even at 0.1 mM UDP the nothofagin formation was only slowed down by 32%. Therefore the corresponding production rates with AcSuSy wild type and variant were 91 and 14-times slower, respectively. Using the AcSuSy variant, a 50-fold higher UDP concentration of 5 mM was required to match the rate of the GmSuSy reaction. The results can be used to calculate a "system K_M" for UDP which describes the apparent dependence of the GT-SuSy cascade on the UDP concentration. The system K_M is a practically useful parameter conducive to reaction optimization via maximizing the nothofagin formation rate, the number of regeneration cycles or a compromise of both. For GmSuSy a value of 46 (± 8) µM was determined. This can be compared to a system K_M of 1.3 (± 0.2) mM and 20 (± 5) mM for AcSuSy in variant and wild-type form, respectively. A high system K_M can be remedied partially by raising the UDP concentration. However, we considered for several reasons that UDP levels above 0.5 mM were not suited for synthetic applica-
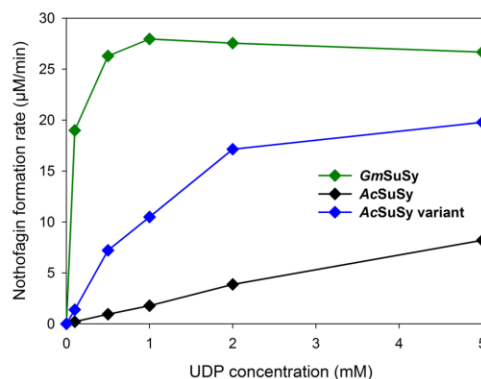

Figure 3. The UDP dependence of the initial nothofagin formation rate in OsCGT-SuSy cascade reactions is shown. Reaction conditions: 5 mM phloretin, 1 M sucrose, 30 µg/mL OsCGT, 10 µg/mL SuSy, 30°C, pH 7.5; standard deviation <2 µM/min.

Biotechnology
Journal
www.biotechnology-journal.com

Biotechnol. J. 2017, 12, 1600557

ADVANCED
SCIENCE NEWS
www.advancedsciencenews.com

tions. The number of regeneration cycles typically defines the efficiency of cofactor/cosubstrate recycling systems [28]. Because the maximum number of transferable glucosyl moieties is restricted by the applicable acceptor concentration, reduction of the UDP concentration is crucial. Therefore, UDP-glucose was regenerated up to 99 and 132 times when the UDP concentration was restricted to 0.1 and 0.5 mM, respectively [19, 34]. High concentrations of the rather costly UDP not only have a negative impact on overall process economy, but product inhibition of GTs by UDP can also strongly reduce reaction rates [16, 29]. To achieve optimal performance the system K_M for UDP has to be well below the respective inhibition constant of the GT module. In the OsCGT-GmSuSy cascade this criterion was clearly fulfilled. By raising UDP from the optimum at 1 mM to 5 mM (≈100-fold above the system K_M) the nothofagin formation rate dropped by less than 5% (Fig. 3).

3.4 *Gm*SuSy shows unexpectedly efficient performance at elevated temperatures

Although *Gm*SuSy clearly performed best among the three SuSys when a temperature of 30°C was chosen, there remained the concern that the plant SuSy might not offer the required thermostability for industrial carbohydrate conversions. These processes are typically performed at elevated temperatures to avoid microbial contaminations [44]. The impact of differences in enzyme thermostability was tested by running the CGT-SuSy cascade reactions between 30 and 60°C (Fig. 4). *Gm*SuSy handled elevated temperatures surprisingly well and nothofagin was formed fastest at 50 and 60°C (Fig. 4A). At both temperatures phloretin was already almost completely converted to nothofagin after around 2 h. At 30 and 40°C similar degrees of conversions were only reached after a substantially longer reaction time of 9 h. The nota-

bly fast conversion at elevated temperatures gave no insights into the long-term stability of *Gm*SuSy under these conditions, but as applied in these reactions, the enzyme was clearly stabilized. We previously showed that at pH 5.5 and 45°C the initial reaction rate of *Gm*SuSy dropped already by more than 50% [11]. At 60°C the enzyme was virtually inactive. Here, not only the pH was increased to 7.5 but also the concentration of the potentially stabilizing sucrose was raised from 0.5 to 1 M [45, 46]. Furthermore, addition of β-CD had a moderately positive effect on both stability and activity of *Gm*SuSy [34]. It was probably a combination of these effects that *Gm*SuSy was sufficiently stabilized to maintain a high UDP-glucose regeneration rate even at 60°C, for at least 2 h.

Cascade reactions with the thermostable *Ac*SuSys could profit even more from elevated reaction temperatures. In line with the previous results, however, the nothofagin formation with *Ac*SuSy was considerably slower than with *Gm*SuSy and it improved significantly at increased temperatures (Fig. 4B). Up to 50°C, nothofagin production was constant over the observed time span of 13 h. At 60°C, however, the phloretin conversion rate gradually decreased over time and the reaction almost completely stopped after 6.5 h. Therefore, the highest final nothofagin concentration was obtained at 50°C and it was just 1.1 mM. Examined under identical conditions, the conversion with *Gm*SuSy was around 75-fold faster. Coupled conversions with the *Ac*SuSy variant were generally faster than the corresponding conversions with the wild-type enzyme (Fig. 4C). However, at 60°C, the nothofagin formation again slowed down during the reaction and fully stopped after 9 h. Once more, 50°C was the optimum temperature and full phloretin conversion was achieved within 6.5 h. Although the CGT-*Gm*SuSy cascade was still more than three times faster, the strongly improved activity of the variant at 50°C could be considered as technologically useful.
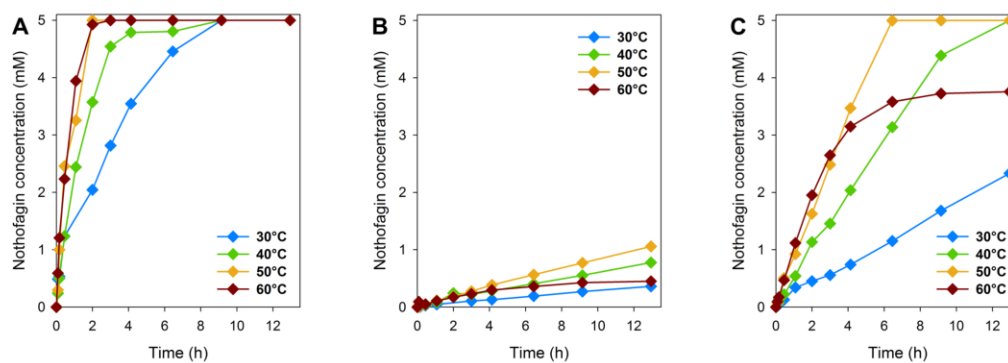

Figure 4. Time-courses of nothofagin formation by *Os*CGT-SuSy cascade conversions are shown for various temperatures. *Gm*SuSy (**A**), *Ac*SuSy (**B**) or *Ac*SuSy variant (**C**) was used. Reaction conditions: 5 mM phloretin, 0.5 mM UDP, 1 M sucrose, 30 µg/mL *Os*CGT, 10 µg/mL SuSy, pH 7.5; standard deviation <0.3 mM.

By analyzing only conversion of phloretin to nothofagin it was difficult to rationalize if it was the CGT or the SuSy reaction that restricted the useable reaction temperature to 50°C. However, analysis of UDP and UDP-glucose concentrations revealed clearly that it was inactivation of the CGT that had caused the insufficient performance at 60°C. In the coupled CGT-AcSuSy reaction the UDP-glucose was constantly below 10 µM during the first 3 h of reaction, but accumulated afterwards to up to 44 µM when the nothofagin formation gradually decreased. Also with the AcSuSy variant UDP-glucose was around 10 µM during the first 30 min, but increased to 125 µM at 6.5 h. Therefore, transition from an initially SuSy-limited reaction to a CGT-restricted conversion occurred between 3 h and 6.5 h of reaction. In conclusion, the comparably low thermostability of GmSuSy had no negative impact on the CGT-SuSy cascade reaction because OsCGT was more sensitive to elevated temperatures than GmSuSy was. This led to the counterintuitive result that full conversion of phloretin at 60°C was only possible with the least thermostable SuSy. Because of the superior UDP-glucose recycling performance of GmSuSy, the synthesis of nothofagin was completed before the CGT was inactivated.

Careful nucleotide analysis showed that degradation of instable UDP-glucose became also problematic at elevated temperatures. Decomposition of UDP-glucose to glucose-1,2-cyclic phosphate and UMP is favored by alkaline pH, at elevated temperatures and when divalent cations like Mg^{2+} are present (Supporting information, Scheme S1) [11, 47]. Despite using a slightly alkaline pH of 7.5 and 13 mM Mg^{2+} in all conversions, UDP-glucose degradation only became critical when the reaction temperature was also increased. Fig. S4 in the Supporting information shows the concentrations of nucleotides and nothofagin after 3 h of coupled conversions at 60°C. Formation of UMP strongly depended on accumulation of UDP-glucose and thereby also on the SuSy activity. Using GmSuSy the concentration of UMP (180 µM) surpassed already after 3 h those of UDP (170 µM) and UDP-glucose (150 µM). Also with the AcSuSy variant relatively high levels of UMP (70 µM) and UDP-glucose (90 µM) were observed. Only with barely active AcSuSy wild-type hardly any UDP-glucose (8 µM) and UMP (14 µM) were formed. Because of the strong dependence of nothofagin formation on UDP-glucose accumulation, sugar donor decomposition also correlated well with nothofagin levels. Even after prolonged incubation of 9 h only 17% of initially applied UDP was degraded to UMP when AcSuSy wild-type was used. However, with GmSuSy and the AcSuSy variant UMP accumulated to 73 and 67%, respectively. Because UMP cannot be recycled by SuSy, the effective concentrations of UDP(-glucose) were reduced from 500 to roughly 150 µM. Such a strong reduction of available UDP should decrease nothofagin formation with the AcSuSy variant by more than three-fold (Fig. 3). To counter UDP-glucose decomposition either more UDP has to be applied or UMP has to be recovered through phosphorylation to UDP [14, 48].

4 Concluding remarks

The current study of nothofagin synthesis via 3'-β-C-glucosylation of phloretin revealed how critically important the efficiency of UDP-glucose regeneration by SuSy was for the overall performance of the GT-SuSy cascade reaction. It showed furthermore that besides the relevant substrate-product pair (here: phloretin and nothofagin), UDP and UDP-glucose had to be quantified throughout the reaction to gain detailed understanding of the cascade's function. A new HPLC method was therefore developed to enable efficient determination of the required reactant concentrations in a single run. In a systematic analysis, for the first time, the impact of kinetic properties of the SuSy on synthetic efficiency of the bienzymatic cascade was demonstrated. Fitness in terms of kinetics, expressed in relatively low K_M-values for UDP and sucrose, made the SuSy from plant source the strongly preferred choice. It superseded enhanced thermostability in the SuSys from bacterial source as criterion of selection. The nothofagin formation rate improved almost 100-fold by choosing the proper SuSy enzyme. A simplified practical approach to kinetic evaluation was determination of the "system K_M" where the dependence of the glycoside formation rate on the UDP concentration was measured. Roughly, a system K_M of just around one-tenth of the added UDP concentration during the reaction seemed ideal and only the plant SuSy fulfilled this criterion. Advancements in NAD(P)H/NAD(P)$^+$ recycling systems for oxidoreductase cascade reactions have demonstrated that coenzyme regeneration poses no longer an intrinsic hindrance to economic biocatalytic processes employing these systems [28, 49]. In the GT-SuSy cascade, however, in order to achieve a significant number of regeneration cycles, the UDP "co-substrate" had to be applied in the µM-range where only GmSuSy worked efficiently. Combined with suitable strategies of solubility enhancement for the hydrophobic acceptor substrate, recycling numbers of about 100 are therefore clearly within reach for UDP-glucose using the SuSy reaction [19, 34]. This strongly supports the use of GT-SuSy cascades in natural product glycosylations targeted at specialty chemical synthesis. A final note of caution concerns the reaction temperature. While increasing the temperature seemed generally useful when thermostable GTs were available, spontaneous decomposition of the nucleotide sugars could become problematic under these conditions.

5   References

[1] Weijers, C. A. G. M., Franssen, M. C. R., Visser, G. M., Glycosyltransferase-catalyzed synthesis of bioactive oligosaccharides. *Biotechnol. Adv.* 2008, *26*, 436–456.

[2] Bowles, D., Lim, E.-K., Poppenberger, B., Vaistij, F. E., Glycosyltransferases of lipophilic small molecules. *Annu. Rev. Plant Biol.* 2006, *57*, 567–597.

[3] Liang, D.-M., Liu, J.-H., Wu, H., Wang, B.-B. et al., Glycosyltransferases: mechanisms and applications in natural product development. *Chem. Soc. Rev.* 2015, *44*, 8350–8374.

[4] De Bruyn, F., Maertens, J., Beauprez, J., Soetaert, W., De Mey, M., Biotechnological advances in UDP-sugar based glycosylation of small molecules. *Biotechnol. Adv.* 2015, *33*, 288–302.

[5] Elling, L., Glycobiotechnology: Enzymes for the synthesis of nucleotide sugars, in: Scheper, T. (Ed.), *New Enzymes for Organic Synthesis*, Springer Berlin, Heidelberg 1997, pp. 89–144.

[6] Tsai, T.-I., Lee, H.-Y., Chang, S.-H., Wang, C.-H. et al., Effective sugar nucleotide regeneration for the large-scale enzymatic synthesis of Globo H and SSEA4. *J. Am. Chem. Soc.* 2013, *135*, 14831–14839.

[7] Koizumi, S., Endo, T., Tabata, K., Ozaki, A., Large-scale production of UDP-galactose and globotriose by coupling metabolically engineered bacteria. *Nat. Biotechnol.* 1998, *16*, 847–850.

[8] Zhang, C., Griffith, B. R., Fu, Q., Albermann, C. et al., Exploiting the reversibility of natural product glycosyltransferase-catalyzed reactions. *Science* 2006, *313*, 1291–1294.

[9] Gutmann, A., Krump, C., Bungaruang, L., Nidetzky, B., A two-step *O*- to *C*-glycosidic bond rearrangement using complementary glycosyltransferase activities. *Chem. Commun.* 2014, *50*, 5465–5468.

[10] Gantt, R. W., Peltier-Pain, P., Cournoyer, W. J., Thorson, J. S., Using simple donors to drive the equilibria of glycosyltransferase-catalyzed reactions. *Nat. Chem. Biol.* 2011, *7*, 685–691.

[11] Gutmann, A., Nidetzky, B., Unlocking the potential of Leloir glycosyltransferases for applied biocatalysis: Efficient synthesis of uridine 5'-diphosphate-glucose by sucrose synthase. *Adv. Synth. Catal.* 2016, *358*, 3600–3609.

[12] Elling, L., Grothus, M., Kula, M.-R., Investigation of sucrose synthase from rice for the synthesis of various nucleotide sugars and saccharides. *Glycobiology* 1993, *3*, 349–355.

[13] Schmölzer, K., Gutmann, A., Diricks, M., Desmet, T., Nidetzky, B., Sucrose synthase: A unique glycosyltransferase for biocatalytic glycosylation process development. *Biotechnol. Adv.* 2016, *34*, 88–111.

[14] Kulmer, S. T., Gutmann, A., Lemmerer, M., Nidetzky, B., Biocatalytic cascade of polyphosphate kinase and sucrose synthase for synthesis of nucleotide-activated derivatives of glucose. *Adv. Synth. Catal.* 2017, *359*, 292–301.

[15] Masada, S., Kawase, Y., Nagatoshi, M., Oguchi, Y. et al., An efficient chemoenzymatic production of small molecule glucosides with in situ UDP-glucose recycling. *FEBS Lett.* 2007, *581*, 2562–2566.

[16] Terasaka, K., Mizutani, Y., Nagatsu, A., Mizukami, H., In situ UDP-glucose regeneration unravels diverse functions of plant secondary product glycosyltransferases. *FEBS Lett.* 2012, *586*, 4344–4350.

[17] Bungaruang, L., Gutmann, A., Nidetzky, B., Leloir glycosyltransferases and natural product glycosylation: Biocatalytic synthesis of the *C*-glucoside nothofagin, a major antioxidant of Redbush herbal tea. *Adv. Synth. Catal.* 2013, *355*, 2757–2763.

[18] Gutmann, A., Bungaruang, L., Weber, H., Leypold, M. et al., Towards the synthesis of glycosylated dihydrochalcone natural products using glycosyltransferase-catalysed cascade reactions. *Green Chem.* 2014, *16*, 4417–4425.

[19] Michlmayr, H., Malachová, A., Varga, E., Kleinová, J. et al., Biochemical characterization of a recombinant UDP-glucosyltransferase from rice and enzymatic production of deoxynivalenol-3-*O*-β-D-glucoside. *Toxins* 2015, *7*, 2685–2700.

[20] Huang, F.-C., Hinkelmann, J., Hermenau, A., Schwab, W., Enhanced production of β-glucosides by in-situ UDP-glucose regeneration. *J. Biotechnol.* 2016, *224*, 35–44.

[21] Hokke, C. H., Zervosen, A., Elling, L., Joziasse, D. H., van den Eijnden, D. H., One-pot enzymatic synthesis of the Galα1→3Galβ1→4GlcNAc sequence with in situ UDP-Gal regeneration. *Glycoconjugate J.* 1996, *13*, 687–692.

[22] Zervosen, A., Elling, L., A novel three-enzyme reaction cycle for the synthesis of *N*-acetyllactosamine with in situ regeneration of uridine 5'-diphosphate glucose and uridine 5'-diphosphate galactose. *J. Am. Chem. Soc.* 1996, *118*, 1836–1840.

[23] Brinkmann, N., Malissard, M., Ramuz, M., Römer, U. et al., Chemoenzymatic synthesis of the Galili epitope Galα(1→3)Galβ(1→4)GlcNAc on a homogeneously soluble PEG polymer by a multi-enzyme system. *Bioorg. Med. Chem. Lett.* 2001, *11*, 2503–2506.

[24] Chen, X., Zhang, J., Kowal, P., Liu, Z. et al., Transferring a biosynthetic cycle into a productive *Escherichia coli* strain: Large-scale synthesis of galactosides. *J. Am. Chem. Soc.* 2001, *123*, 8866–8867.

[25] Liu, Z., Lu, Y., Zhang, J., Pardee, K., Wang, P. G., P1 trisaccharide (Galα1,4Galβ1,4GlcNAc) synthesis by enzyme glycosylation reactions using recombinant *Escherichia coli*. *Appl. Environ. Microbiol.* 2003, *69*, 2110–2115.

[26] Engels, L., Henze, M., Hummel, W., Elling, L., Enzyme module systems for the synthesis of uridine 5'-diphospho-α-D-glucuronic acid and non-sulfated human natural killer cell-1 (HNK-1) epitope. *Adv. Synth. Catal.* 2015, *357*, 1751–1762.

[27] Rupprath, C., Kopp, M., Hirtz, D., Müller, R., Elling, L., An enzyme module system for in situ regeneration of deoxythymidine 5'-diphosphate (dTDP)-activated deoxy sugars. *Adv. Synth. Catal.* 2007, *349*, 1489–1496.

[28] Zhao, H., van der Donk, W. A., Regeneration of cofactors for use in biocatalysis. *Curr. Opin. Biotechnol.* 2003, *14*, 583–589.

[29] Durren, R. L., McIntosh, C. A., Flavanone-7-*O*-glucosyltransferase activity from *Petunia hybrida*. *Phytochemistry* 1999, *52*, 793–798.

[30] Diricks, M., De Bruyn, F., Van Daele, P., Walmagh, M., Desmet, T., Identification of sucrose synthase in nonphotosynthetic bacteria and characterization of the recombinant enzymes. *Appl. Microbiol. Biotechnol.* 2015, *99*, 8465–8474.

[31] Diricks, M., Gutmann, A., Debacker, S., Dewitte, G. et al., Sequence determinants of nucleotide binding in Sucrose Synthase: improving the affinity of a bacterial Sucrose Synthase for UDP by introducing plant residues. *Protein Eng. Des. Sel.* 2017, *30*, 141–148.

[32] Figueroa, C. M., Asención Diez, M. D., Kuhn, M. L., McEwen, S. et al., The unique nucleotide specificity of the sucrose synthase from *Thermosynechococcus elongatus*. *FEBS Lett.* 2013, *587*, 165–169.

[33] Schmölzer, K., Lemmerer, M., Gutmann, A., Nidetzky, B., Integrated process design for biocatalytic synthesis by a Leloir glycosyltransferase: UDP-glucose production with sucrose synthase. *Biotechnol. Bioeng.* 2017, *114*, 924–928.

[34] Bungaruang, L., Gutmann, A., Nidetzky, B., β-Cyclodextrin improves solubility and enzymatic *C*-glucosylation of the flavonoid phloretin. *Adv. Synth. Catal.* 2016, *358*, 486–493.

[35] Ku, S.-K., Kwak, S., Kim, Y., Bae, J.-S., Aspalathin and nothofagin from Rooibos (*Aspalathus linearis*) inhibits high glucose-induced inflammation in vitro and in vivo. *Inflammation* 2015, *38*, 445–455.

Biotechnology
Journal
www.biotechnology-journal.com

Biotechnol. J. 2017, 12, 1600557

ADVANCED
SCIENCE NEWS
www.advancedsciencenews.com

[36] Schloms, L., Storbeck, K.-H., Swart, P., Gelderblom, W. C. A., Swart, A. C., The influence of *Aspalathus linearis* (Rooibos) and dihydroch-alcones on adrenal steroidogenesis: Quantification of steroid inter-mediates and end products in H295R cells. *J. Steroid Biochem. Mol. Biol.* 2012, *128*, 128–138.

[37] Snijman, P. W., Swanevelder, S., Joubert, E., Green, I. R., Gelderblom, W. C. A., The antimutagenic activity of the major flavonoids of rooibos (*Aspalathus linearis*): Some dose-response effects on mutagen activa-tion-flavonoid interactions. *Mutat. Res.* 2007, *631*, 111–123.

[38] Brazier-Hicks, M., Evans, K. M., Gershater, M. C., Puschmann, H. et al., The *C*-glycosylation of flavonoids in cereals. *J. Biol. Chem.* 2009, *284*, 17926–17934.

[39] Gutmann, A., Nidetzky, B., Switching between *O*- and *C*-glycosyl-transferase through exchange of active-site motifs. *Angew. Chem. Int. Ed.* 2012, *51*, 12879–12883.

[40] Wahl, C., Hirtz, D., Elling, L., Multiplexed capillary electrophoresis as analytical tool for fast optimization of multi-enzyme cascade reactions – synthesis of nucleotide sugars. *Biotechnol. J.* 2016, *11*, 1298–1308.

[41] Zheng, Y., Anderson, S., Zhang, Y., Garavito, R. M., The structure of sucrose synthase-1 from *Arabidopsis thaliana* and its functional implications. *J. Biol. Chem.* 2011, *286*, 36108–36118.

[42] Curatti, L., Porchia, A. C., Herrera-Estrella, L., Salerno, G. L., A prokaryotic sucrose synthase gene (*susA*) isolated from a filamen-tous nitrogen-fixing cyanobacterium encodes a protein similar to those of plants. *Planta* 2000, *211*, 729–735.

[43] Shao, J., Hayashi, T., Wang, P. G., Enhanced production of α-galactosyl epitopes by metabolically engineered *Pichia pastoris*. *Appl. Environ. Microbiol.* 2003, *69*, 5238–5242.

[44] Desmet, T., Soetaert, W., Bojarová, P., Křen, V. et al., Enzymatic glycosylation of small molecules: challenging substrates require tailored catalysts. *Chem. Eur. J.* 2012, *18*, 10786–10801.

[45] Lee, J. C., Timasheff, S. N., The stabilization of proteins by sucrose. *J. Biol. Chem.* 1981, *256*, 7193–7201.

[46] De Winter, K., Verlinden, K., Křen, V., Weignerová, L. et al., Ionic liquids as cosolvents for glycosylation by sucrose phosphorylase: balancing acceptor solubility and enzyme stability. *Green Chem.* 2013, *15*, 1949–1955.

[47] Baroja-Fernández, E., Muñoz, F. J., Li, J., Bahaji, A. et al., Sucrose synthase activity in the *sus1/sus2/sus3/sus4 Arabidopsis* mutant is sufficient to support normal cellulose and starch production. *Proc. Natl. Acad. Sci. U.S.A.* 2012, *109*, 321–326.

[48] Zervosen, A., Stein, A., Adrian, H., Elling, L., Combined enzymatic synthesis of nucleotide (deoxy) sugars from sucrose and nucleoside monophosphates. *Tetrahedron* 1996, *52*, 2395–2404.

[49] Goldberg, K., Schroer, K., Lütz, S., Liese, A., Biocatalytic ketone reduction—a powerful tool for the production of chiral alcohols—part I: processes with isolated enzymes. *Appl. Microbiol. Biotechnol.* 2007, *76*, 237–248.

Biotechnology Journal

Glycosyltransferase cascades for natural product glycosylation: Use of plant instead of bacterial sucrose synthases improves the UDP-glucose recycling from sucrose and UDP

Alexander Gutmann, Alexander Lepak, Margo Diricks, Tom Desmet, Bernd Nidetzky
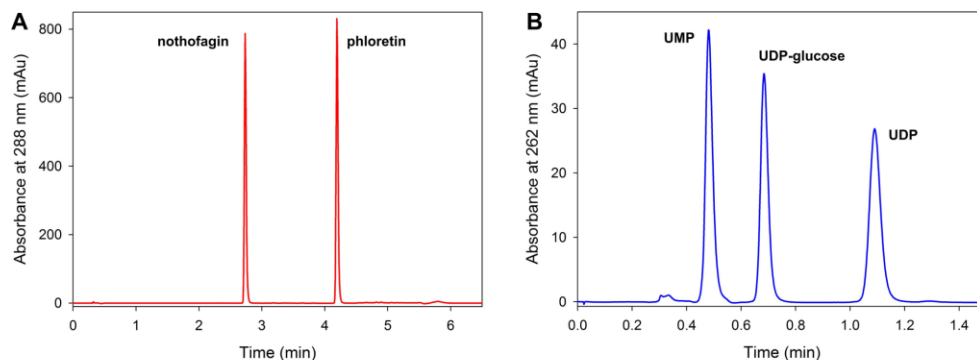
Figure S1. A) Nothofagin and phloretin were quantified by reversed-phase C-18 HPLC using a water/acetonitrile gradient. B) UMP, UDP-glucose and UDP were quantified by isocratic reversed-phase C-18 HPLC using a TBAB based ion-pairing protocol.

Scheme S1. UDP-glucose can spontaneously decompose to UMP and glucose-1,2-cyclic phosphate.
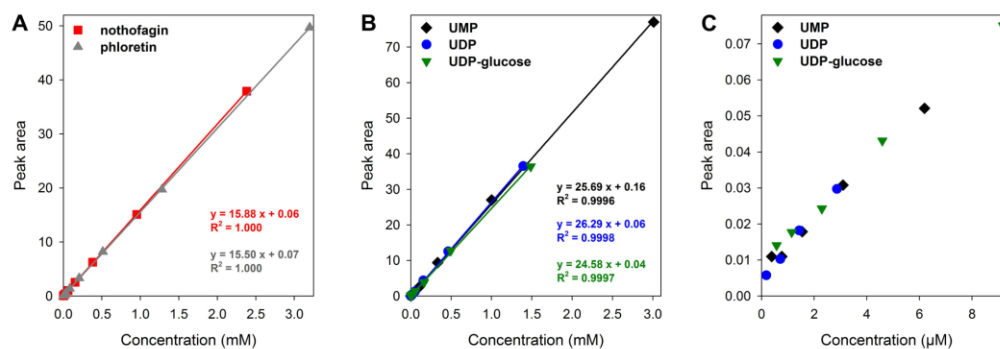
Figure S2. The linear dynamic range (A, B) and limits of quantitation and detection (C) of the reversed-phase ion-pairing gradient HPLC assay were determined. The injection volume was fixed to 5 μL and compounds were monitored by UV-detection at 288 (A) or 262 nm (B, C).
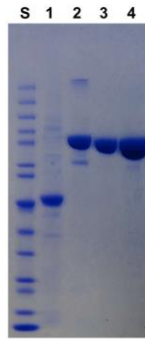
Figure S3. SDS-PAGE of affinity purified enzymes from *E. coli* overexpression cultures. S: PageRuler™ Prestained Protein Ladder (Thermo Scientific); 1: *Os*CGT; 2: *Gm*SuSy; 3: *Ac*SuSy; 4: *Ac*SuSy variant. *Os*CGT and *Gm*SuSy were purified by *Strep*-tag affinity chromatography. *Ac*SuSy wildtype and variant were purified by His-tag affinity chromatography.
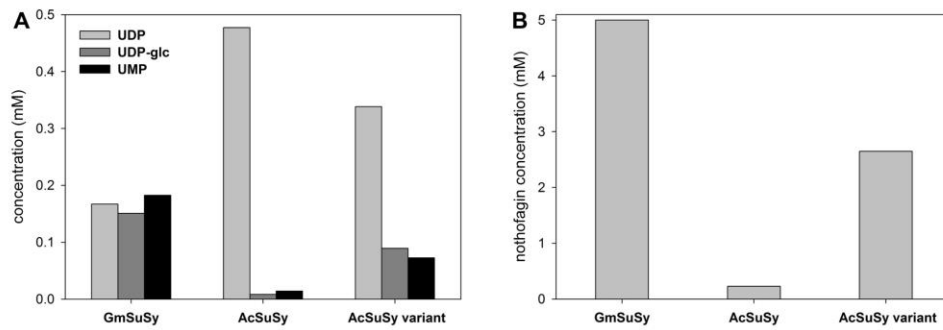
Figure S4. Concentrations of UDP, UDP-glucose, UMP (A) and nothofagin (B) after 3h of *Os*CGT-SuSy cascade conversions at 60°C are shown. Data was extracted from reactions displayed in Fig. 4. Reaction conditions: 5 mM phloretin, 0.5 mM UDP, 1 M sucrose, 30 µg/mL *Os*CGT, 10 µg/mL SuSy, pH 7.5.

S3

Chapter 4:
Biochemical characterization and mechanistic analysis of the
levoglucosan kinase from *Lipomyces starkeyi*

Biochemical Characterization and Mechanistic Analysis of the Levoglucosan Kinase from *Lipomyces starkeyi*

Christina Rother[+],[a] Alexander Gutmann[+],[a] Ramakrishna Gudiminchi,[a, b] Hansjörg Weber,[c] Alexander Lepak,[a] and Bernd Nidetzky*[a, b]

Levoglucosan kinase (LGK) catalyzes the simultaneous hydrolysis and phosphorylation of levoglucosan (1,6-anhydro-β-D-glucopyranose) in the presence of Mg^{2+}–ATP. For the *Lipomyces starkeyi* LGK, we show here with real-time in situ NMR spectroscopy at $10\,^\circ$C and pH 7.0 that the enzymatic reaction proceeds with inversion of anomeric stereochemistry, resulting in the formation of α-D-glucose-6-phosphate in a manner reminiscent of an inverting β-glycoside hydrolase. Kinetic characterization revealed the Mg^{2+} concentration for optimum activity (20–50 mM), the apparent binding of levoglucosan ($K_m = 180$ mM) and ATP ($K_m = 1.0$ mM), as well as the inhibition by ADP ($K_i = 0.45$ mM) and D-glucose-6-phosphate ($IC_{50} = 56$ mM). The enzyme was highly specific for levoglucosan and exhibited weak ATPase activity in the absence of substrate. The equilibrium conversion of levoglucosan and ATP lay far on the product side, and no enzymatic back reaction from D-glucose-6-phosphate and ADP was observed under a broad range of conditions. 6-Phospho-α-D-glucopyranosyl fluoride and 6-phospho-1,5-anhydro-2-deoxy-D-*arabino*-hex-1-enitol (6-phospho-D-glucal) were synthesized as probes for the enzymatic mechanism but proved inactive with the enzyme in the presence of ADP. The pyranose ring flip $^4C_1 \rightarrow {}^1C_4$ required for 1,6-anhydro-product synthesis from D-glucose-6-phosphate probably presents a major thermodynamic restriction to the back reaction of the enzyme.

Introduction

Levoglucosan kinase (LGK) catalyzes the conversion of the anhydrosugar levoglucosan (1,6-anhydro-β-D-glucopyranose, **1**) in the presence of adenosine triphosphate (ATP) and Mg^{2+}.[1] The enzymatic reaction involves combined hydrolysis and phosphorylation of the substrate to yield D-glucose-6-phosphate (**2**) and adenosine diphosphate (ADP) as the products (Scheme 1). Together with AnmK (1,6-anhydro-*N*-acetylmuramic acid kinase), LGK forms a subfamily of anhydrosugar kinases within the large hexokinase protein family.[1a,2] LGKs from *Lipomyces starkeyi*,[1a,3] *Aspergillus niger*,[4] and *Sporobolomyces salmonicolor*[5] have been purified and characterized. Furthermore, LGK activity was detected in the cell extracts of various yeasts and filamentous fungi.[1b,6]

[a] C. Rother,[+] Dr. A. Gutmann,[+] Dr. R. Gudiminchi, A. Lepak, Prof. Dr. B. Nidetzky
    Institute of Biotechnology and Biochemical Engineering
    Graz University of Technology, NAWI Graz
    Petersgasse 12, 8010 Graz (Austria)
    E-mail: bernd.nidetzky@tugraz.at

[b] Dr. R. Gudiminchi, Prof. Dr. B. Nidetzky
    Austrian Centre of Industrial Biotechnology
    Petersgasse 14, 8010 Graz (Austria)

[c] Prof. H. Weber
    Graz University of Technology, NAWI Graz
    Stremayrgasse 9, 8010 Graz (Austria)

[+] These authors contributed equally to this work.

    Supporting Information and the ORCID identification numbers for the authors of this article can be found under https://doi.org/10.1002/cbic.201700587.
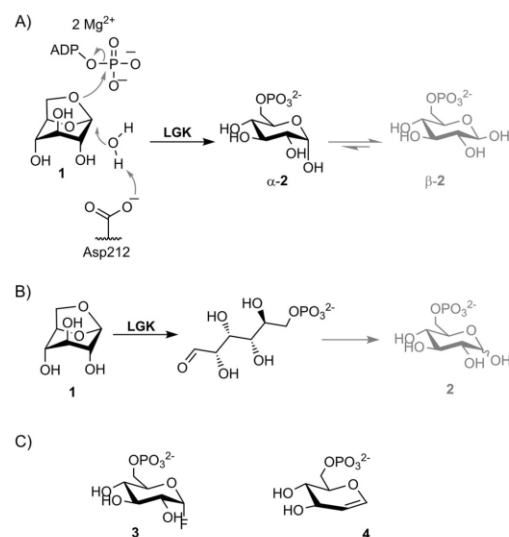
Scheme 1. Mechanistic proposals suggest formation of A) α-**2** or B) an open-chain form of **2** as products of the LGK reaction.[1a] Anomerization of **2** (shown in gray) most probably occurs in solution. C) Compounds **3** and **4** were synthesized as mechanistic probes of the enzymatic reaction in the reverse direction.

Because of the anhydro linkage, the pyranose ring of **1** is locked in the otherwise energetically highly unfavorable $^1C_4$ chair conformation.[7] Compound **1** is a major product of ligno-

cellulose pyrolysis and accounts for up to 30 wt% of the liquid fraction obtained in the process.[8] Therefore, the use of **1** as a substrate for bioconversion by microbial fermentation is of rapidly growing interest.[8a, 9]

Besides the biocatalytic application of LGK, the enzyme's catalytic mechanism has also drawn considerable attention. How LGK couples hydrolysis of anhydro substrate **1** with phosphorylation from ATP is a problem of fundamental significance. LGK and AnmK uniquely combine the catalytic features of a glycoside hydrolase (GH) with those of an ATP-dependent sugar kinase.[10] The structures of the LGK from *L. starkeyi* (LsLGK)[1a] and the AnmK from *Pseudomonas aeruginosa*[10a, 11] have been determined. In both enzymes, a conserved Asp was suggested to serve as catalytic base for attack of water on the substrate's anomeric carbon atom. On hydrolytic opening of the 1,6-anhydro ring, phosphorylation at C6 also occurs. A stereochemical reaction course was proposed for LGK (Scheme 1A), in which the α-anomer of **2** is the product. Accordingly, the anomeric configuration changes between β in **1** and α in **2**, which is consistent with the stereochemical course of the reaction of an inverting β-glycoside hydrolase.[12] A NMR spectroscopy time-course experiment was used to analyze the formation of **2** by LsLGK.[1a] However, product **2** was detected as an equilibrium mixture of anomers (38% α, 62% β).[13] It was therefore suggested that the LGK reaction might involve a pyranose ring-opening step, as shown in Scheme 1B.[1a] This might facilitate release of ring strain present in **1**. Ring closure would then yield a mixture of anomers of **2** at equilibrium. However, the very fast mutarotation of **2** complicates this analysis.[13] Interestingly, NMR spectroscopy analysis of the reaction of AnmK revealed formation of the α-anomer in a stereochemically inverting reaction.[10a]

The current study was performed to advance mechanistic understanding of LsLGK on the basis of three lines of experiments. The problem of the catalytic reaction's stereochemical course was reassessed by real-time NMR spectroscopy under conditions that reduce the rate of mutarotation. 6-Phospho-α-D-glucopyranosyl fluoride (**3**) and 6-phospho-1,5-anhydro-2-deoxy-D-*arabino*-hex-1-enitol (**4**) were synthesized and evaluated as mechanistic probes of LGK, potentially functioning as analogues of **2** in the reverse reaction of the enzyme. Compounds **3** and **4** were of particular interest, because unlike **2**, which must release water to enable 1,6-anhydro ring formation (Scheme 1A), they could become "activated" much more easily by expulsion of fluoride or protonation, respectively. It was also considered that the conversion of **3** into **1** might thermodynamically (owing to the exergonic release of fluoride) be much more favored than the conversion of **2**. Finally, although LsLGK was characterized biochemically in some detail before,[1a, 3] certain properties of the enzyme and the reaction catalyzed by it were not clear. This required a systematic re-investigation and new evidence was obtained in its course.

Results and Discussion

Purification and stabilization of LsLGK

LsLGK equipped with a N- or C-terminal His-tag was produced in *Escherichia coli*. The N-terminally tagged protein accumulated mostly in aggregates and was hardly active. The N termini in the dimeric assembly of LsLGK are in close proximity to each other (Figure S1 in the Supporting Information). Steric conflict may therefore prevent accommodation of the His-tag in a still-native enzyme structure. In contrast, the C-terminally tagged protein was expressed well and could be purified conveniently (Figure S2). A specific activity of (16.9 ± 2.2) U mg^{-1} was determined at pH 7.8 under optimized assay conditions. The far-UV circular dichroism (CD) spectrum of the purified LsLGK (Figure S3) indicated a properly folded protein. The relative composition of secondary structural elements in LsLGK calculated from the CD data was in agreement with that found in the enzyme's crystal structure.[1a]

Purified LsLGK proved to be a highly unstable protein. Differential scanning fluorimetry (DSF) revealed a melting temperature (T_m) of just 37°C. The enzyme was not stabilized by one of its substrates or products, except for ATP (2 mM) together with MgCl$_2$ (10 mM), which increased T_m by 3°C. Consistent with the low T_m, LsLGK was quickly inactivated under typical reaction conditions (Figure S4A). At 30°C and pH 7.5, all activity was lost within just 2 h. A temperature decrease to 25°C and the addition of 1 mg mL^{-1} bovine serum albumin (BSA) enhanced the enzyme's half-life to approximately 4 h. Raising the BSA concentration to 5 mg mL^{-1} did not further enhance the stability (Figure S4B). Stability was pH dependent, with an optimum at approximately pH 7.9 (Figure S4C). Limited stability of LsLGK was noted before,[14] but in another study,[3] the enzyme was found to be stable (for 30 min) at 30°C in the pH range of 7 to 9. With the main factors of LsLGK stability unknown at this time, we proceeded forward taking carefully into account that fast activity loss restricted the usable incubation times in biochemical assays for enzyme characterization.

Kinetic characterization and substrate specificity of LsLGK

The pH–activity profile of LsLGK has a sharp optimum at pH 8 (Figure S5), which corresponds well to the optimum of stability at pH 7.9 (Figure S4C). Therefore, HEPES buffer (pH 7.8) was used in all subsequent studies.

The active site of LsLGK contains two Mg^{2+} ions that interact with the phosphate groups of ATP.[1a] The ratio of Mg^{2+}/ATP strongly influences the activity of LsLGK. Using 4 and 8 mM Mg^{2+}, the activity was highest at a Mg^{2+}/ATP ratio of 2 and drastically dropped if the ratio was less than 1.[1a] We considered that the ideal concentration of Mg^{2+} might also depend on the concentrations of ATP and LsLGK used. Employing a constant ATP concentration (2.5 mM), the enzyme activity increased with the Mg^{2+} concentration until a Mg^{2+}/ATP ratio of 8 was reached at 20 mM MgCl$_2$ (Figure 1A). Any further increase in Mg^{2+} led to a moderate decrease in activity. It was previously shown for a cyclin-dependent kinase, which similar
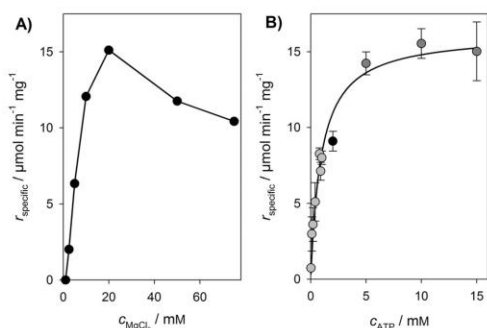
Figure 1. The dependence of the LsLGK reaction rate on Mg^{2+} and ATP concentration was determined (500 mM **1**, 0.02 mg mL^{-1} LsLGK, pH 7.8, 30 °C). A) Reaction in the presence of 2.5 mM ATP is shown. B) To determine the K_m for ATP, the Mg^{2+} concentration was adjusted to the ATP concentration used: 10 mM (light gray), 20 mM (black), and 50 mM (dark gray) MgCl$_2$ were used. The solid line shows a hyperbolic fit of the data.
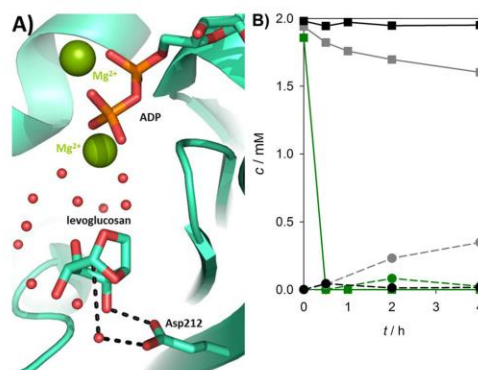

Figure 2. A) The active site of LsLGK is shown (PDB ID: 4ZLU, chain A).[1a] Only the catalytic base Asp212 forms a hydrogen bond with C4 of **1**, which is surrounded by water molecules. Bound **1** appears to adopt a $^1C_4$ pyranosyl ring conformation. B) ATP consumption (squares, solid lines) and release of free phosphate (circles, dashed lines) were measured to study the canonical LsLGK reaction and error hydrolysis of ATP. Reactions in the absence of LsLGK (black) or **1** (gray) were compared to a standard conversion (green). Reaction conditions: 0/100 mM **1**, 2 mM ATP, 0/1 mg mL^{-1} LsLGK, 10 mM MgCl$_2$, pH 7.8, 30 °C.

to LsLGK requires the binding of two magnesium ions for activity,[15] that the complex of 2 Mg^{2+} and ADP binds more tightly to the enzyme than does the corresponding ATP complex. The same effect might also lead to "Mg^{2+} substrate" inhibition of LsLGK.[1a, 15]

The influence of Mg^{2+} was considered in the determination of the K_m for ATP (Figure 1 B). Adjusted to the ATP concentration used, the MgCl$_2$ concentration was varied from 10 to 50 mM to avoid conditions too high or too low in Mg^{2+}. The resulting K_m was (1.0 ± 0.2) mM, as shown in Figure 1 B. Somewhat lower K_m values of (0.20 ± 0.02)[3a] and (0.68 ± 0.06) mM[3b] were previously obtained by using a constant MgCl$_2$ concentration of 10 mM. At the time of these earlier measurements, rate reduction due to Mg^{2+} limitation was not yet reported and so might have been mistaken for substrate saturation by ATP. The K_m for **1** was (177 ± 36) mM (Figure S6). It was again slightly higher but comparable with earlier reported values (69–119 mM).[1a, 3] Product inhibition of LGK was tested in reactions containing 300 mM **1** and 2.5 mM ATP. ADP was a strong competitive inhibitor against ATP with a K_i of (0.45 ± 0.03) mM (Figure S7 A). Similarly strong inhibition by ADP was reported for the LGKs from *S. salmonicolor* ($K_i = 0.15$ mM)[5] and *A. niger* ($K_i = 0.20$ mM).[4a] Product **2** also inhibited the LsLGK, albeit only weakly [IC$_{50} = (56 \pm 10)$ mM, Figure S7 B]. D-Glucose, D-mannose, and D-fructose hardly inhibited the enzyme (Figure S8).

The crystal structure of the LsLGK reveals a relatively "open" substrate-binding site (Figure 2 A).[1a] Somewhat surprisingly for a carbohydrate-active enzyme, **1** forms only a single hydrogen bond with the LsLGK (Asp212). Otherwise, **1** is surrounded by numerous water molecules. This raises concern that hydrolysis might compete effectively with the canonical use of ATP for phosphorylation of **1**. Hexokinases are generally known to exhibit ATPase activity in the absence of substrate.[16] Indications of minor "error" hydrolysis of ATP were also found with the LGK from *S. salmonicolor*.[5] To distinguish between synthesis of

2 and hydrolysis of ATP in the enzymatic reaction, we analyzed the release of ADP and inorganic phosphate on the conversion of ATP. We compared utilization of ATP in the presence of **1** to its utilization if **1** or the LsLGK was lacking (Figures 2 B and S9). If **1** was present, all ATP (2 mM) was rapidly converted into ADP and no hydrolysis took place. In the absence of **1**, however, ATP was degraded by hydrolysis. The activity of the LsLGK for ATP hydrolysis (≈ 4 mU mg^{-1}) was approximately 400-fold slower than the activity for the conversion of **1** into **2**. The gradual slowdown in ATP conversion is explainable by the limited stability of the LsLGK. The ADP produced initially was not hydrolyzed further. A control lacking the LsLGK showed that ATP was stable under the applied conditions (Figure 2 B).

The binding of **1** by the LsLGK (Figure 2 A) inspires the belief that the enzyme might utilize anhydrosugars other than **1** as its substrates. However, biochemical evidence suggests LGKs to be rather specific for **1**.[3b, 4a, 5] We tested a series of anhydrosugars and also D-glucose for utilization by the LsLGK in the presence of ATP (Figure S10). In all cases, no conversion of the putative substrate was detected. Consumption of ATP was identical to the control comprising only enzyme and ATP. Formation of ADP was always closely matched to the release of phosphate. High substrate specificity of the LsLGK is supported.

In situ NMR spectroscopy supports an inverting mechanism of LsLGK

The use of real-time ^{1}H NMR spectroscopy to study the stereochemical course of the LGK reaction is complicated by the fast mutarotation of released **2**. At 25 °C and pH 7, the half-lives of α- and β-**2** were reported to be just 5 and 10 s, respectively.[13]

65

On lowering the temperature to 0 °C, the half-life of β-**2** was increased to 2.1 min.[17] Changing the solvent from H_2O to D_2O also enhanced anomer stability approximately 1.7-fold.[18] Conditions suitable for in situ ¹H NMR spectroscopy analysis therefore involve enzymatic conversion of **1** into **2** that proceeds faster than the subsequent mutarotation. We carefully optimized temperature (10 °C), pD (7.0), LsLGK (2 mg mL⁻¹), and substrate concentrations (100 mм **1**, 10 mм ATP) to this end.

The results of the real-time monitoring of the enzymatic reaction are shown in Figure 3 A. The first NMR spectrum recorded 5 min after the start reveals a clear excess of α-**2** (56 %)
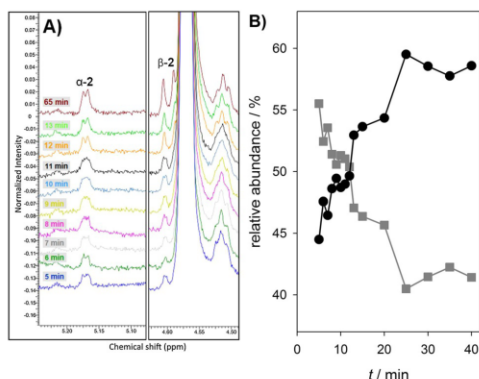

Figure 3. Stereochemical course for the conversion of **1** by LsLGK was analyzed by using real-time in situ NMR spectroscopy (100 mм **1**, 10 mм ATP, 20 mм $MgCl_2$, 2 mg mL⁻¹ LsLGK, pD 7.0, 10 °C). A) Superposition of the ¹H NMR spectra for the conversion of **1** into **2**. Characteristic signals for α-**2** and β-**2** are indicated in the spectra. B) Relative portions of the α- and β-anomers of **2** formed during the conversion of **1** as a function of reaction time.

over β-**2** (44 %). The relative abundance of β-**2** subsequently increases until after approximately 25 min the mutarotation equilibrium is reached (Figure 3 B). The seemingly conflicting observation of an anomeric mixture of **2** by Bacik et al. is probably due to increased mutarotation at the higher reaction temperature (20 °C instead of 10 °C) and longer reaction time (4 h instead of 40 min).[1a] In fact, during the less emphasized first minutes of the conversion, some accumulation of α-**2** was also visible in the previous in situ NMR spectroscopy study.[1a] The distinct patterns for the formation of α-**2** and β-**2** in dependence of time (Figure 3 B) support a reaction scenario in which the LsLGK is specific for the formation of the α-anomer of **2**. Because glucopyranose ring opening (Scheme 1 B) is not in line with initial accumulation of α-**2**, we propose an inverting (β → α) stereochemical reaction course of LsLGK, similar to that of AnmK (Scheme 1 A). Accordingly, LsLGK combines mechanistic features of an ATP-dependent kinase and an inverting glycoside hydrolase.[10c, d] From the enzyme's structure,[1a] Asp212 might function as a general base by facilitating attack of water in the catalytic mechanism (Scheme 1 A, Figure 2 A).

The LsLGK reaction is experimentally quasi-irreversible

Conversion of **1** into **2** is expected to release a large amount of free energy, resulting from ATP conversion and from relaxation of conformational strain in **1** upon hydrolysis of the β-1,6-glycosidic bond. Just the phosphorylation of D-glucose from ATP releases approximately 20 kJ mol⁻¹ of energy if performed at neutral pH in the presence of Mg^{2+}.[19] Enzymatic reactions at varied substrate concentrations (ATP: 2–10 mм; **1**: 2–100 mм; pH 7.8) confirmed that the equilibrium of the LsLGK reaction lay far on the side of product **2**. However, despite the high enzyme concentrations used in the experiments (3 mg mL⁻¹), the time to reaction equilibrium exceeded the lifetime of the enzyme. Determination of the equilibrium constant (K_{eq}) was therefore not pursued. To investigate the formation of **1** by the enzymatic back reaction, a broad range of concentrations of ADP (0.2–20 mм) and **2** (1–150 mм) was tested. The pH was also varied in the range of 6 to 9, considering its possible effect on K_{eq} in phosphoryl-transfer reactions.[20] No **1** was detected under the different conditions used. Detection of **1** was performed routinely by HPLC by using refractive index detection. To enhance sensitivity, high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) was used, and the samples were pretreated to decrease the large excess of **2** and ADP. Sample preparation was validated with 0.3 mм **1** as an internal standard (Figure S11). The detection limit of **1** was indicated to be approximately 50 μм.

Analysis of mechanistic probes of the LGK reaction

Compounds **3** and **4** (Scheme 1 C) were synthesized as potential probes of the reverse reaction of the LsLGK. Their use was inspired by the fact that glycosyl fluorides and glycals have proven useful as alternative substrates of different glycoside hydrolases from both the retaining and the inverting mechanistic types.[21] However, enzymatic reactivity with these analogues can vary over a broad range in various glycoside hydrolases, and especially the glycals proved to be potent inhibitors.[21g, 22] We nonetheless considered that formation of the intramolecular β-1,6-glycosidic bond could be facilitated by replacing the difficult water-release step of the canonical reaction of **2** by a different, probably more easily achieved chemical step: release of fluoride from **3** and protonation of the double bond at C2 in **4** (Scheme 2). Both **3** and **4** do not rely, in principle, on the protonation of Asp212 for reactivity.
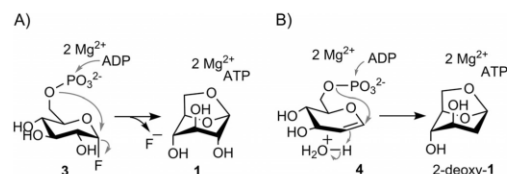
Scheme 2. Proposed mechanisms for the LsLGK-catalyzed conversion of A) **3** and B) **4** in the presence of ADP.

66

Although conversion of **4** also involves a protonation step, it occurs at C2 and not at the OH leaving group at C1, which forms a hydrogen bond with Asp212 (Scheme 1 A). Therefore, we envisioned that the numerous water molecules in the active site of the LsLGK (Figure 2 A) could act as shuttles for long-range proton transfer from an alternative general acid to the C2 of **4**.

To examine the reactivity of **3**, the LsLGK (1.5 mg mL^{-1}) was added directly to a reaction mixture obtained by phosphorylation of α-D-glucopyranosyl fluoride from ATP by using hexokinase. Whereas the formation of **3** was confirmed unambiguously by using ^{19}F NMR spectroscopy analysis (Figure S12), ^{1}H NMR spectroscopy analysis revealed that no further conversion of **3** into **1** occurred over 24 h (Figure S13). Also, no transformation of **3** into any other product was detectable.

The reactivity of **4** was examined in a manner similar to that described above for **3**. Compound **4** was synthesized by enzymatic phosphorylation of "D-glucal" (1,5-anhydro-2-deoxy-D-arabino-hex-1-enitol) and was isolated by anion-exchange chromatography and barium acetate precipitation. To avoid thermodynamic restrictions during the conversion of **4**, ATP was removed in situ by adenylate kinase from *E. coli* (EcADK).[23] The enzyme catalyzes the reversible formation of two ADP molecules from ATP and AMP. Therefore, besides 160 mM **4** and 10 mM ADP, the reaction mixture also contained 90 mM AMP. To compensate potentially low reaction rates with the substrate analogue, 10 mg mL^{-1} LsLGK was used, and the reaction was followed for 24 h by in situ ^{1}H NMR spectroscopy (Figure S14). However, 2-deoxy-**1** was not detected, and also, the conversion of **4** into alternative products was not observed.

Generally, these results suggest that "activation" of substrate in **3** and **4** compared to **2** could not be exploited by the enzyme to overcome an apparently very large energy barrier to β-1,6-anhydro ring formation. Note that release of fluoride from **3** is expected to release a large amount of free energy,[21b] which, in principle, could be utilized to promote the formation of **1**. Interestingly, therefore, uncoupling of hydrolysis from phosphoryl transfer to form **1** from **3** or 2-deoxy-**1** from **4** was also not observed in the reactions of the LsLGK.

Mechanistic proposal: A hypothetical catalytic itinerary of sugar puckering conformations

Considering Hammond's postulate,[24] the large decrease in standard Gibbs free energy (ΔG_0) upon the conversion of **1** into **2** would imply an early transition state, the structure of which would resemble **1** more than **2**. Besides cleavage of the anhydro linkage and transfer of a phosphate group from ATP, conversion of **1** by the LsLGK also has to involve a $^{1}C_4 \rightarrow {}^{4}C_1$ pyranose ring flip. This rearrangement must happen through an itinerary of sugar ring conformations, the variety of which is summarized in the Cremer–Pople spherical coordinate system (Figure 4 A).[25] Another point of consideration was that a glycoside hydrolase like reaction mechanism of the LsLGK would plausibly involve an oxocarbenium ion like transition state.[26] Such transition states are characterized by partial change in hy-
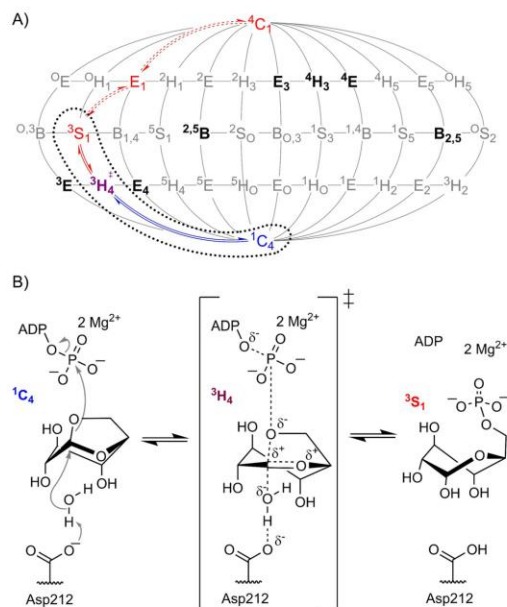

Figure 4. Mechanistic proposal for the LsLGK reaction. A) Suggested catalytic itinerary of sugar puckering conformations is displayed on the Cremer–Pople spherical coordinate system. Conformations in bold can adopt oxocarbenium ion like transition states. The proposed $^{3}H_4$ transition state involving cleavage of the anhydro bond with coincident change from the β configuration (blue) to α configuration (red) is shown in purple. The conformational itinerary, indicated by the dotted line, is expected to happen on the enzyme. The additional conformational rearrangements could also occur in solution. A tentative route is indicated. B) Mechanistic proposal for anhydro bond cleavage. The $^{1}C_4$ chair conformation is observed for **1** bound to LsLGK.[1a]

bridization (sp$^3 \rightarrow$ sp^2) at the anomeric carbon atom.[27] The implied coplanar orientation of C5, O5, C1, and C2 can only be adopted by two boat ($B_{2,5}$ and $^{2,5}B$), two half-chair ($^{4}H_3$ and $^{3}H_4$), and four envelope (^{3}E, E_3, ^{4}E, and E_4) conformations.[26b, 27a] Considering the conformational itineraries utilized by glycoside hydrolases, reviewed in refs. [26] and [27], we suggest a catalytic reaction coordinate for LGK, as shown in Figure 4.

Among the sugar ring conformations directly accessible from the $^{1}C_4$ chair (Figure 4 A), only the $^{3}H_4$ and the closely related ^{3}E and E_4 conformations can accommodate oxocarbenium ion character. A $^{3}H_4$ transition state is tentatively suggested for the LGK reaction (Figure 4 B). The $^{3}H_4$ half-chair limits ring strain from the 1,6-anhydro linkage and has typically a somewhat lower energy than the ^{3}E and E_4 conformations.[26b] It is structurally close to the $^{1}C_4$ chair of **1** observed in the substrate-bound structure of the LsLGK.[1a] Its formation in the transition state would be in line with the principle of least nuclear motion in catalysis. The $^{1}C_4 \rightarrow {}^{3}H_4^{+} \rightarrow {}^{3}S_1$ itinerary suggested for the LsLGK (Figure 4 B) has precedence from inverting mannosidases of glycoside hydrolase families GH 47 and GH 134.[28] It is also found in retaining α-fucosidases and sialida-

67

ses.[26b, 29] In the LGK reaction, conversion of α-**2** from the post-catalytic $^3S_1$ conformation into the most stable $^4C_1$ conformation might happen on the enzyme or in solution. An itinerary via E_1, or alternatively ^OE, is feasible.[25b] The large barrier for conversion of the $^4C_1$ conformation into the $^3S_1$ conformation (≈ 32 kJ mol^{-1} for α-D-glucose)[25b] would restrict enzymatic reaction of **2** or **3** to **1**.

Conclusion

Conversion of **1** into **2** by the levoglucosan kinase from *L. starkeyi* (LsLGK) involves configurational inversion at the anomeric carbon atom, β (equatorial) in the substrate and α (axial) in the product. This stereochemical course is consistent with the hybrid kinase–hydrolase mechanism depicted in Figure 4. In this mechanism, opening of the β-1,6-anhydro ring is promoted by nucleophilic attack of O6 on the terminal γ phosphorus atom of ATP. The water nucleophile reacts under general base catalytic assistance from Asp212. The transition state is expected to have substantial oxocarbenium ion like character. A hypothetical conformational itinerary for conversion of the $^1C_4$ chair of **1** into the $^4C_1$ chair of α-**2** involves a $^3H_4^{\ddagger}$ half-chair transition state and a $^3S_1$ skew-boat conformation. Absence of detectable reaction of **2**, or its "activated" analogue **3**, to yield **1** involves a large energy barrier for converting the $^4C_1$ chair into the $^3S_1$ conformation. Lack of reactivity of **4** might reflect inability to adopt the $^3S_1$ conformation. Besides the mechanistic analysis, new biochemical data for the LsLGK provide an advanced knowledge basis for this interesting class of enzymes.

Experimental Section

Materials: ATP, ADP, (anhydro)sugars, and 2,3,4,6-tetra-*O*-acetyl-α-D-glucopyranosyl fluoride were from Carbosynth (Berkshire, UK). Phosphoenolpyruvic acid (PEP), D-glucal, pyruvate kinase from rabbit muscle, hexokinase type III from *Saccharomyces cerevisiae* (25 U mg^{-1}), and glucose-6-phosphate dehydrogenase (G6P-DH) from *Leuconostoc mesenteroides* were from Sigma–Aldrich.

Preparation of LsLGK: Detailed procedures for creation of LsLGK expression strains, recombinant production in *E. coli* and protein purification are found in the Supporting Information. In short, for expression with N- and C-terminally fused His-tag the codon-optimized LsLGK gene (GenBank EU751287) was inserted into the pQE-30 and pET-41b(+) vectors, respectively. LsLGK from overnight expression in *E. coli* BL21-Gold (DE3) was affinity purified by Ni-NTA chromatography.

LsLGK activity assays: Either conversion of ATP into ADP was monitored by HPLC or formation of **2** was quantified photometrically. Aliquots withdrawn for HPLC analysis were stopped by addition of an equal volume of acetonitrile. ADP and ATP separation on a Kinetex C$_{18}$ column (5 µm, 100 Å, 50×4.6 mm; Phenomenex, Aschaffenburg, Germany) is described elsewhere in detail.[30] A flow rate of 2 mL min^{-1} [40 mM tetra-*n*-butylammonium bromide (TBAB), 20 mM KH$_2$PO$_4$, pH 5.9] was applied, and nucleotides were quantified by UV detection at $\lambda = 259$ nm. G6P-DH was used to couple synthesis of **2** to equimolar formation of NADH, which was spectrophotometrically quantified at $\lambda = 340$ nm ($\varepsilon = 6.22$ mM^{-1} cm^{-1}). Unless mentioned otherwise, reaction mixtures for continuous measurement of **2** formation contained **1**

(300 mM), ATP (2.5 mM), NAD$^+$ (1.4 mM), LsLGK (0.02 mg mL^{-1}), G6P-DH (10 U mL^{-1}), MgCl$_2$ (20 mM), BSA (1 mg mL^{-1}), and HEPES (50 mM, pH 7.8). Conversions were performed in microplates at 30°C and started by the addition of **1**. Alternatively, the assay was used discontinuously. To quantify **2**, aliquots were withdrawn from LsLGK reactions, and the enzyme was inactivated by heating to 95°C for 5 min. After removal of precipitated protein by centrifugation, samples were diluted tenfold with phosphate buffer (50 mM, pH 7.0), containing NAD$^+$ (1.4 mM), MgCl$_2$ (10 mM), and ethylenediaminetetraacetic acid (EDTA; 10 mM). The concentration of **2** was inferred from differences in absorption at $\lambda = 340$ nm before and after incubation with G6P-DH (10 U mL^{-1}) for 30 min at 30°C.

Stability of LsLGK: Verification of proper folding of the LsLGK by far-UV CD spectroscopy and evaluation of enzyme stabilization by DSF are described in detail in the Supporting Information. Time-dependent inactivation was tested by measuring LsLGK activity with the continuous G6P-DH assay after various incubation times ranging from 1 to 24 h. Unless mentioned otherwise, LsLGK (1 mg mL^{-1}) was incubated at 25°C in HEPES buffer (20 mM, pH 7.5), containing NaCl (50 mM) and dithiothreitol (DTT, 0.5 mM). An increase in temperature from 25 to 30°C and the addition of BSA (1 and 2 mg mL^{-1}) were evaluated. The effect of pH was tested in the presence of BSA (1 mg mL^{-1}). 2-(*N*-morpholino)ethanesulfonic acid (MES; 50 mM, pH 6.4), HEPES (50 mM, pH 7.9), or glycine (50 mM, pH 7.0 and 9.6) was used. To determine the operational stability of LsLGK, time courses for **2** formation were analyzed in the presence of 0 to 5 mg mL^{-1} BSA. Compound **1** (300 mM) and ATP (2.5 mM) were converted by LsLGK (0.01 mg mL^{-1}) at 25°C. Reactions were buffered in HEPES (50 mM; pH 7.8) containing MgCl$_2$ (20 mM). Throughout the conversions, samples were withdrawn, and the concentration of **2** was determined by the discontinuous G6P-DH assay.

Kinetic characterization of LsLGK: Unless indicated otherwise, reactions were performed at 30°C and pH 7.8. Reaction mixtures contained LsLGK 0.02 mg mL^{-1}), HEPES (50 mM), MgCl$_2$ (20 mM), and BSA (5 mg mL^{-1}). If not mentioned elsewise, ion-pairing HPLC was used to monitor conversion of ATP into ADP. Initial reaction rates were determined from measurements at four distinct points in time.

A pH–activity profile was recorded by following conversion of **1** (100 mM) and ATP (2 mM) by LsLGK (0.06 mg mL^{-1}) with the discontinuous G6P-DH assay. Reactions contained MgCl$_2$ (10 mM) and buffer (50 mM; MES: pH 5.6–6.8, HEPES: pH 6.6–8.3, glycine pH 8.7–10.5). The influence of the MgCl$_2$ concentration (1–75 mM) was tested in reactions containing **1** (500 mM) and ATP (2.5 mM). The K_m of the LsLGK for ATP was determined in the presence of **1** (500 mM) and BSA (1 mg mL^{-1}). The MgCl$_2$ concentration was adjusted to the ATP level. 0.02–1 mM ATP/10 mM MgCl$_2$; 2 mM ATP/20 mM MgCl$_2$; 5–20 mM ATP/50 mM MgCl$_2$. To determine the K_m for **1** in the presence of ATP (2.5 mM) and MgCl$_2$ (10 mM), the concentration of **1** was varied from 10 to 750 mM. Product inhibition by ADP (0.1–5 mM) or **2** (10–190 mM) was studied with ATP (2.5 mM) and **1** (300 mM). The inhibiting effect of monosaccharides was studied with the continuous G6P-DH assay **1** (300 mM), ATP (2.5 mM), BSA (1 mg mL^{-1}). Sugar concentrations were 10 times higher than the corresponding K_m of *S. cerevisiae* hexokinase:[31] 3.3 mM fructose, 1.2 mM glucose, or 0.5 mM mannose.

Substrate specificity and error hydrolysis of ATP: Phosphorylation of various (anhydro)sugars by the LsLGK was evaluated: **1**, D-glucose, 1,6-anhydro-β-D-mannopyranose, 1,6-anhydro-β-D-galactopyranose, 1,6-anhydro-β-D-cellobiose, and 1,6-anhydro-β-D-mal-

68

tose. At the same time the influence of these sugars on ATP hydrolysis was tested. Reaction mixtures contained sugar (100 mM), ATP (2 mM), MgCl$_2$ (10 mM), and HEPES (50 mM, pH 7.8). Conversions at 30 °C for 24 h were started by the addition of LsLGK (1 mg mL^{-1}). Control reactions contained either no sugar or no LsLGK. Conversion of ATP into ADP was quantified by HPLC, and release of inorganic phosphate was measured spectrophotometrically at $\lambda = 850$ nm, as reported by Saheki et al.[32]

Real-time in situ NMR spectroscopy: Reaction mixtures contained **1** (100 mM), ATP (10 mM), MgCl$_2$ (20 mM), LsLGK (2 mg mL^{-1}), and BSA (1 mg mL^{-1}). The pD was set to 7.0, and the temperature was 10 °C. Spectra were recorded with a Varian (Agilent) INOVA 500-MHz spectrometer (Agilent Technologies) using the VNMRJ 2.2D software. ^{1}H NMR spectra were measured at 499.98 MHz with a 5 mm indirect detection PFG probe. The water signal was presaturated by a shaped pulse by using a standard presaturation sequence: relaxation delay 2 s, 90° proton pulse, acquisition time 2.048 s, spectral width 8 kHz, number of points 32 k. For integration of the signal of the β-anomer of **2**, the clearly separated half of the doublet signal was integrated, and the resulting area was multiplied by two.

Reaction equilibrium: Concentrations of **1** (2–100 mM) and ATP (2–10 mM) were varied, resulting in ratios of **1**/ATP ranging from 0.7 to 16.7. Reaction mixtures contained LsLGK (3 mg mL^{-1}), BSA (1 mg mL^{-1}), MgCl$_2$ (20 mM), and HEPES (100 mM, pH 7.8). Over the 24 h long conversions at 30 °C, aliquots were withdrawn at certain times. The concentrations of ADP and ATP were determined by HPLC, and **2** was quantified photometrically by the discontinuous G6P-DH assay. Residual **1** was calculated from initially applied **1** and the amount of **2** formed.

Reversal of the LsLGK reaction: To identify possible conditions for the synthesis of **1**, various concentrations of **2** (1–150 mM) and ADP (0.2–20 mM) were tested. Also, the Mg^{2+} concentration was varied from 1 to 50 mM, and a pH range of 6 to 9 was investigated. Reactions were performed at 25 °C and 300 rpm in the presence of BSA (1 mg mL^{-1}). Conversions were started by the addition of LsLGK (0.5–1.1 mg mL^{-1}) and were followed for up to 24 h. HPLC analysis was used to detect **1**. A HPX-87H column (Bio-Rad, Vienna, Austria) was operated at a flow rate of 0.8 mL min^{-1} (5 mM H$_2$SO$_4$, 65 °C), and refractive index detection was employed. For more sensitive detection of **1**, selected samples were also analyzed by HPAEC-PAD, as described in detail in the Supporting Information.

Synthesis and conversion of 6-phospho-α-D-glucopyranosyl fluoride (3): Preparation of α-D-glucopyranosyl fluoride from 2,3,4,6-tetra-O-acetyl-α-D-glucopyranosyl fluoride is described in the Supporting Information. α-D-Glucopyranosyl fluoride (≈ 100 mM) was used in a reaction with ATP (100 mM), MgCl$_2$ (100 mM), and hexokinase (4 mg mL^{-1}) in D$_2$O. The pD was adjusted to 7.0 with concentrated NaOH. ^{19}F NMR spectroscopy analysis showed that 23 mM **3** was formed after 24 h of incubation at 30 °C. LsLGK (1.5 mg mL^{-1}) was added, and after another 24 h at 25 °C, the reaction mixture was analyzed by ^{1}H NMR spectroscopy.

Synthesis and conversion of 6-phospho-D-glucal (4): The synthesis of **4** by hexokinase-catalyzed phosphorylation of D-glucal was adopted from Chenault et al.[33] Anion-exchange chromatography and barium precipitation were applied to isolate **4**. Preparation of **4** is described in detail in the Supporting Information.

Conversion of **4** (160 mM) and ADP (10 mM) by LsLGK (10 mg mL^{-1}) in D$_2$O was followed by in situ ^{1}H NMR spectroscopy. The reaction was buffered to pD 7.8 with phosphate buffer (10 mM) containing

MgCl$_2$ (20 mM) and BSA (1 mg mL^{-1}). For ATP removal, AMP (90 mM) and EcADK (1 mg mL^{-1}) were added (see the Supporting Information for details on EcADK preparation). Conversion at 30 °C was monitored for 24 h with a Varian INOVA 500 MHz spectrometer, as described for the canonical LsLGK reaction by recording one spectrum per hour.

Acknowledgements

Conflict of Interest

The authors declare no conflict of interest.

[1] a) J.-P. Bacik, J. R. Klesmith, T. A. Whitehead, L. R. Jarboe, C. J. Unkefer, B. L. Mark, R. Michalczyk, *J. Biol. Chem.* **2015**, *290*, 26638–26648; b) Y. Kitamura, Y. Abe, T. Yasui, *Agric. Biol. Chem.* **1991**, *55*, 515–521.

[2] T. Uehara, K. Suefuji, N. Valbuena, B. Meehan, M. Donegan, J. T. Park, *J. Bacteriol.* **2005**, *187*, 3643–3649.

[3] a) J. Dai, Z. Yu, Y. He, L. Zhang, Z. Bai, Z. Dong, Y. Du, H. Zhang, *World J. Microbiol. Biotechnol.* **2009**, *25*, 1589–1595; b) J. Ning, Z. Yu, H. Xie, H. Zhang, G. Zhuang, Z. Bai, S. Yang, Y. Jiang, *World J. Microbiol. Biotechnol.* **2008**, *24*, 15–22.

[4] a) X. Zhuang, H. Zhang, *Protein Expression Purif.* **2002**, *26*, 71–81; b) H.-J. Xie, X.-L. Zhuang, H.-X. Zhang, Z.-H. Bai, H.-Y. Qi, *FEMS Microbiol. Lett.* **2005**, *251*, 313–319.

[5] Y. Kitamura, T. Yasui, *Agric. Biol. Chem.* **1991**, *55*, 523–529.

[6] H. Xie, X. Zhuang, Z. Bai, H. Qi, H. Zhang, *World J. Microbiol. Biotechnol.* **2006**, *22*, 887–892.

[7] a) M. K. Dowd, A. D. French, P. J. Reilly, *Carbohydr. Res.* **1994**, *264*, 1–19; b) L. Smrčok, M. Sládkovičová, V. Langer, C. C. Wilson, M. Koóš, *Acta Crystallogr. Sect. B* **2006**, *62*, 912–918.

[8] a) L. R. Jarboe, Z. Wen, D. Choi, R. C. Brown, *Appl. Microbiol. Biotechnol.* **2011**, *91*, 1519–1523; b) Z. U. Islam, Y. Zhisheng, E. B. Hassan, C. Dongdong, Z. Hongxun, *J. Ind. Microbiol. Biotechnol.* **2015**, *42*, 1557–1579; c) H. B. Mayes, M. W. Nolte, G. T. Beckham, B. H. Shanks, L. J. Broadbelt, *ACS Sustainable Chem. Eng.* **2014**, *2*, 1461–1473.

[9] a) J. Lian, M. Garcia-Perez, S. Chen, *Bioresour. Technol.* **2013**, *133*, 183–189; b) Z. Yang, Z. Bai, H. Sun, Z. Yu, X. Li, Y. Guo, H. Zhang, *Microb. Cell Fact.* **2014**, *13*, 182; c) X. L. Zhuang, H. X. Zhang, J. Z. Yang, H. Y. Qi, *Bioresour. Technol.* **2001**, *79*, 63–66.

[10] a) J.-P. Bacik, G. E. Whitworth, K. A. Stubbs, A. K. Yadav, D. R. Martin, B. A. Bailey-Elkin, D. J. Vocadlo, B. L. Mark, *J. Biol. Chem.* **2011**, *286*, 12283–12291; b) J. P. Bacik, L. R. Jarboe, *IUBMB life* **2016**, *68*, 700–708; c) G. Davies, B. Henrissat, *Structure* **1995**, *3*, 853–859; d) D. Pollard-Knight, A. Cornish-Bowden, *Mol. Cell. Biochem.* **1982**, *44*, 71–80.

[11] J.-P. Bacik, M. Tavassoli, T. R. Patel, S. A. McKenna, D. J. Vocadlo, M. Khajehpour, B. L. Mark, *J. Biol. Chem.* **2014**, *289*, 4504–4514.

[12] a) M. L. Sinnott, *Chem. Rev.* **1990**, *90*, 1171–1202; b) D. L. Zechel, S. G. Withers, *Curr. Opin. Chem. Biol.* **2001**, *5*, 643–649.

69

[13] K. J. Schray, S. J. Benkovic, *Acc. Chem. Res.* **1978**, *11*, 136–141.

[14] J. R. Klesmith, J.-P. Bacik, R. Michalczyk, T. A. Whitehead, *ACS Synth. Biol.* **2015**, *4*, 1235–1243.

[15] D. M. Jacobsen, Z.-Q. Bao, P. O'Brien, C. L. Brooks, III, M. A. Young, *J. Am. Chem. Soc.* **2012**, *134*, 15357–15370.

[16] A. Kaji, S. P. Colowick, *J. Biol. Chem.* **1965**, *240*, 4454–4462.

[17] E. Gernert, A. S. Keston, *Arch. Biochem. Biophys.* **1974**, *161*, 420–425.

[18] K. J. Schray, E. E. Howell, *Arch. Biochem. Biophys.* **1978**, *189*, 102–105.

[19] E. A. Robbins, P. D. Boyer, *J. Biol. Chem.* **1957**, *224*, 121–135.

[20] a) J. W. Lawson, R. L. Veech, *J. Biol. Chem.* **1979**, *254*, 6528–6537; b) A. Gutmann, B. Nidetzky, *Adv. Synth. Catal.* **2016**, *358*, 3600–3609.

[21] a) Y. Tanaka, W. Tao, J. S. Blanchard, E. J. Hehre, *J. Biol. Chem.* **1994**, *269*, 32306–32312; b) S. J. Williams, S. G. Withers, *Carbohydr. Res.* **2000**, *327*, 27–46; c) E. J. Hehre in *Enzymatic Degradation of Insoluble Carbohydrates, Vol. 618* (Eds.: J. N. Saddler, M. H. Penner), American Chemical Society, Washington, D.C., **1996**, pp. 66–78; d) P. Wildberger, L. Brecker, B. Nidetzky, *Carbohydr. Res.* **2012**, *356*, 224–232; e) G. Legler, K. R. Roeser, H. K. Illig, *Eur. J. Biochem.* **1979**, *101*, 85–92; f) B. Lougheed, H. D. Ly, W. W. Wakarchuk, S. G. Withers, *J. Biol. Chem.* **1999**, *274*, 37717–37722; g) J. C. Díaz Arribas, A. G. Herrero, M. Martín-Lomas, F. J. Cañada, S. He, S. G. Withers, *Eur. J. Biochem.* **2000**, *267*, 6996–7005.

[22] a) A. G. Santana, G. Vadlamani, B. L. Mark, S. G. Withers, *Chem. Commun.* **2016**, *52*, 7943–7946; b) M. Petricevic, L. F. Sobala, P. Z. Fernandes, L. Raich, A. J. Thompson, G. Bernardo-Seisdedos, O. Millet, S. Zhu, M. Sollogoub, J. Jiménez-Barbero, C. Rovira, G. J. Davies, S. J. Williams, *J. Am. Chem. Soc.* **2017**, *139*, 1089–1097.

[23] M. Brune, R. Schumann, F. Wittinghofer, *Nucleic Acids Res.* **1985**, *13*, 7139–7151.

[24] G. S. Hammond, *J. Am. Chem. Soc.* **1955**, *77*, 334–338.

[25] a) D. Cremer, J. A. Pople, *J. Am. Chem. Soc.* **1975**, *97*, 1354–1358; b) H. B. Mayes, L. J. Broadbelt, G. T. Beckham, *J. Am. Chem. Soc.* **2014**, *136*, 1008–1022.

[26] a) G. J. Davies, V. M.-A. Ducros, A. Varrot, D. L. Zechel, *Biochem. Soc. Trans.* **2003**, *31*, 523–527; b) G. Speciale, A. J. Thompson, G. J. Davies, S. J. Williams, *Curr. Opin. Struct. Biol.* **2014**, *28*, 1–13.

[27] a) A. Ardèvol, C. Rovira, *J. Am. Chem. Soc.* **2015**, *137*, 7528–7547; b) G. J. Davies, A. Planas, C. Rovira, *Acc. Chem. Res.* **2012**, *45*, 308–316.

[28] a) K. Karaveg, A. Siriwardena, W. Tempel, Z.-J. Liu, J. Glushka, B.-C. Wang, K. W. Moremen, *J. Biol. Chem.* **2005**, *280*, 16197–16207; b) A. J. Thompson, J. Dabin, J. Iglesias-Fernández, A. Ardèvol, Z. Dinev, S. J. Williams, O. Bande, A. Siriwardena, C. Moreland, T. C. Hu, D. K. Smith, H. J. Gilbert, C. Rovira, G. J. Davies, *Angew. Chem. Int. Ed.* **2012**, *51*, 10997–11001; *Angew. Chem.* **2012**, *124*, 11159–11163; c) Y. Jin, M. Petricevic, A. John, L. Raich, H. Jenkins, L. Portela De Souza, F. Cuskin, H. J. Gilbert, C. Rovira, E. D. Goddard-Borger, S. J. Williams, G. J. Davies, *ACS Cent. Sci.* **2016**, *2*, 896–903.

[29] a) A. Lammerts van Bueren, A. Ardèvol, J. Fayers-Kerr, B. Luo, Y. Zhang, M. Sollogoub, Y. Blériot, C. Rovira, G. J. Davies, *J. Am. Chem. Soc.* **2010**, *132*, 1804–1806; b) G. Sulzenbacher, C. Bignon, T. Nishimura, C. A. Tarling, S. G. Withers, B. Henrissat, Y. Bourne, *J. Biol. Chem.* **2004**, *279*, 13119–13128; c) M. A. F. Amaya, A. G. Watts, I. Damager, A. Wehenkel, T. Nguyen, A. Buschiazzo, G. Paris, A. C. Frasch, S. G. Withers, P. M. Alzari, *Structure* **2004**, *12*, 775–784.

[30] S. T. Kulmer, A. Gutmann, M. Lemmerer, B. Nidetzky, *Adv. Synth. Catal.* **2017**, *359*, 292–301.

[31] R. Fernandez, P. Herrero, F. Moreno, *J. Gen. Microbiol.* **1985**, *131*, 2705–2709.

[32] S. Saheki, A. Takeda, T. Shimazu, *Anal. Biochem.* **1985**, *148*, 277–281.

[33] H. K. Chenault, R. F. Mandes, *Bioorg. Med. Chem.* **1994**, *2*, 627–629.

70

Supporting Information

Biochemical Characterization and Mechanistic Analysis of the Levoglucosan Kinase from *Lipomyces starkeyi*

Christina Rother[+],[a] Alexander Gutmann[+],[a] Ramakrishna Gudiminchi,[a, b] Hansjörg Weber,[c] Alexander Lepak,[a] and Bernd Nidetzky*[a, b]

cbic_201700587_sm_miscellaneous_information.pdf

1. Experimental Section

1.1 Preparation of LsLGK

The LsLGK gene (GenBank EU751287) was codon optimized for *E. coli* expression and purchased as gBlocks® Gene Fragment from Integrated DNA Technologies (Coralville, IA, US). The gene was ligated into the *Bam*HI and *Hind*III sites of the pQE-30 plasmid for expression with N-terminally fused His-Tag. For C-terminal fusion of the His-tag the stop codon was removed from the LsLGK gene before it was inserted into the *Nde*I and *Not*I sites of the pET-41b(+) expression vector. Expression strains were created by transforming electro-competent *E. coli* BL21-Gold (DE3) cells with the described plasmids.

LsLGK expression strains were grown at 37°C and 100 rpm in LB medium. 100 µg mL^{-1} ampicillin and 50 µg mL^{-1} kanamycin were added for strains harboring the pQE-30 and pET-41b(+) plasmid, respectively. At an OD$_{600}$ of 0.7 enzyme expression was induced by addition of 1 mM IPTG (isopropyl-D-1-thiogalactopyranoside) and the temperature was lowered to 30°C. After 3 h cells were harvested by centrifugation (20 min, 5000 rpm, 4°C). Cell pellets were dissolved in 50 mM HEPES, pH 7.8 and frozen at -20°C until cell disruption by sonication.

An ÄktaPrime plus system (GE Healthcare, Vienna, Austria) equipped with a 5 mL HisTrap HP Ni Sepharose column (GE Healthcare) was used for LsLGK purification from the cleared lysate. A constant flow rate of 1 mL min^{-1} was applied. 20 mM Na-PO$_4$ buffer, pH 7.4 containing 500 mM NaCl and 20 mM imidazole was used to equilibrate the column and wash off unbound protein. LsLGK was eluted by a 50 mL long gradient from 20 to 250 mM imidazole. The concentrated elution fractions were dialyzed against 20 mM HEPES buffer, pH 7.5 containing 50 mM NaCl and 0.5 mM DTT before storage in small aliquots at -20°C. Protein concentrations were determined photometrically at 280 nm and protein purities were assessed by SDS-PAGE.

Aggregation of LsLGK with N-terminal His-tag was analyzed by size exclusion chromatography with a HiLoad 16/60 Superdex 200 prep grade column (GE Healthcare). 50 mM HEPES buffer, pH 7.2 with 150 mM NaCl was used at a constant flow rate of 1 mL min^{-1}.

1.2 Circular dichroism (CD) spectroscopy

Far-UV CD spectra were acquired on a Jasco J-175 spectropolarimeter (Jasco Analytical Instruments, Groß-Umstadt, Germany) over a spectral window of 190-260 nm. 0.5 mg mL^{-1} LsLGK were dissolved in 20 mM HEPES, pH 7.5. Secondary structures were calculated by the DichroWeb software using the CDSSTR method and the SMP180 reference set.[1]

1.3 Differential scanning fluorimetry (DSF)

Potential stabilizing effects of substrates, products and MgCl$_2$ were analyzed by DSF. Test solutions contained 0.75 µg mL^{-1} SYPRO® Orange, 0.75 mg mL^{-1} LsLGK and 50 mM HEPES, pH 7.8. Unless mentioned otherwise, 10 mM MgCl$_2$ was added. The effect of ADP alone (2 mM) and in combination with **2** (D-glucose-6-phosphate) (20 mM each) was analyzed. Addition of 100 mM **1** (levoglucosan) or **2** was studied. Finally supplementation of 2 mM ATP was tested with and without 10 mM MgCl$_2$. Using a real-time PCR instrument the temperature was increased in a range of 20-95°C every 5 s by 0.5°C. Fluorescence was continuously measured as ratio of extinction at 492 nm and emission at 610 nm.

2

1.4 Calculation of K_i for ADP

Conversion of 300 mM **1** and 2.5 mM ATP by 0.02 mg mL^{-1} LsLGK was supplemented with 0.1-5 mM ADP to study product inhibition by ADP. K_i was determined by fitting the experimental data with Formula 1 which describes competitive enzyme inhibition. K_m for ATP (1.0 mM) and r_{max} (16.2 µmol min^{-1} mg^{-1}) were extracted from Figure 1B. The r-values are reaction rates at a given ADP concentration (c_{ADP}). The ATP concentration (c_{ATP}) was 2.5 mM.

$$(1) \qquad r = \frac{r_{max} \cdot c_{ATP}}{K_m \left(1 + \frac{c_{ADP}}{K_i}\right) + c_{ATP}}$$

1.5 Detection of 1 by HPAEC-PAD

High-performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) was applied for sensitive detection of **1** in samples obtained from reverse reactions of LsLGK. A Carbopac PA10 column (Thermo Fisher, Waltham, MA, US) was used for separation by isocratic elution with 40 mM NaOH at a flow rate of 0.8 mL min^{-1}. After each run the column was washed with 1 M sodium acetate (10 min), 100 mM NaOH (10 min) and 30 mM NaOH (20 min). The column was equilibrated with 40 mM NaOH for 10 min. Sample preparation for HPAEC-PAD involved removal of residual **2** and ADP by anion exchange chromatography. A 1 mL Toyopearl SuperQ-650 M column (Tosoh Bioscience, Tokyo, Japan) was mounted on an ÄKTA FPLC system (GE Healthcare). Uncharged **1** was eluted with water at 1 mL min^{-1}. After each run the column was washed with 2 M NaCl.

1.6 Preparation of α-D-glucopyranosyl fluoride

The α-D-glucopyranosyl fluoride was prepared from 2,3,4,6-tetra-O-acetyl-α-D-glucopyranosyl fluoride using a slightly adapted protocol from Steinmann et al.[2] Shortly, the peracetylated substrate was deprotected by Zemplén deacetylation[3] using sodium methoxide in methanol under stirring at 0°C. After neutralization to pH 7 using Amberlite IR-120 (10 min presoaked in MeOH), the solution was filtered through cotton before removal of solvents in oil pump vacuum.

1.7 Preparation of 6-phospho-D-glucal (4)

4 was synthesized by hexokinase catalyzed phosphorylation of D-glucal (1,5-anhydro-2-deoxy-D-$arabino$-hex-1-enitol).[4] 40 mM D-glucal and 40 mM ATP were converted by 1 mg mL^{-1} hexokinase in presence of 10 mM MgCl$_2$ and 50 mM HEPES, pH 7.5. After 30 h of incubation at 37°C ultrafiltration (10 kDa cut-off) was used to stop the reaction by removing hexokinase.

4 was purified by anion-exchange chromatography on a 5 mL SuperQ-650 M column. Water was used as mobile phase A and mobile phase B was 1 M sodium acetate, pH 4.2. The flow rate was fixed to 5 mL min^{-1}. Uncharged compounds were washed off with 20 mL of mobile phase A before **4** was separated from other phosphorylated compounds during a 120 mL long gradient from 0-25% B. Fractions were analyzed by HPLC to monitor nucleotides and TLC (thin layer chromatography) was used to detect D-glucal and **4**. Silica gel 60 F$_{254}$ plates (Merck, Darmstadt, Germany) were applied with 1-butanol:1-propanol:ethanol:H$_2$O (3:3:1:1) as mobile phase. Compounds were stained with orcinol (40 mg orcinol, 80 mL acetonitrile, 4 mL concentrated H$_2$SO$_4$). Fractions containing pure **4** were pooled and concentrated under reduced pressure.

Excess of acetate from anion-exchange chromatography was removed by barium precipitation of **4**. 1.5 equivalents of barium acetate and 4 volumes of ethanol were added to precipitate **4** overnight at -20°C. After centrifugation (5000 rpm, 4°C, 20 min) the precipitate was dried for 2 h at 40°C. The pellet was resuspended in water containing 1.5 equivalents of Na$_2$SO$_4$. After 2 h of incubation at 4°C precipitated Ba$_2$SO$_4$ was removed by centrifugation (5000 rpm, 4°C, 20 min). The supernatant was freeze-dried and product identity was confirmed by ^{1}H-NMR.

3

1.8 Preparation of adenylate kinase from *Escherichia coli* (EcADK)

The gene encoding *Ec*ADK (EC: 2.7.4.3; GenBank AAB40228.1) was amplified from the genomic DNA of *E. coli* DH5 alpha using the primers EcADK_NdeI_fw (catatgcgtatcattctgcttggcgctcc) and EcADK_XhoI_rv (ctcgagttagccgaggattttttccagatcagcg). For expression with N-terminally fused *Strep*-tag the gene was inserted into *Nde*I and *Hind*III sites of the expression vector pET-STRP3.[5] The expression strain was created by transformation of electro-competent *E. coli* BL21-Gold (DE3) cells. Enzyme expression in LB-medium supplemented with 50 µg mL^{-1} kanamycin was performed as described for LsLGK but overnight enzyme expression at 25°C was induced by 0.5 mM IPTG. EcADK was purified from cell free extract by Strep-tag affinity chromatography using a gravity flow *Strep*-Tactin® Sepharose® column (IBA GmbH, Göttingen, Germany).[6] The enzyme was buffer exchanged to 25 mM HEPES and stored at -20°C.
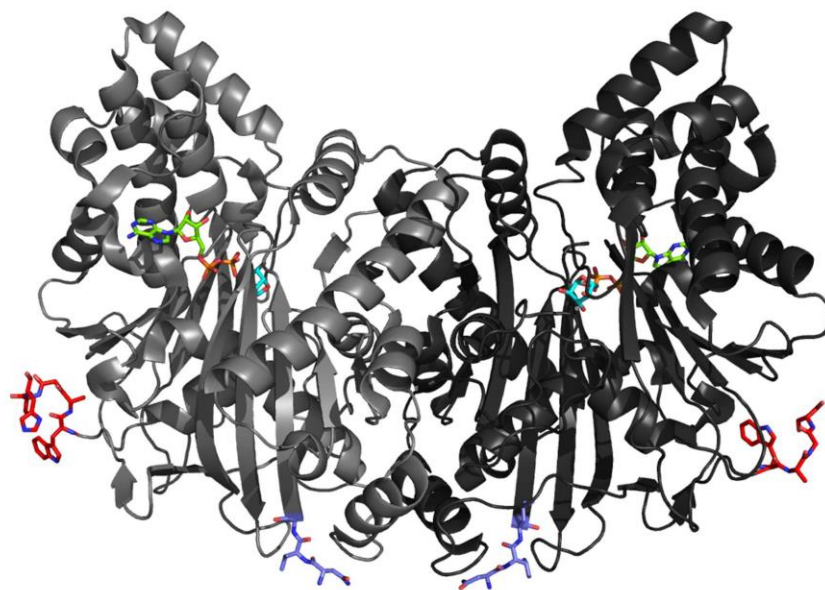
4

2. Results

Figure S1. The crystal structure of LsLGK (PDB ID: 4ZLU) is shown.[7] Chain A and B of the dimer are displayed in dark and light grey, respectively. The N-termini of both chains are marked in blue and the C-termini are colored in red. ADP and **1** bound to the active sites are shown in green and cyan, respectively.

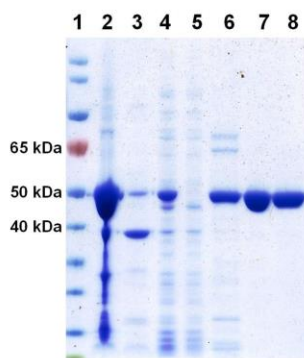**Figure S2.** Affinity purification of LsLGK with C-terminal His-tag (49.7 kDa) was analyzed by SDS PAGE. 1: PageRuler™ Prestained Protein Ladder (Thermo Scientific); 2-3: distinct insoluble fractions after cell lysis; 4: soluble fraction after cell lysis; 5: unbound protein from His-tag affinity purification; 6-8: distinct elution fractions from affinity His-tag purification; 7 and 8 were pooled and used to study LsLGK.
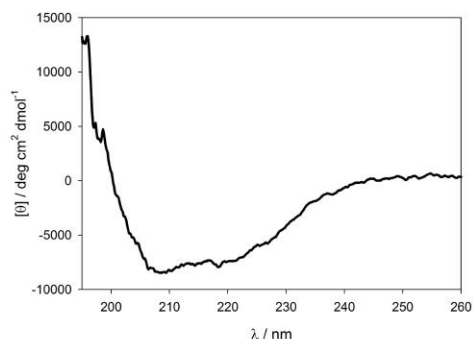
5

Figure S3. The far-UV CD spectrum of purified LsLGK is shown. 20 mM HEPES buffer, pH 7.5 was used to dissolve 0.5 mg mL^{-1} LsLGK. Analysis of the spectrum with DichroWeb suggested a composition of 44% α-helices, 17% β-strands, 14% turns and 25% unordered structures. This was in good agreement with experimental data from the LsLGK crystal structure (PDB ID: 4ZLU) which was composed of 48% α-helices, 24% β-strands and 28% unordered structures.
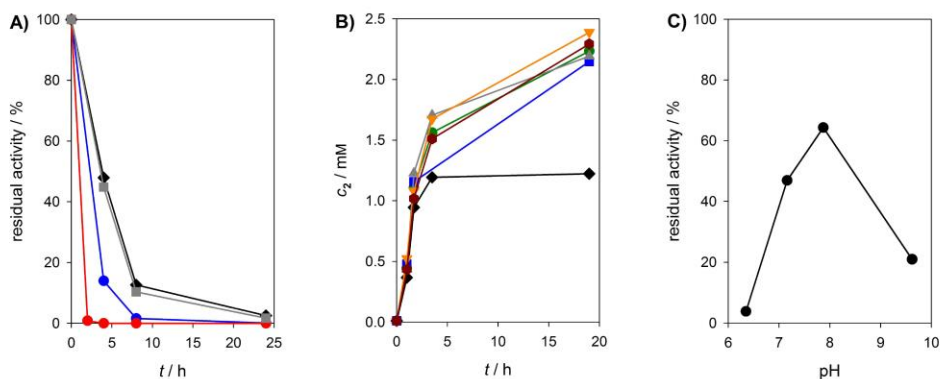

Figure S4. The influences of temperature, pH and BSA addition on stability of LsLGK were analyzed. A) Time courses of LsLGK (1 mg mL^{-1}) inactivation at pH 7.5 are shown. The temperature was 30 (red) or 25°C (blue). At 25°C also addition of BSA (grey: 1 mg mL^{-1}; black: 2 mg mL^{-1}) was tested. B) Time courses of **2** formation at various BSA concentrations were compared: 0 mg mL^{-1} (black), 0.5 mg mL^{-1} (grey), 1 mg mL^{-1} (blue), 2 mg mL^{-1} (green), 3.5 mg mL^{-1} (orange), 5 mg mL^{-1} (red). Reaction conditions: 300 mM **1**, 2.5 mM ATP, 0.01 mg mL^{-1} LsLGK, 20 mM MgCl$_2$, pH 7.8, 25°C. C) The pH dependency of LsLGK inactivation was studied in presence of 1 mg mL^{-1} BSA. After 2 h of incubation at 25°C residual activity from 1 mg mL^{-1} LsLGK was determined. Following buffers were used at a concentration of 50 mM: MES pH 6.4, glycine pH 7.0, HEPES pH 7.9, glycine pH 9.6.

6
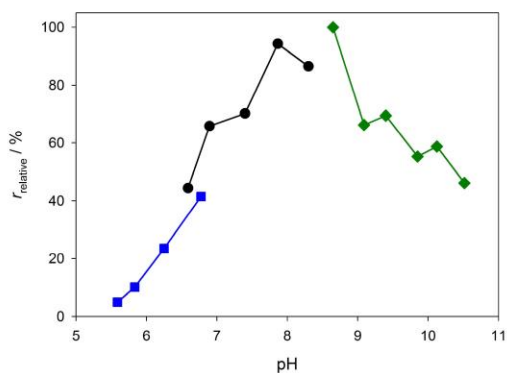
Figure S5. The pH-activity profile for synthesis of **2** by LsLGK is shown. Reactions were buffered by MES (pH 5.6–6.8; blue), HEPES (pH 6.6–8.3; black) or glycine (pH 8.7–10.5; green). Reaction conditions: 100 mM **1**, 2 mM ATP, 10 mM $MgCl_2$, 50 mM buffer, 0.06 mg mL^{-1} LsLGK, 5 mg mL^{-1} BSA, 30°C.

Figure S6. The K_m of LsLGK for **1** was determined to be 177 ± 36 mM. The solid line represents a hyperbolic fit of the experimental data and error bars show the standard deviation of three independent experiments. Reaction conditions: 2.5 mM ATP, 10 mM $MgCl_2$, 0.02 mg mL^{-1} LsLGK, 5 mg mL^{-1} BSA, pH 7.8.

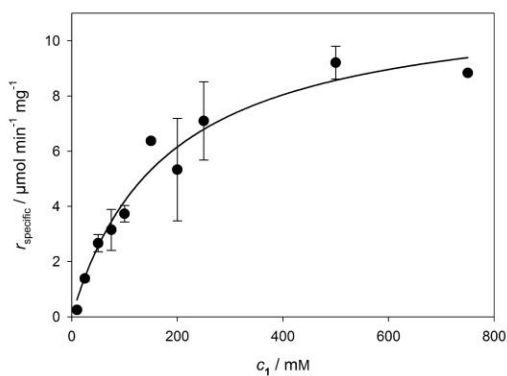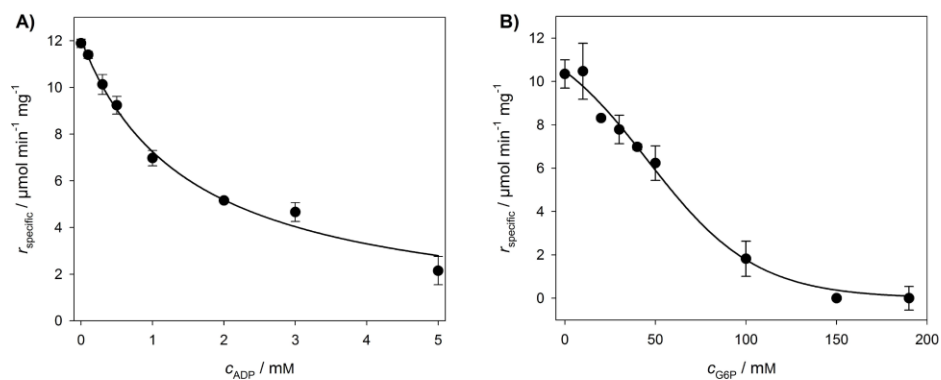**Figure S7.** The influence of product inhibition on LsLGK activity was tested (300 mM **1**, 2.5 mM ATP, 20 mM $MgCl_2$, 0.02 mg mL^{-1} LsLGK, 5 mg mL^{-1} BSA, pH 7.8, 30°C). A) ADP was a competitive inhibitor and a K_i of 0.45 ± 0.03 mM was calculated from the fit shown as a solid line. See the experimental section for details on fitting K_i. B) Inhibition by **2** was fitted by a sigmoidal curve (solid line). A half maximal inhibitory concentration (IC_{50}) of 56 ± 10 mM was calculated.

Figure S8. Inhibition of LsLGK by various hexoses was tested. Concentrations of D-fructose, D-glucose and D-mannose were 10 times higher than the corresponding K_m of *S. cerevisiae* hexokinase.[8] Reaction conditions: 300 mM **1**, 2.5 mM ATP, 20 mM $MgCl_2$, 0.02 mg mL^{-1} LsLGK, 1 mg mL^{-1} BSA, pH 7.8, 30°C.
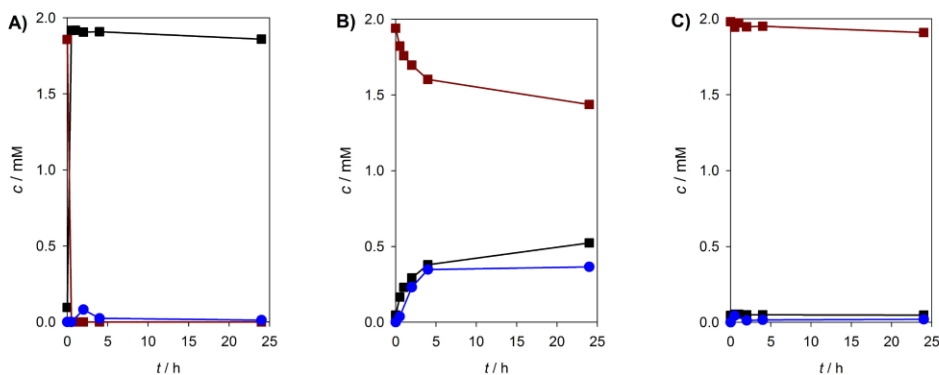
Figure S9. Time courses of ATP hydrolysis by LsLGK were studied (red: ATP; black: ADP; blue: free phosphate). A) ATP conversion by LsLGK in presence of **1**; B) ATP conversion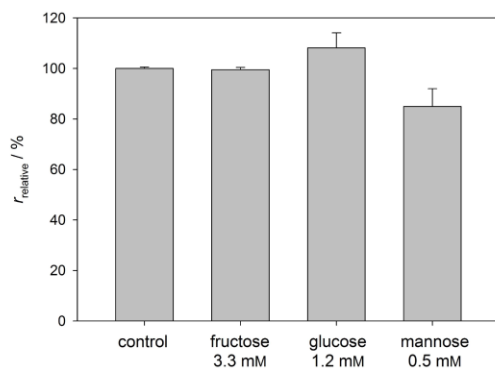 by LsLGK in absence of **1**; C) Incubation of **1** and ATP in absence of LsLGK. Reaction conditions: 0/100 mM **1**, 2 mM ATP, 0/1 mg mL^{-1} LsLGK, 10 mM MgCl$_2$, pH 7.8, 30°C.
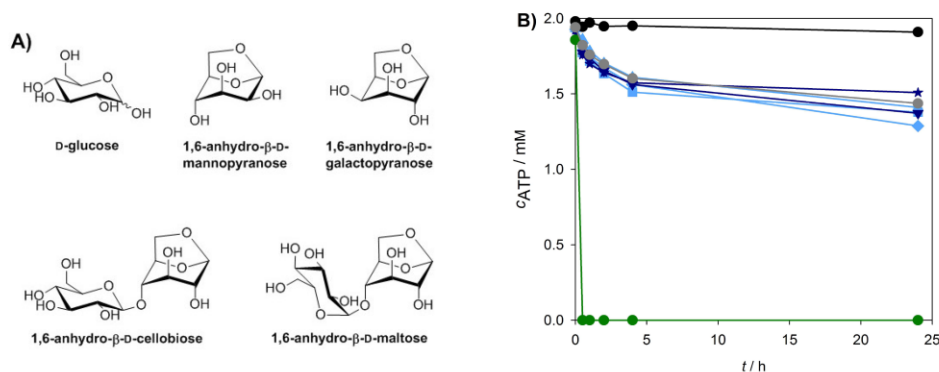
Figure S10. The canonical LsLGK reaction and error hydrolysis of ATP were investigated in presence of various analogues of **1**. A) The chemical structures of the tested substrates are shown. B) Time courses of ATP consumption were compared. 2 mM ATP and 100 mM of the respective sugar were converted with 1 mg mL^{-1} LsLGK (10 mM MgCl$_2$, pH 7.8, 30°C). **1** (green circle); monosaccharide analogues (light blue): D-glucose (diamond), 1,6-anhydro-β-D-mannopyranose (square), 1,6-anhydro-β-D-galactopyranose (triangle up); disaccharide analogues (dark blue): 1,6-anhydro-β-D-cellobiose (star), 1,6-anhydro-β-D-maltose (triangle down); no sugar (grey circle), no LsLGK (black circle).

9
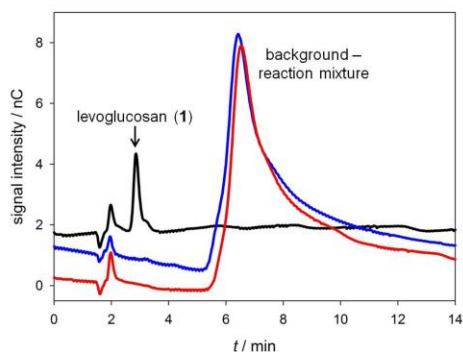
Figure S11. HPAEC-PAD analysis was applied to test if **1** can be formed by reversal of the LsLGK reaction (150 mM **2**, 10 mM ADP, 20 mM MgCl$_2$, 1.13 mg mL^{-1} LsLGK, 1 mg mL^{-1} BSA, 50 mM MES, pH 6.0). The reaction mixture was analyzed before (red) and after (blue) 24 h of incubation at 25°C. A solution of 0.3 mM **1** was used as positive control (black). To remove phosphate containing compounds all samples were passed through an anion exchange column prior to analysis.
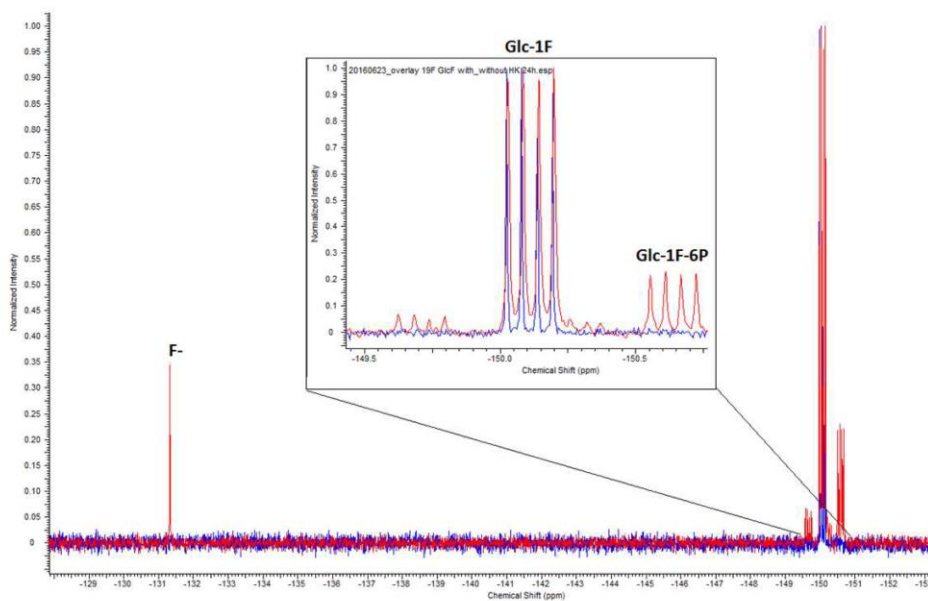
Figure S12. Hexokinase catalyzed phosphorylation of 100 mM α-D-glucopyranosyl fluoride to 6-phospho-α-D-glucopyranosyl fluoride (**3**) was monitored by ^{19}F-NMR (100 mM ATP, 100 mM MgCl$_2$, 4 mg mL^{-1} hexokinase, D$_2$O, pD 7.0). After 24 h of incubation at 30°C the reaction mixture (red) was compared with an equally treated negative control lacking hexokinase (blue). Only in presence of hexokinase signals for free fluoride and two additional sugar fluorides were observed. Around 23% of α-D-glucopyranosyl fluoride were phosphorylated to the desired main product **3** (-150.7 ppm).

10
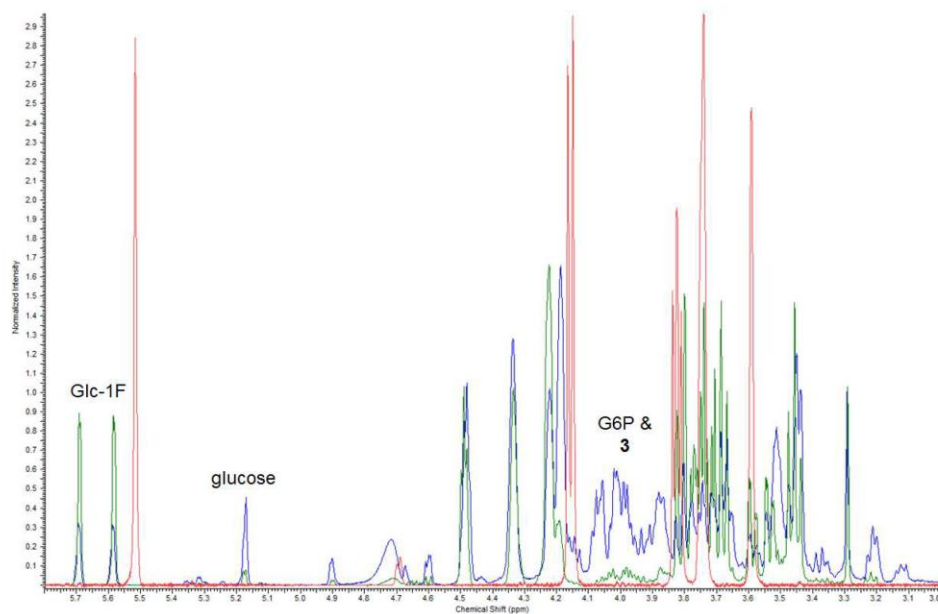
Figure S13. Synthesis of **1** from α-D-glucopyranosyl fluoride (Glc-1F) by consecutive conversion with hexokinase and LsLGK was monitored by ¹H-NMR. The spectrum of a levoglucosan standard (red) is overlaid with spectra of the reaction mixture before enzyme addition (green) and after incubation with hexokinase and LsLGK (blue). Phosphorylation of α-D-glucopyranosyl fluoride by hexokinase yielded about 23 mM **3** (~23% conversion, see Figure S12). To convert **3** to **1** 1.5 mg mL⁻¹ LsLGK were added. After 24h of incubation at 25°C signals for **3** and D-glucose-6-phosphate (G6P) were clearly visible (blue trace). However, the missing H-1 signal at 5.53 ppm shows that no **1** was formed.
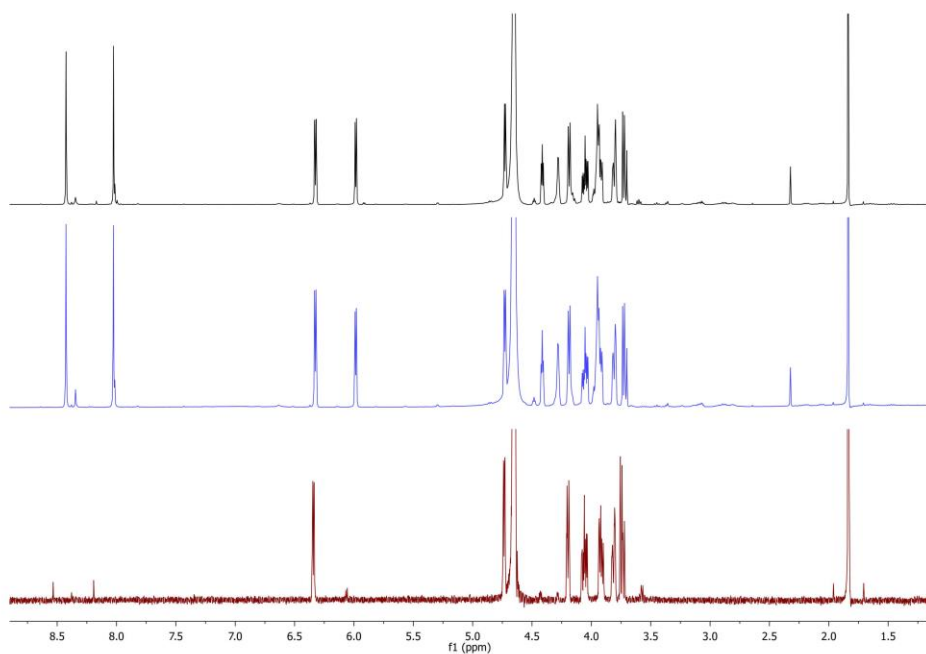
Figure S14. Attempted conversion of 160 mM **4** and 10 mM ADP with 10 mg mL^{-1} LsLGK was monitored by in situ ^{1}H-NMR. To prevent accumulation of ATP, 1 mg mL^{-1} EcADK and 90 mM AMP were added. The stacked ^{1}H-NMR spectra show the substrate **4** (red) and the reaction mixture before (blue) and after (black) 24 h of incubation at 30°C. Despite high enzyme concentration and long incubation time no significant formation of products was observed. The characteristic signals for the CH$_2$-protons of the anticipated product 2-deoxy-**1** would be clearly visible around 2 ppm.

References

[1] L. Whitmore, B. A. Wallace, *Nucleic Acids Res.* **2004**, *32*, W668-W673.

[2] A. Steinmann, J. Thimm, M. Matwiejuk, J. Thiem, *Macromolecules* **2010**, *43*, 3606-3612.

[3] G. Zemplén, E. Pacsu, *Ber. dtsch. Chem. Ges. A/B* **1929**, *62*, 1613-1614.

[4] H. K. Chenault, R. F. Mandes, *Bioorg. Med. Chem.* **1994**, *2*, 627-629.

[5] D. P. Dixon, T. Hawkins, P. J. Hussey, R. Edwards, *J. Exp. Bot.* **2009**, *60*, 1207-1218.

[6] A. Gutmann, B. Nidetzky, *Angew. Chem. Int. Ed.* **2012**, *51*, 12879-12883; *Angew. Chem.* **2012**, *124*, 13051–13056.

[7] J.-P. Bacik, J. R. Klesmith, T. A. Whitehead, L. R. Jarboe, C. J. Unkefer, B. L. Mark, R. Michalczyk, *J. Biol. Chem.* **2015**, *290*, 26638-26648.

[8] R. Fernandez, P. Herrero, F. Moreno, *J. Gen. Microbiol.* **1985**, *131*, 2705-2709.

13

Chapter 5:

New flavanol O-glycosides in grape and wine

New flavanol *O*-glycosides in grape and wine

Marie Zerbib[a], Jean-Paul Mazauric[b], Emmanuelle Meudec[b], Christine Le Guernevé[b], Alexander Lepak[d], Bernd Nidetzky[d], Véronique Cheynier[b], Nancy Terrier[c], Cédric Saucier[a,*]

[a] SPO, Univ Montpellier, INRA, Montpellier Supagro, Montpellier, France
[b] SPO, INRA, Montpellier Supagro, Univ Montpellier, Montpellier, France
[c] AGAP, INRA, CIRAD, Montpellier SupAgro, Univ Montpellier, Montpellier, France
[d] Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, Graz, Austria

ARTICLE INFO

ABSTRACT

The presence of monomeric and dimeric flavan-3-ol monohexosides was investigated in grapes and wines. Polyphenol extracts were prepared from grape seeds and skins (Syrah, Merlot, and Cabernet-Sauvignon) sampled at three different developmental stages. Different wines (Tannat, Alicante, Syrah, Merlot, and Grenache) were also studied. The different samples obtained were analyzed by UPLC-ESI-IT-MS. Specific molecular ions corresponding to flavan-3-ol hexosides were detected by using targeted molecular ions with specific *m/z* values: 451 for (*epi*)catechin hexosides, and 739 for procyanidin dimer hexosides. 4′-*O*-β-glucosyl-(+)-catechin and 7-*O*-β-glucosyl-(+)-catechin were obtained by using a glucosyl transferase from apple, UGT71A15, and their structures determined by NMR. These glucosylated monomers and the dimers were identified in all analyzed grape seed and several skin extracts at the different stages and in wines made from different varieties.

1. Introduction

Flavonoids are a category of polyphenols present in a wide variety of natural products such as grape, tea, and cocoa (Quideau, Deffieux, Douat-Casassus, & Pouysegu, 2011). They have health properties and are involved in organoleptic properties such as the color and the taste of wine and play a role in plant defense mechanisms (Corder et al., 2006). The common molecular backbone of these secondary metabolites is a C6-C3-C6 type. Flavan-3-ols, called flavanols here, are an important subclass of flavonoids. They are present in grape and wine as monomers or polymers called proanthocyanidins (PAs) or condensed tannins. The monomeric units are linked through C_4-C_6 or C_4-C_8 linkages (B type) with sometimes additional C2-O-C5 or C2-O-C7 bond (A type). In *Vitis vinifera* berries, four principal units of PAs are usually found: (+)-catechin, (−)-epicatechin, (−)-epigallocatechin and (−)-epicatechin gallate. They are synthesized in skins and seeds at the earlier developmental stages after flowering, and the monomer and polymer concentrations remain stable after *veraison*, the onset of ripening (Downey, Harvey, & Robinson, 2003; Terrier, Ollé, Verriès, & Cheynier, 2009).

The common flavonoid biosynthesis sequence is well described in the literature (Bogs et al., 2005; Fujita et al., 2005; Tanner & Kristiansen, 1993; Xie, Sharma, Paiva, Ferreira, & Dixon, 2003).

However, the chemical or biochemical mechanism of the polymerization steps generating PAs remains unknown. In 2008, Dixon and collaborators published a work about the expression of a glucosyltransferase, UGT72L1, in the seed coat of *Medicago truncatula*. This enzyme catalyzes the formation of epicatechin 3′-*O*-glucoside, named here E3′OG, in the cytoplasm and its over-expression results in increased accumulation of PA-like compounds in the vacuole (Pang et al., 2013). Further studies have suggested that E3′OG rather than free epicatechin is the substrate for the tonoplast-localized MATE, transporters MATE1 from *M. truncatula* and TT12 from *A. thaliana*, involved in PA biosynthesis (Zhao & Dixon, 2009). This glucosylated flavan-3-ol was postulated to be an intermediate in flavan-3-ol polymerization in *Medicago truncatula* (Pang, Peel, Sharma, Tang, & Dixon, 2008).

Polyphenols substituted with different hexoses (glucose, galactose, mannose, …), including stilbenes (Ali & Strommer, 2003; Versari, Parpinello, Tornielli, Ferrarini, & Giulivo, 2001), phenolic acids (Monagas, Bartholomé, & Gomez-Cordovés, 2007), flavonols and flavones (Castillo-Muñoz et al., 2009; Hmamouchi, Es-Safi, Lahrichi, Fruchier, & Essassi, 1996) have been described in the literature. Also, several flavan-3-ol glucosides were found in a wide range of plants : (+)-catechin-7-*O*-β-glucoside in elm (Kim et al., 2004), buckwheat (Smaali, Maugard, Limam, Legoy, & Marzouki, 2007; Watanabe, 1998),

cowpea seeds (Cui et al., 2016), barley and malt (Friedrich & Galensa, 2002), (+)-catechin 3-*O*-glucoside in lentils (Dueñas, Sun, Hernández, Estrella, & Spranger, 2003), (+)-catechin 5-*O*-, 7-*O*-, 3′-*O*-, and 4′-*O*-glucosides in Chinese peony (Tanaka, Kataoka, Tsuboi, & Kouno, 2000), (+)-catechin 7-*O*- β- and 4′-*O*-β-glucosides in Douglas-fir bark (Foo & Karchesy, 1989), (+)-catechin 5-*O*-β-glucoside and (−)-catechin 7-*O*-β-glucoside in rhubarb and *Rhaphiolepis umbellata* bark (Nonaka, Ezaki, Hayashi, & Nishioka, 1983). The characterization of flavan-3-ol hexosides was investigated in literature, and to this aim, some of these compounds were produced. In fact, different strategies have been used to obtain these standards: organic synthesis (Raab et al., 2010), enzymatic reactions (Smaali et al., 2007; Vic, Biton, Le Beller, Michel, & Thomas, 1994). Some purifications from plants or foods may be performed as well. Raab and collaborators developed the chemical synthesis of (+)-catechin 3′-*O*-β-D-glucopyranoside, (+)-catechin 5-*O*-β-D-glucopyranoside, and (+)-catechin 3-*O*-β-D-glucopyranoside. They also obtained (+)-catechin 4′-*O*-β-D-glucopyranoside and (+)-catechin 7-*O*-β-D-glucopyranoside by using enzyme extracts from lentil and barley (Raab et al., 2010). Another work on a glucosyl-transferase from apple called UGT71A15 has some affinity for flavonoids such as phlorizin (phloretin 2′-*O*-glucoside) (Lepak, Gutmann, Kulmer, & Nidetzky, 2015). In addition, unpublished preliminary assays showed that this enzyme has some activity on flavanols.

The presence of hexosylated monomers of flavanols was already reported in Merlot grape seeds and wine but not in grape skins (Delcambre & Saucier, 2012). By using an ESI/Q-TOF/MS with a targeted MS/MS strategy, 14 flavanol hexosides were detected. MS fragmentation patterns indicated that these molecules are *O*-hexosides of the four major monomers found in grape (Simirgiotis, Silva, Becerra, & Schmeda-Hirschmann, 2012). However, the nature and position of the hexoside substituents have not been determined and whether these compounds can be found as constitutive units of proanthocyanidins is unknown.

The main objective of the present work was to detect by mass spectrometry new flavan-3-ol hexosides compounds in grapes and wines. Also, grape tissues, grape stages of development and variety influences were investigated. Different grapes varieties at three different stages of development and wine were selected.

In this publication, we wanted to test the following hypotheses:

– Several isomers of dimeric flavanol hexosides are present in different grape varieties and wines
– The content of these compounds is influenced by the grape variety, the tissue (skin or seed) and the grape development stage
– Some of the compounds are β glucosides of flavanols

2. Materials and methods

2.1. Reagents and samples

Deionized water was purified with a Milli-Q purification system (Millipore, Molsheim, France). Acetone and diethyl ether were obtained from Analytic Lab (St Mathieu de Treviers, France). Formic acid and HPLC grade methanol were purchased from Sigma Aldrich (St Louis, MO, USA).

Grapes from Unité Expérimentale de Pech Rouge in Gruissan (France) were harvested at three different dates corresponding to different development stages.

Merlot: 31/07/2015 (3.4 °Brix), 14/08/2015 (20.2 °Brix), 09/09/2015 (23.1 °Brix)
Syrah: 31/07/2015 (2.9 °Brix), 14/08/2015 (11.1 °Brix), 09/09/2015 (20.4 °Brix)
Cabernet Sauvignon (CS): 14/08/2015 (3.2 °Brix), 09/09/2015 (15.1 °Brix), 24/09/2015(22.9 °Brix).

The first development stage sampling point of each grape is named *phase I* (P-I), the second one, *phase II* (P-II), and the last one is at maturity, called *phase III* (P-III). The samples were stored at −80 °C before analysis.

Languedoc-Roussillon red wines selected for a screening were: a Tannat press wine from cave de Crouseilles, Madiran-2016, an Alicante Bouschet from Chateau de Pinet-2015, a Merlot from Mas Pellier-2014, a Syrah from Mas de Valbrune-2013 and a Grenache from Château Flaugergues, Montpellier-2016. Red wines were stored at −80 °C before analysis.

2.2. Preparation of grape tannin extracts

For each grape variety, seeds and skins were removed by hand on ice from 100 berries. Ten grams of each tissue were extracted using 200 mL of acetone/water 7:3, under nitrogen with mechanical agitation, overnight. Solutions were filtered on filter paper, evaporated under reduced pressure at 37 °C. The residues were dissolved in deionized water and lyophilized. Tannin extract powders were stored at −20 °C.

2.3. UPLC-DAD-ESI-IT-MS method

All analyses were performed on a Waters Acquity UPLC-DAD system (Milford, MA) coupled with a Bruker Daltonics Amazon (Bremen, Germany) mass spectrometer equipped with an electrospray source and an ion trap mass analyzer. Separation was achieved using a C18 column (Waters, Milford, BEH, 1.7 μm, 1 × 150 mm) at temperature of 35 °C. The binary mobile phase consisted of deionized water with 1% of formic acid (A) and methanol with 1% of acid formic (B). The 45-minutes-long gradient was as follows: 0–1 min, isocratic, 2% B; 1–10 min, 2–30% B; 10–12 min, isocratic, 30% B; 12–25 min, 30–75% B; 25–30 min, 75–90% B; 30–35 min, isocratic, 90% B; 35–45 min, isocratic 2% B. The flow rate was 0.08 mL.min^{-1}.

For UPLC-DAD-ESI-IT-MS analysis, grape extracts were prepared in methanol/H$_2$O 1:1 to a final concentration of 1 mg.mL^{-1}.

Red wines were all centrifuged at 13,500 rpm, during 15 min at 4 °C. The supernatants were recovered and injected.

The MS instrument was operated in the negative mode with fragment ion scan mode focused on specific precursor ions at *m/z* 451 and 739 (end plate off set: −500 V; temperature, 200 °C; nebulizer gas: 10 psi and dry gas, 5 L.min^{-1}; capillary voltage, 4.5 kV). Collision energy for fragmentation used for MS^2 experiments were set at 1.

A UGT714A15 reaction kinetic study was done, as shown on Appendix 1. The proportions of (+)-catechin 4′-*O*-β-D-glucoside and (+)-catechin 7-*O*-β-D-glucoside were monitored with the assumption that they had the same UV absorption coefficient as (+)-catechin at 280 nm.

Monomeric and dimeric flavanol hexosides relative intensities were measured based on the peak height of their associated *m/z* Extract Ion Chromatogram (EIC). Each sample was analyzed in triplicate. Samples were analyzed in several batches within a 2-month period and the average intensities and standard deviations (error bars) calculated.

2.4. Heterologous production of UGT71A15 and enzymatic reactions

Cloning and transformation of the glucosyltransferase UGT71A15 (GenBank: DQ103712) have been described elsewhere (Gosch, Halbwirth, Schneider, Hölscher, & Stich, 2010). The heterologous expression of N-terminal Strep-tagged UGT71A15 in *E. coli* and purification by Strep-affinity chromatography were performed as described earlier (Lepak et al., 2015).

UGT71A15 was assayed in a reaction volume of 200 μL containing 50 mM HEPES pH 8.0, 50 mM KCl, 13 mM MgCl$_2$, 50 μg/mL enzyme with different acceptor reagents, (+)-catechin and (−)-epicatechin, and UDPG concentrations (1 mM, 2.5 mM and 5 mM) at 30 °C. To obtain

a sufficient amount for structural characterization by NMR and MS, the reaction was scaled up to 4 mL using the same buffer conditions. (+)-catechin 4′-O-β-glucoside was then purified by preparative HPLC (940 LC Varian) using a C18 reversed phase (Microsorb C18 100-3, 100 * 21.4 mm (L * ID), Dynamax Agilent). Samples were filtered on a 0.2 μm regenerated cellulose membrane filter and 1 mL was injected. The binary mobile phase consisted of deionized water with 1% of formic acid (A) and acetonitrile, water and acid formic (80:19:1) (B). The 16-minutes-long gradient was: 0–1 min, 10% B; 1–10 min, 16% B; 10–12 min, 80% B; 12–14 min, 10% B; 14–16 min, 10% B. The flow rate was 10 mL.min^{-1}.

2.5. NMR spectroscopy experiments

All experiments were performed on an Agilent 500 MHz DD2 NMR spectrometer (Agilent Technologies, Santa Clara, CA, USA) operating at 500.05 MHz for 1H and 125.75 MHz for ^{13}C and equipped with a 5 mm indirect detection Z-gradient probe. The chemical shifts were reported to that of methanol-d_4, the organic solvent used for the experiments (3.31 ppm and 49.1 ppm for ^{1}H and ^{13}C respectively). All data were collected at 25 °C with active sample temperature regulation.

Identification and structure characterization of products were performed using both 1D and 2D NMR, homonucluear ^{1}H (DOSY, COSY, TOCSY and ROESY) and heteronuclear ^{1}H/^{13}C (HSQC and HMBC) experiments. Data were processed and analyzed using VNMRJ software.

2.5.1. (+)-Catechin-4′-O-β-glucoside (C4′OG)
^{1}H NMR (25 °C, MetOD, ppm) δ 4.629 (d, J = 7,3, C2); 3.982 (m, C3), 2.516 (dd, 7.9; 16.5, C4), 2.817 (dd, 5.5; 16.5, C4), 5.932 (d, 2.2, A6), 5.868 (d, 2.2, A8), 6.906 (d, 2.0, B2′), 7.183 (d, 8.4, B5′), 6.831 (d, 2.0; 8.4, B6′), 4.771 (d, 7.5, Glc1), 3.492 (m, Glc2), 3.474 (m, Glc3), 3.408 (m, Glc4), 3.410 (m, Glc5), 3.716 (m, Glc6,6′), 3.895 (bd, 11.9 Hz, Glc6,6′).
^{13}C NMR (25 °C, MetOD) δ 82.45 (C2), 68.69 (C3), 28.35 (C4), 157.59 (A5), 96.33 (A6), 157.80 (A7), 95.45 (A8), 156.73 (A9), 100.68 (A10), 136.39 (B1′), 115.79 (B2′), 148.25 (B3′), 146.51 (B4′), 118.58 (B5′), 119.88 (B6′), 104.29 (Glc1), 74.90 (Glc2), 77.51 (Glc3), 71.32 (Glc4), 78.26 (Glc5), 62.39 (Glc6,6′).

2.5.2. (+)-Catechin-7′-O-β-glucoside (C7′OG)
^{1}H NMR (25 °C, MetOD, ppm) δ 4.659 (d, J = 7,4, C2) ; 3.996 (m, C3), 2.537 (dd, 8.2 ; 16.1, C4), 2.817 (dd, 5.2 ; 16.5, C4), 6.201 (d, 2.4, A6), 6.157 (d, 2.4, A8), 6.829 (d, 1.8, B2′), 6.758 (d, 8.1, B5′), 6.711 (d, 1.8 ; 8.1, B6′), 4.815 (d, 7.6, Glc1), 3.35–3.45 (m, Glc2, Glc3, Glc4, Glc5), 3.696 (dd, 4.5 ; 12.0, Glc6,6′), 3.884 (dd, 2.0 ; 12.0, Glc6,6′).

3. Results and discussion

3.1. Detection of hexosylated flavanols by targeted UPLC-DAD-ESI-IT-MS

The presence of hexosylated flavanol monomers and dimers was investigated in analyzed wine and Merlot grape extracts. Specific molecular ions corresponding to (epi)catechin and procyanidin dimer hexosides were targeted by ion-trap LC–MS. Respectively, five and seven peaks eluted at different retention times were detected in the extracted ion chromatograms at m/z 451 and m/z 739 (Appendix 5). These isomers may correspond to different substitution positions on the flavanols (catechin or epicatechin), by different hexoses or different conformations. Indeed, the sugar can be bound on any of the flavanol hydroxyl positions or on the C6 or C8 position.

3.1.1. Theoretical MS fragmentation of flavanol monomer hexosides
Specific MS/MS fragmentations were also searched beside the molecular ions to confirm the monomeric flavanol hexosides structures. According to the hexose position, some theoretical fragmentation can be described as illustrated in Fig. 1. The different possible
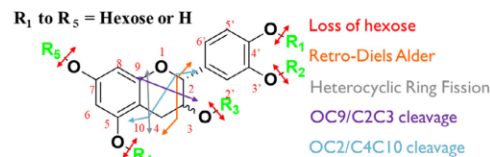

Fig. 1. Fragmentation pathways of (epi)catechin hexosides (). adapted from Delcambre and Saucier (2012)

fragmentations are then retro-diels Alder (RDA), Heterocyclic ring fission (HRF) (Gu et al., 2003) and different bond cleavages. The most common detected fragment is (epi)catechin aglycone detected at m/z 289, corresponding to the loss of hexose attached on a hydroxyl group (−162), represented by red arrows. This fragment ion is present for each possible hexose position. In addition, other fragments from different configurations may have the same mass and do not permit to conclude on the hexose position. For example, fragments at m/z 299 and 151 and m/z 137 and 313 obtained from RDA and OC9/C2C3 cleavage (respectively orange and purple arrows) can be detected if the hexose is bound on A, B or C ring (Delcambre & Saucier, 2012). NMR analysis and molecular standards are then necessary to confirm the nature of the hexose and its position on the molecule.

3.1.2. Theoretical MS fragmentation of flavanol dimer hexosides
Fig. 2 presents the theoretical fragmentations of a dimeric flavanol glycoside with a hexose (for instance a glucose) on the position 5 of the upper unit A ring. The loss of the glucose will yield the proanthocyanidin dimer aglycone ion (m/z 577) for any position of glucose. RDA fragmentation can occur twice on the C ring of the dimer upper unit and on the C′ ring of the lower unit, yielding an ion fragment at m/z 587 and 435 or 587 and 273 with the loss of both B rings (−152 x2). This fragmentation is specific of the hexose on the A ring. HRF may also take place on the upper unit to yield a fragment ion at m/z of 451 and 287. The last classical presented fragmentation is the quinone methide (QM) cleavage also called interflavan cleavage. If the hexose is bound on the upper unit, like on Fig. 2, fragment ions at m/z of 449 Da and at 289 Da are formed. If the hexose is on the lower unit, ions at m/z 451 and 287 are obtained. Depending on the hexose position, other MS^2 fragmentations are summarized in Appendix 2. Other fragments such as m/z 587 and 721 are common for all flavanol dimer hexosides.

3.1.3. Characterization of flavanol dimer and monomer hexosides in grape and wine using MS fragmentation
The use of MS^2 fragmentation has allowed (Appendix 5) the determination of the hexose position on the upper or the lower unit of proanthocyanidin hexosides but not its exact position. In the case of monomeric flavanol hexosides, the fragment at m/z 289 was observed from all isomers (Fig. 3a) and did not indicate the hexose position. To better identify the monomeric flavanol hexosides structures, and for further semi-quantification studies, flavanol glucoside enzymatic hemisynthesis pathways were investigated.

Concerning observed dimer isomers, as presented in Fig. 3b, the principal fragmentation patterns detected in our samples were m/z 721 (−18, loss of water 18), 587, 569 (dehydration from 587), 451, 435, 417 (dehydration from 435), which are common for all dimer hexoside configurations, and also at 449, 339 and 289 which are more specific of some hexose position. Indeed, MS^2 fragments at m/z 449 and 289 would be obtained from a QM, indicating that the hexose position would be attached on the upper unit. MS^3 fragment at 339 represents the loss of the B cycle (−110) from the previous pattern at 449 showing that the hexose was not on it. The other possibility was that the hexose was attached on the C ring (position 3). But fragments from dehydrations were detected and commonly occur on the C3 (Li & Deinzer, 2007), then the hexose is attached on other hydroxyl. To
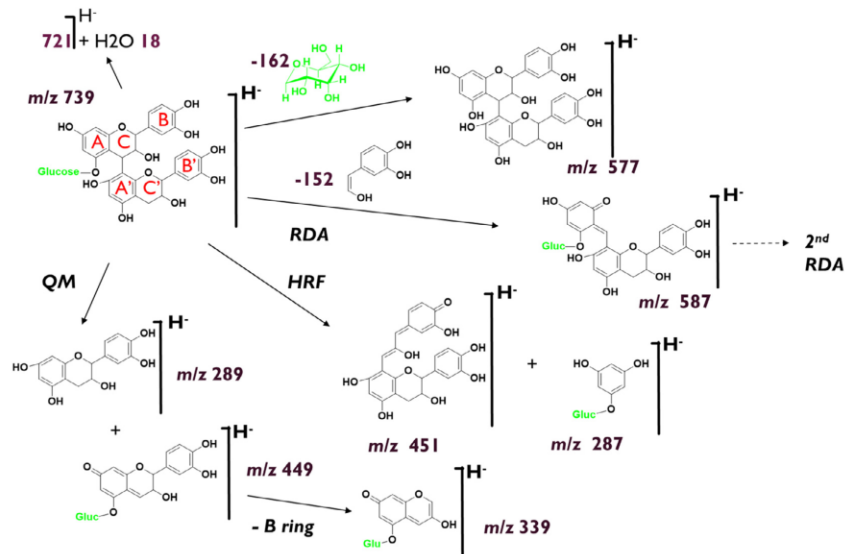
Fig. 2. Fragmentation patterns for a proanthocyanidin glucoside with the glucose attached on the A cycle of the upper unit.

conclude, the fragmentation observed is in favor of a hexose attached on the upper unit and on the A ring (position 5 or 7) as illustrated in Fig. 2.

3.2. Enzymatic production and characterization of flavan-3-ol glycosides: (+)-catechin 4′-O-β-glucoside and catechin-7-O-β-glucoside

A high-efficiency single-step hemisynthesis of C4′OG and C7′OG, catalyzed by the UGT71A15 glucosyl transferase, is depicted below (Fig. 4a). Previous work by Nidetsky and Lepak showed that the UGT71A15 (GenBank: DQ103712) from apples catalyzes site-selective 5-O-β-glucosylation of phlorizin, using UDPG as the glucose donor (Lepak et al., 2015). Similar experimental conditions were tested for glucosylation of (+)-catechin and (−)-epicatechin. Like resveratrol, (+)-catechin was an efficient acceptor substrate showing a high level of

conversion after 24 h. Two signals on the UV chromatogram at 280 nm were observed which corresponded to a molecular ion mass at m/z 451 with a m/z at 289 MS/MS fragment for the two retention times. NMR analysis, in accordance with literature (Raab et al., 2010), allowed us to determine the molecular configuration of (+)-catechin 4′-O-β-glucoside (C4′OG) and catechin-7-O-β-glucoside (C7OG) showed in Appendix 3. Chemical shifts and coupling constants are in accordance with those find in the literature (Raab et al., 2010; Vihakas, Tahtinen, Ossipov, & Salminen, 2012). Conversely, (−)-epicatechin was not converted by the enzyme to its corresponding glucoside. A kinetic study was undertaken to optimize the (+)-catechin concentration: 1 mM, 2.5 mM and 5 mM concentrations were compared (Appendix 1). The reaction with 2.5 mM reached a plateau at 24 h for C7OG but not for C4′OG. The initial concentration of 1 mM was more effective than others based on the yield obtained. In fact, the C4′OG amount represented approximately
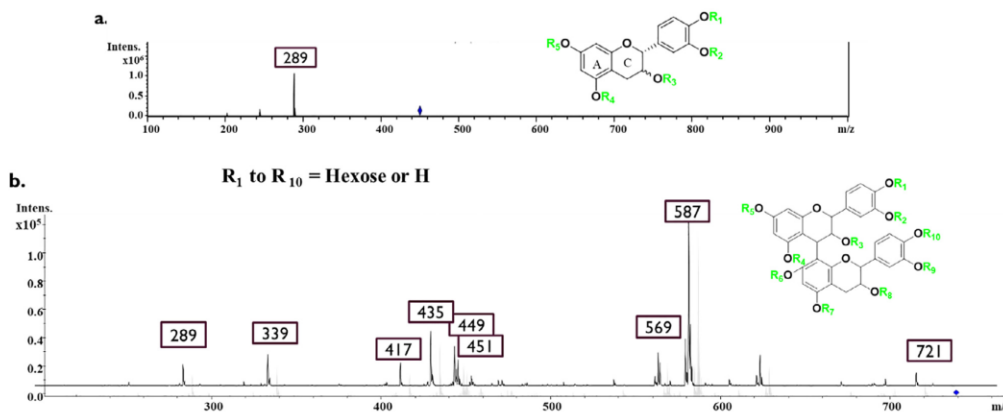
Fig. 3. Examples of flavan-3-ols hexosides MS2/3 fragmentations in merlot grape seed extract from P-III in the negative ion mode with (a). Hexoside (epi)catechin (from m/z 451, at 10.4 min), (b) Hexoside (epi)catechin dimer (from m/z 739, at 12.2 min).

444
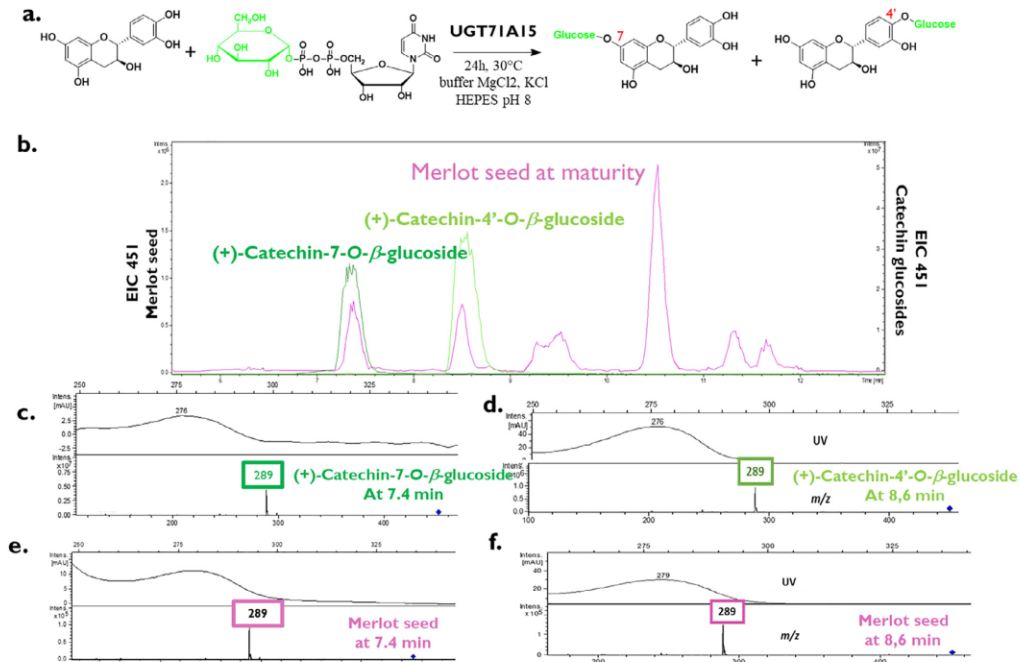
Fig. 4. (a) Enzymatic hemisynthesis of catechin 4′-*O*-β-glucoside (C4′OG) and catechin 7-*O*-β-glucoside by the glucosyltransferase UGT71A15 (b) *m/z* 451 EIC of C4′OG (light green), C7OG (dark green) and merlot seed (pink) at the end of veraison. UV chromatogram and MS[2] from *m/z* 451 EIC of (c) C7OG, (c) C4′OG and molecules (e) at 7.5 min and (f) at 8.6 min from merlot seed at the end of veraison. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
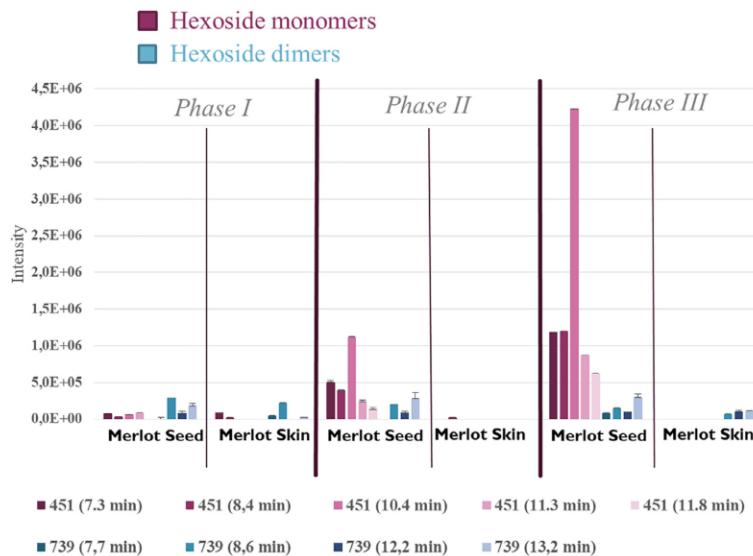
Fig. 5. EIC intensities for hexosylated monomers in purple (*m/z* 451) and dimers in blue (*m/z* 739) from grape polyphenols extracts from seed and skin of Merlot at three different stages (P-I, P-II and P-III). Error bars correspond to MS standard deviation for triplicate analyses.

445
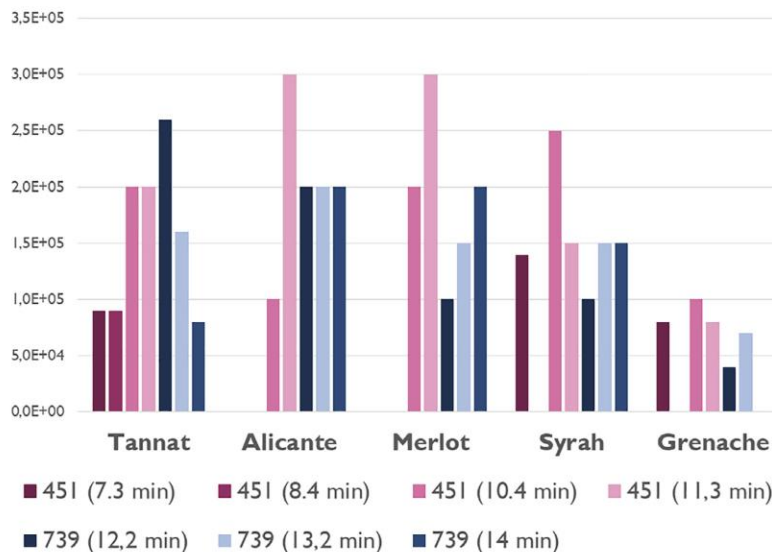
Fig. 6. EIC intensities for hexosides monomers in purple (*m/z* 451) and dimers hexosides in blue (*m/z* 739) from five red wines from different grape varieties.

40% of the initial catechin concentration in the reaction medium for 1 mM compared to 28% and 10% for respectively 2.5 mM and 5 mM catechin. The reaction containing 1 mM and 5 mM catechin reached a plateau after 24 h for C4′OG and C7′OG. The same result was observed for C7OG yield: 25% for 1 mM and 13% and 4.5% for respectively, 2.5 mM and 5 mM. The higher the reagent concentration, the lower the reaction yield. We can hypothesize an excess of (+)-catechin can inhibit the enzyme activity. A substrate concentration at 1 mM was then chosen for further synthesis. On Fig. 4b, the overlaid extracted ion chromatograms (EIC) at *m/z* 451 of the enzymatic solution containing C4′OG and C7OG and of the merlot seed extract in P-II were compared. Based on the retention times and mass spectrometry results, our synthesized compounds would appear to be naturally present in the seed extracts. In fact, MS^2 fragmentation depicted on Fig. 4c, e. and on Fig. 4d, f show similar fragmentation patterns with aglycone formation at *m/z* 289 for both of our compounds, even if a coelution of different isomers cannot be completely ruled out. These monomers are the most polar of flavanol hexosides and they are the major isomers. The less polar isomer could be epicatechin hexosides derivatives but unfortunately, (−)-epicatechin was not a substrate for UGT71A15.

3.3. Study of Merlot grape composition at three different stages of development

Several isomers of monomeric and dimeric flavanol hexosides were detected in the skin and seed extracts of the three grape varieties, Merlot, Cabernet-Sauvignon and Syrah. Comparison was based on the assumption that compounds showing the same retention time and MS data were identical although some isomers might co-elute. In Fig. 5 and Appendix 4, histograms represent the ion intensities of targeted flavanol hexosides in seed and skin extracts at three different stages of grape development. Since the standards for these molecules are not available, these analyses allowed qualitative comparisons between samples rather than absolute quantification.

Merlot grape seed and skin extracts were first investigated. In seed, an increase of all monomeric flavanol hexosides was observed during development; these include the compounds eluted at 7.3 and 8.4 min, tentatively attributed respectively to C4′OG and C7OG based on

retention time and MS/MS fragmentation (loss of glucose), and also those eluted at 10.4, 11.3 and 11.8 min. The EIC intensities of the five monomeric flavanol hexosides increased from 10^4 order of magnitude in P-I, to around 10^5–10^6 in P-II, and reached 4.10^6 at maturity, P-III, for the major isomer. In Merlot skins, a small amount of monomer hexosides were detected in P-I. Both isomers of monomeric flavanol hexosides at 7.3 min and 8.4 min were distinguished at the earlier stage. The isomer attributed to C4′OG (7.3 min) was still present in P-II and finally no more detected at maturity.

Four flavanol dimer hexosides were detected in Merlot seed extracts and their distribution monitored during ripening. The most abundant one, eluted at 13.2 min, accumulated in seeds until P-II and was detected in skins in P-I and in P-III. Another isomer eluted at 8.6 min was present in higher amount in the seeds in P-I. The other ones were detected in trace amounts in seeds and skins in P-I (elution times at 12.2 and 13.2 min).

In Merlot seeds, five isomers of monomer hexosides and four isomers of dimer hexosides were detected, and in Merlot skin, two monomers and three dimers in small amount. For CS seeds, five monomers and three dimers were detected and one monomer and three dimers in CS skin. Finally, five monomers and three dimers were observed in Syrah seed extracts and three dimers and no monomer in skin (Appendix 3).

All the monohexosylated flavanols were detected in the different grape varieties except the flavanol dimer hexoside eluted at 7.7 min which was only detected in our Merlot sample. To compare tissues, in seeds, an accumulation of all flavanol monomer hexosides was measured during maturation and the isomer at 10.4 min was the most abundant in all grape varieties. The isomer of monomeric flavanol hexoside at 7.3 min was the second most present in CS and Syrah seed extracts but not in merlot seed. In skin, the targeted hexosides were detected in trace amounts with a major quantity of isomers detected during the P-I in all grape varieties.

The presence of flavanol hexoside monomers in Merlot seeds in P-III was in accordance with previous studies (Delcambre & Saucier, 2012) and was also evidenced in two other grape varieties in this work. In addition, dimeric flavanol hexosides were detected in grape seed and skin for the first time in our experiments. Additionally, the presence of

(*epi*)gallocatechin (*m/z* 467) and epicatechin gallate (*m/z* 603) hexosides were also investigated, but, their concentrations were below their detection levels in our samples.

3.4. Study of five red wine from different grape varieties

Analysis of five red wines from Tannat, Alicante, Merlot, Syrah and Grenache grape varieties showed that they contain similar levels of flavanol hexosides as shown on Fig. 6.

Concerning flavanol monomer hexosides, the wine from the Tannat variety had the highest number of isomers eluted at 7.3, 8.4, 10.4, and 11.3 min. Three flavanol monomer hexosides were detected in Syrah and Grenache wines at 7.3, 8.4 and 10.4 min. Also, Alicante and Merlot wines included only two isomers, at 10.4 and 11.3 min. All these flavanol hexosides were detected in grape seed at different levels. The flavanol monomeric hexoside at 10.4 min was one of the most abundant in all analyzed seed grapes and wines. On the contrary, the monomer isomer at 11.3 min was one of the most abundant compounds in wine but was not the major one in our grape seed extracts. Finally, the monomer hexoside detected at 11.8 min in seed extracts from all grape varieties seems to be absent in our wine samples.

The same three flavanol dimer hexosides at 12.2, 13.2 and 14 min were detected in the wines, except for the Grenache wine in which the dimer detected at 14 min was not present. The two dimeric flavanol hexosides detected at 12.2 and 13.2 min in wine were also present in Merlot, CS and Syrah grape seeds. But, the isomer at 14 min was absent in all grape extracts. Finally, we detected mostly a dimer at 13.2 in Syrah, Merlot and CS seeds in P-III and in wines.

To conclude, our results based on similarity of retention times and MS data suggest that the majority of flavanol hexosides detected in wines have a grape origin. Other flavanol hexosides (such as the dimer detected at 14 min) are present in wine but not in grape. Other biochemical pathways (production or degradations) from microorganisms (yeasts and bacteria) are also possible. Concerning the grapes, the genetic ability to produce these compounds seems to be shared amongst the different varieties but they were detected in different proportions which might be a result of different genetic expressions of some glycosyl transferase genes.

4. Conclusion

The presence of monomeric flavanol hexosides was confirmed in grape and red wines by mass spectrometry. Proanthocyanidin dimer hexosides were successfully detected in our samples with specific MS/MS fragmentations. The flavanol monomer and dimer hexosides were present in wines from different grape varieties (Tannat, Alicante, Merlot, Syrah, and Grenache) and also in Merlot, Syrah and CS grape berries. One of the flavanol monomer hexoside was found in grapes but not in wines. All (*epi*)catechin hexosides concentrations increased from P-I until P-III in seeds and decreases or were absent in skin. Standards of (+)-catechin 4′-*O*-β-glucoside and catechin-7-*O*-β-glucoside were obtained through an enzymatic hemisynthesis with the use of glucosyl transferase UGT71A15. The presence of these compounds in analyzed grape seeds, skins and wines was evidenced based on HPLC–MS analysis. The mechanisms of formation of those hexosylated flavanols and their possible roles in proanthocyanidin biosynthesis in grapes remain to be elucidated.

Funding

References

Ali, A., & Strommer, J. (2003). A simple extraction and chromatographic system for the simultaneous analysis of anthocyanins and stilbenes of Vitis species. *Journal of Agricultural and Food Chemistry, 51*(25), 7246–7251.

Bogs, J., Downey, M. O., Harvey, J. S., Ashton, A. R., Tanner, G. J., & Robinson, S. P. (2005). Proanthocyanidin synthesis and expression of genes encoding leucoanthocyanidin reductase and anthocyanidin reductase in developing grape berries and grapevine leaves. *American Society of Plant Biologists, 139*(2), 652–663.

Castillo-Muñoz, N., Gómez-Alonso, S., García-Romero, E., Gómez, M. V., Velders, A. H., & Hermosín-Gutiérrez, I. (2009). Flavonol 3-O-glycosides series of Vitis vinifera Cv. petit verdot red wine grapes. *Journal of Agricultural and Food Chemistry, 57*(1), 209–219.

Corder, R., Mullen, W., Khan, N. Q., Marks, S. C., Wood, E. G., Carrier, M. J., & Crozier, A. (2006). Oenology: Red wine procyanidins and vascular health. *Nature, 444*, 566.

Cui, E.-J., Song, N.-A., Shrestha, S., Chung, I.-S., Kim, J.-Y., Jeong, T.-S., & Baek, N.-I. (2016). Flavonoid glycosides from cowpea seeds (Vigna sinensis K.) inhibit LDL. *Food Science and Biotechnology, 21*(2), 619–624.

Delcambre, A., & Saucier, C. (2012). Identification of new flavan-3-ol monoglycosides by UHPLC-ESI-Q-TOF in grapes and wine. *Journal of Mass Spectrometry, 47*(6), 727–736.

Downey, M. O., Harvey, J. S., & Robinson, S. P. (2003). Analysis of tannins in seeds and skins of Shiraz grapes throughout berry development. *Australian Journal of Grape and Wine Research, 9*(1), 15–27.

Dueñas, M., Sun, B., Hernández, T., Estrella, I., & Spranger, M. I. (2003). Proanthocyanidin composition in the seed coat of lentils (Lens culinaris L.). *Journal of Agricultural and Food Chemistry, 51*(27), 7999–8004.

Foo, L. Y., & Karchesy, J. J. (1989). Polyphenolic glycosides from Douglas fir inner bark. *Phytochemistry, 28*(4), 1237–1240.

Friedrich, W., & Galensa, R. (2002). Identification of a new flavanol glucoside from barley (Hordeum vulgar L.) and malt. *European Food Research and Technology, 214*(5), 388–393.

Fujita, A., Soma, N., Goto-Yamamoto, N., Shindo, H., Kakuta, T., Koizumi, T., & Hashizume, K. (2005). Anthocyanidin reductase gene expression and accumulation of flavan-3-ols in grape berry. *American Journal of Enology and Viticulture, 56*(4), 336–342.

Gosch, C., Halbwirth, H., Schneider, B., Hölscher, D., & Stich, K. (2010). Cloning and heterologous expression of glycosyltransferases from Malus × domestica and Pyrus communis, which convert phloretin to phloretin 2′-O-glucoside (phloridzin). *Plant Science, 178*(3), 299–306.

Gu, L., Kelm, M. A., Hammerstone, J. F., Beecher, G., Holden, J., Haytowitz, D., & Prior, R. L. (2003). Screening of foods containing proanthocyanidins and their structural characterization using LC–MS/MS and thiolytic degradation. *Journal of Agricultural and Food Chemistry, 51*(25), 7513–7521.

Hmamouchi, M., Es-Safi, N., Lahrichi, M., Fruchier, A., & Essassi, E. M. (1996). Flavones and flavonols in leaves of some Moroccan Vitis vinifera cultivars. *American Journal of Enology and Viticulture, 47*(2), 186–192.

Kim, H.-J., Yeom, S.-H., Kim, M.-K., Shim, J.-G., Lim, H.-W., & Lee, M.-W. (2004). Nitric oxide and prostaglandin E2 synthesis inhibitory activities of flavonoids from the barks of Ulmus macrocarpa. *Natural Product Sciences, 10*(6), 344–346.

Lepak, A., Gutmann, A., Kulmer, S. T., & Nidetzky, B. (2015). Creating a water-soluble resveratrol-based antioxidant by site-selective enzymatic glucosylation. *A European Journal of Chemical Biology, 16*(13), 1870–1874.

Li, H.-J., & Deinzer, M. L. (2007). Tandem mass spectrometry for sequencing proanthocyanidins. *Analytical Chemistry, 79*(4), 1739–1748.

Monagas, M., Bartholomé, B., & Gomez-Cordovés, C. (2007). Updated knowledge about the presence of phenolic compounds in wine. *Critical Reviews in Food Science and Nutrition, 45*(2), 85–118.

Nonaka, G.-I., Ezaki, E., Hayashi, K., & Nishioka, I. (1983). Flavanol glucosides from rhubarb and Rhaphiolepis umbellata. *Phytochemistry, 22*(7), 1659–1661.

Pang, Y., Cheng, X., Huhman, D. V., Ma, J., Peel, G. J., Yonekura-Sakakibara, K., ... Dixon, R. A. (2013). Medicago glucosyltransferase UGT72L1: Potential roles in proanthocyanidin biosynthesis. *Planta, 238*, 139–154.

Pang, Y., Peel, G. J., Sharma, S. B., Tang, Y., & Dixon, R. A. (2008). A transcript profiling approach reveals an epicatechin-specific glucosyltransferase expressed in the seed coat of Medicago truncatula. *PNAS, 105*(37), 14210–14215.

Quideau, S., Deffieux, D., Douat-Casassus, C., & Pouysegu, L. (2011). Plant polyphenols: Chemical properties, biological activities, and synthesis. *A Journal of the Gesellschaft Deutscher Chemiker, 50*(3), 586–621.

Raab, T., Barron, D., Arce Vera, F., Crespy, V., Oliveira, M., & Williamson, G. (2010). Catechin glucosides: Occurrence, synthesis, and stability. *Food Chemistry, 58*(4), 2138–2149.

Simirgiotis, M. J., Silva, M., Becerra, J., & Schmeda-Hirschmann, G. (2012). Direct characterisation of phenolic antioxidants in infusions from four Mapuche medicinal plants by liquid chromatography with diode array detection (HPLC-DAD) and electrospray ionisation tandem mass spectrometry (HPLC-ESI–MS). *Food Chemistry, 131*(1), 318–327.

Smaali, I., Maugard, T., Limam, F., Legoy, M.-D., & Marzouki, N. (2007). Efficient synthesis of gluco-oligosaccharides and alkyl-glucosides by. *World Journal of Microbiology and Biotechnology, 23*(1), 145–149.

Tanaka, T., Kataoka, M., Tsuboi, N., & Kouno, I. (2000). New monoterpene glycoside esters and phenolic constituents of Paeoniae radix, and increase of water solubility of proanthocyanidins in the presence of paeoniflorin. *Chemical and Pharmaceutical Bulletin, 48*(2), 201–207.

Tanner, G. J., & Kristiansen, K. N. (1993). Synthesis of 3,4-cis-[3H]leucocyanidin and enzymatic reduction to catechin. *Analytical Biochemistry, 209*(2), 274–277.

Terrier, N., Ollé, D., Verriès, C., & Cheynier, V. (2009). biochemical and molecular aspects of flavan-3-ol synthesis during berry development. *Grapevine Molecular Physiology & Biotechnology,* 365–388.

Versari, A., Parpinello, G. P., Tornielli, G. B., Ferrarini, R., & Giulivo, C. (2001). Stilbene

compounds and stilbene synthase expression during ripening, wilting, and UV treatment in grape cv. corvina. *Journal of Agricultural and Food Chemistry, 49*(11), 5531–5536.

Vic, G., Biton, J., Le Beller, D., Michel, J. M., & Thomas, D. (1994). Enzymatic glucosylation of hydrophobic alcohols in organic medium by the reverse hydrolysis reaction using almond-β-ᴅ-glucosidase. *Biotechnology and Bioengineering, 46*(2), 109–116.

Vihakas, M., Tahtinen, P., Ossipov, V., & Salminen, J.-P. (2012). Flavonoid metabolites in the hemolymph of European pine sawfly (Neodiprion sertifer) Larvae|SpringerLink. *Journal of Chemical Ecology, 38*(5), 538–546.

Watanabe, M. (1998). Catechins as antioxidants from buckwheat (Fagopyrum esculentum Moench) Groats. *Journal of Agricultural and Food Chemistry, 46*(3), 839–845.

Xie, D.-Y., Sharma, S. B., Paiva, N. L., Ferreira, D., & Dixon, R. A. (2003). Role of anthocyanidin reductase, encoded by BANYULS in plant flavonoid biosynthesis. *Science, 299*(5605), 396–399.

Zhao, J., & Dixon, R. A. (2009). MATE transporters facilitate vacuolar uptake of epicatechin 3′-O-glucoside for proanthocyanidin biosynthesis in Medicago truncatula and Arabidopsis. *Plant Cell, 21*(8), 2323–2340.

- **(+)-catechin 4'-*O*-β-glucoside (C4'OG)**
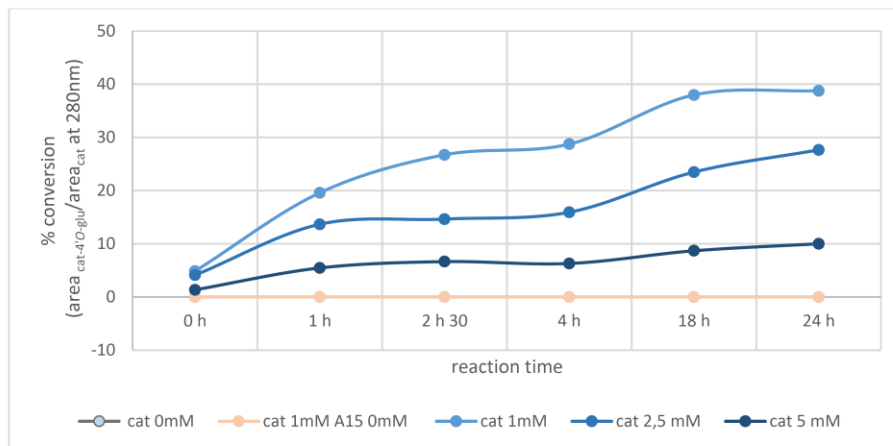
Cat : (+)-catechin

A15 : UGT71A15

- **(+)-catechin 7-*O*-β-glucoside (C7OG)**
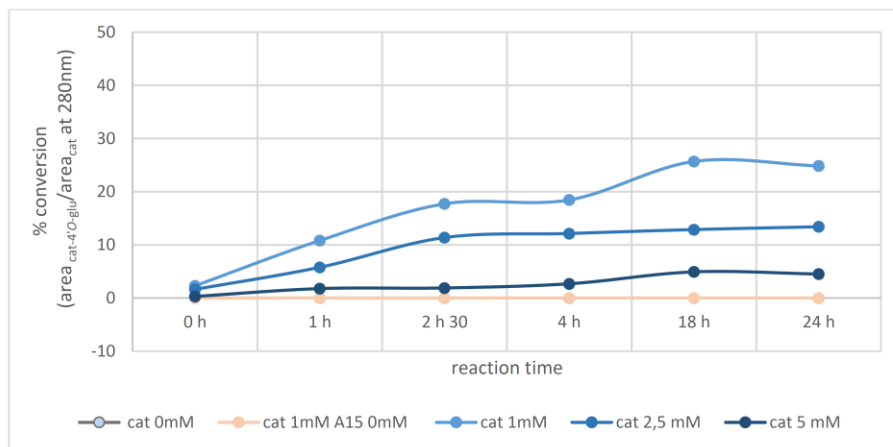
Controls   *cat 0 mM : reaction without (+)- catechin concentration*

        cat 1 mM A15 O mM : reaction without A15 enzym

Appendix 1. Catechin conversion to glucoside: (+)-catechin 4'-O-β-glucoside (C4'OG) and (+)-catechin 7-O-β-glucoside (C7OG).

Upper unit {

Lower unit }

R$_{1\ to\ 10}$ = Hexose

FRAGMENTS MS$^{2/3}$	LOWER UNIT A-RING	LOWER UNIT A-RING	UPPER UNIT B-RING	UPPER UNIT A-RING
273	X		X	
287	X	X		X
289			X	X
339				X
425	X	X	X	X
435		X		X
449			X	X
451	X	X	X	X
575				X
577	X	X	X	X
587	X	X	X	X
613	X	X	X	
721	X	X	X	X

Example 1: Lower Unit, B-ring
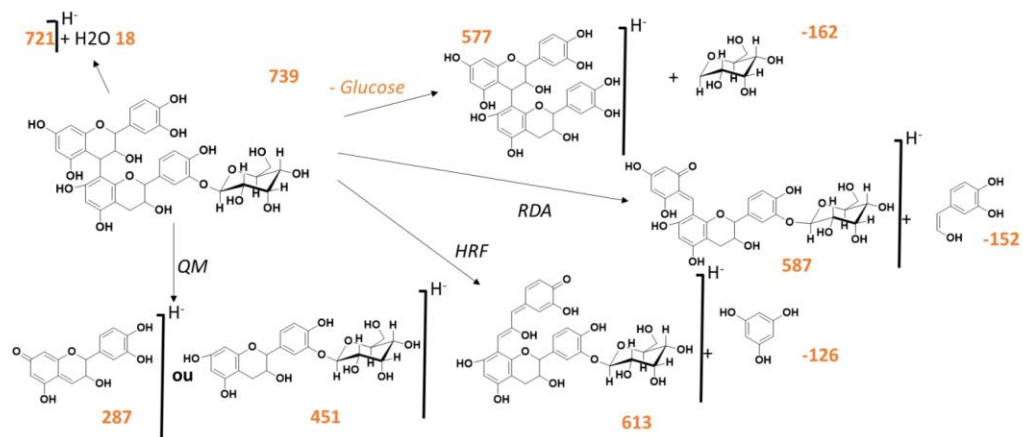
Example 2: Lower Unit, A-ring
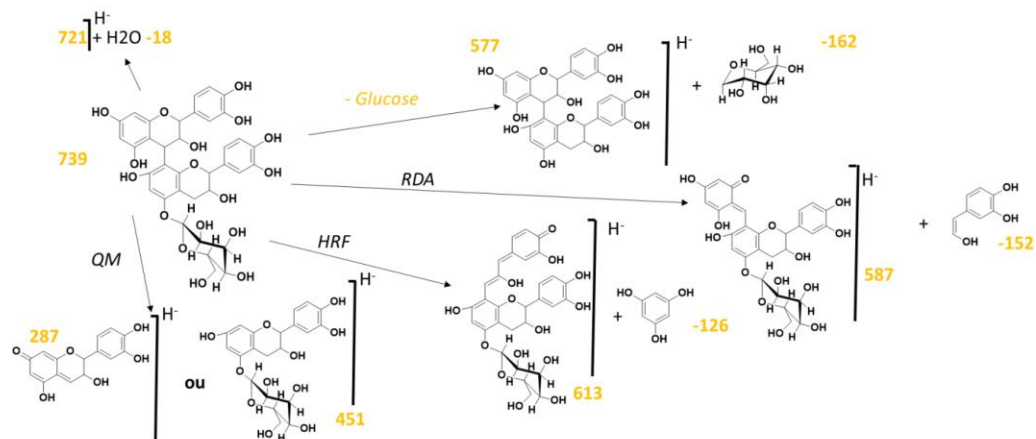
Example 3: Upper Unit, B-ring
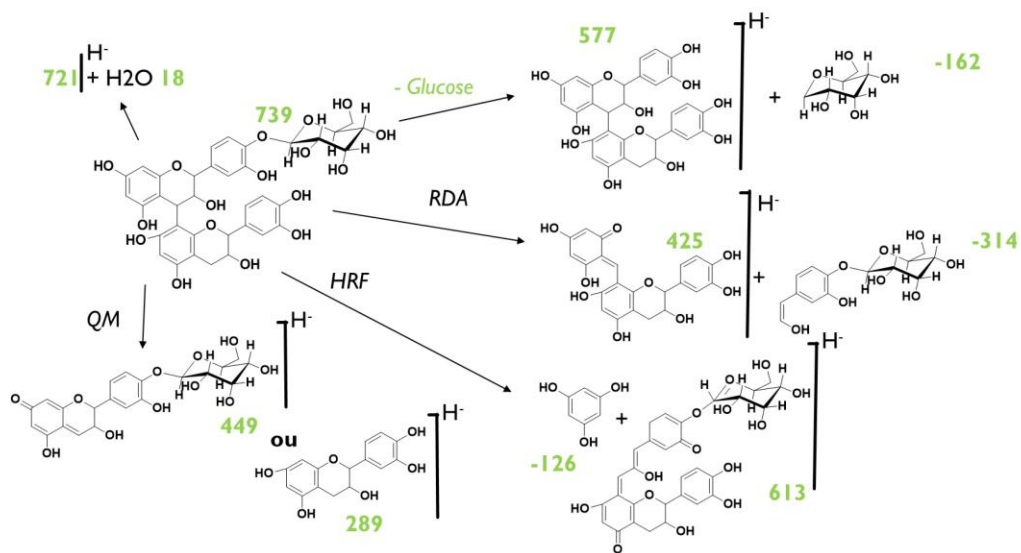
Appendix 2. Proanthocyanidin hexosides fragments depending on glucose position.

- **(+)-catechin-4'-*O*-β-glucoside**
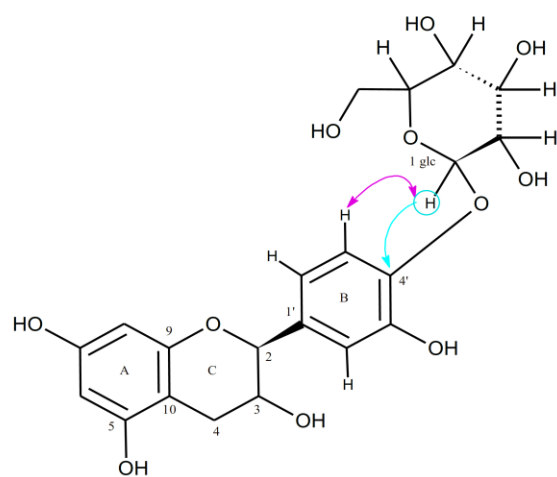
<u>^{1}H NMR :</u>
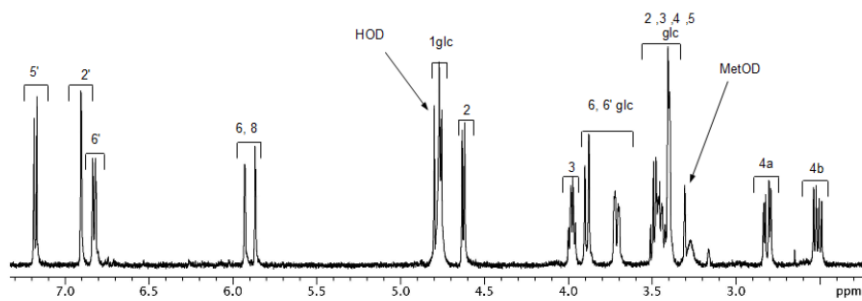
ROESY:

Showing a correlation between the anomeric glucose proton (H1glc) and the proton H5'B of the catechin adduct
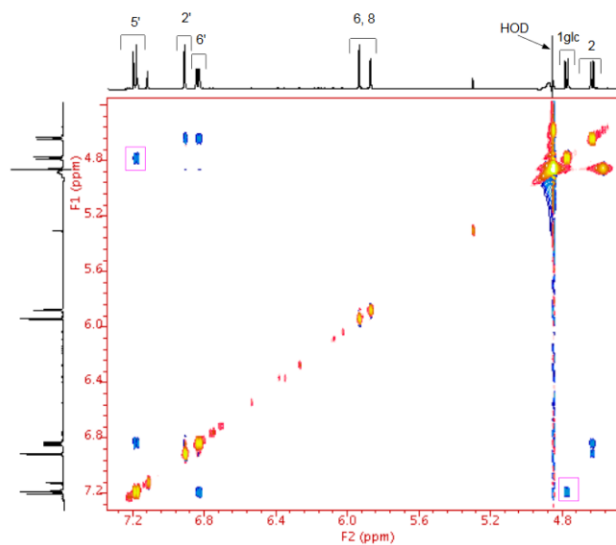
2D HMBC ^{1}H/^{13}C:

Showing a correlation between the anomeric glucose proton (H1glc) and the carbon C4'B of the catechin adduct
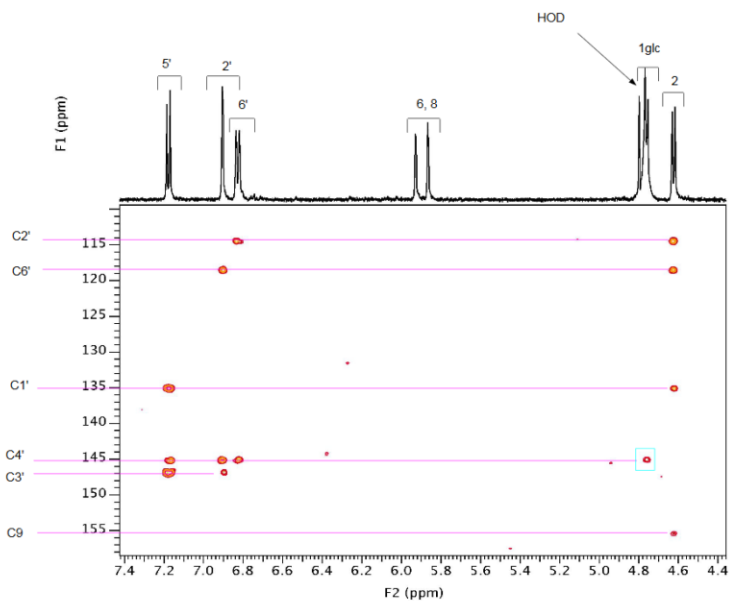
- **(+)-catechin-7-*O*-*β*-glucoside**
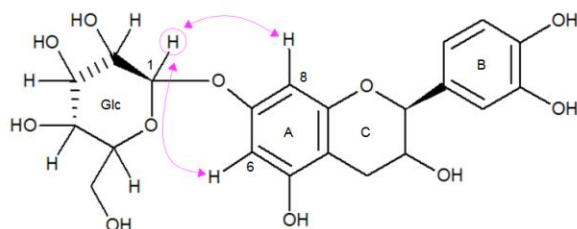
ROESY

Showing the correlations between the anomeric glucose proton (H1glc) and the proton H6A and H8A of catechin adduct
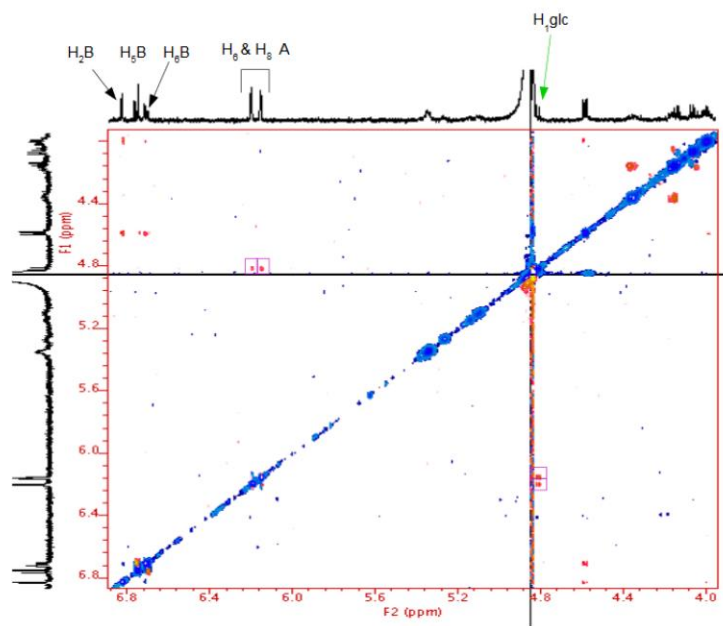
<u>TOCSY (Top: 1D ^{1}H NMR spectrum):</u>

Showing correlations implying the aliphatic protons of catechin adduct (in black) and the protons of glucose moieties (in green)
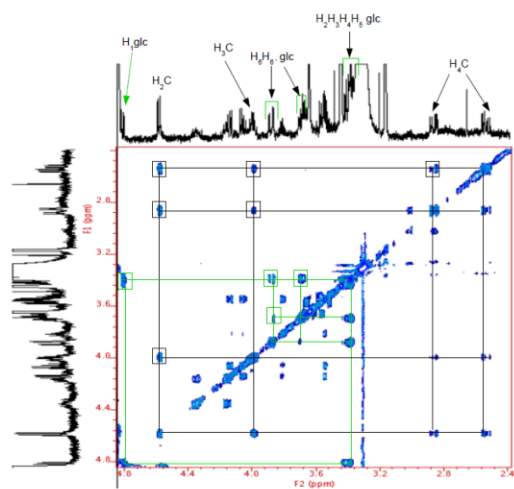
Showing the correlations implying the aromatic protons of catechin adduct
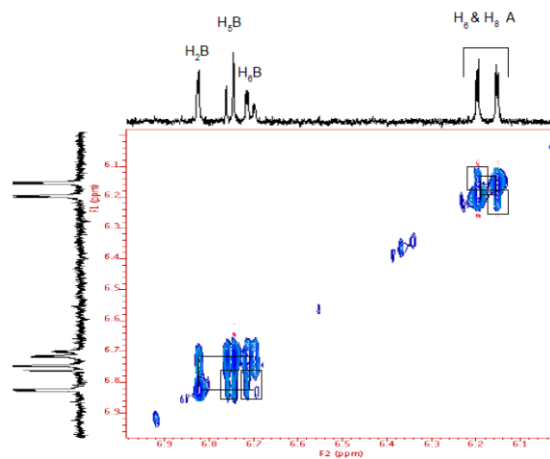
Appendix 3.
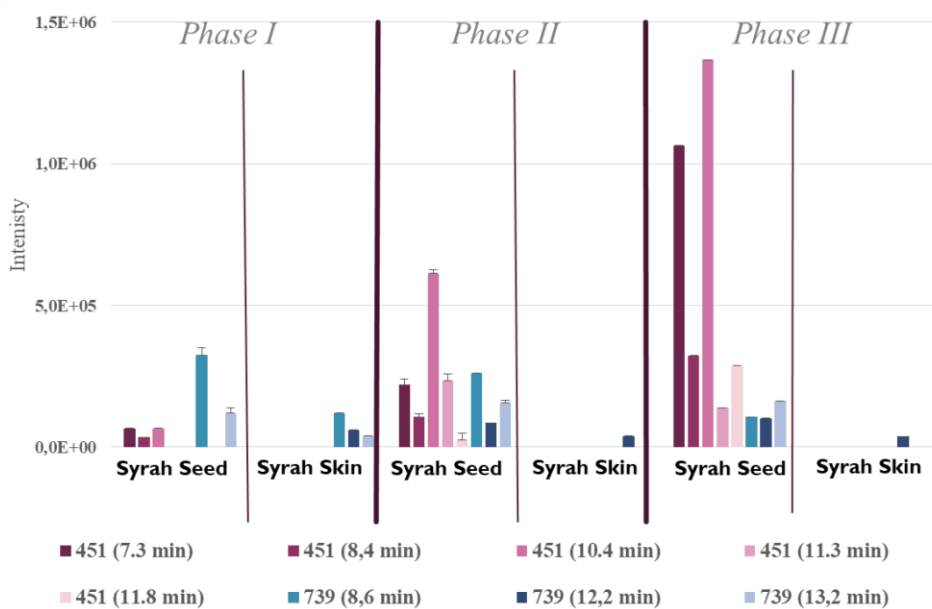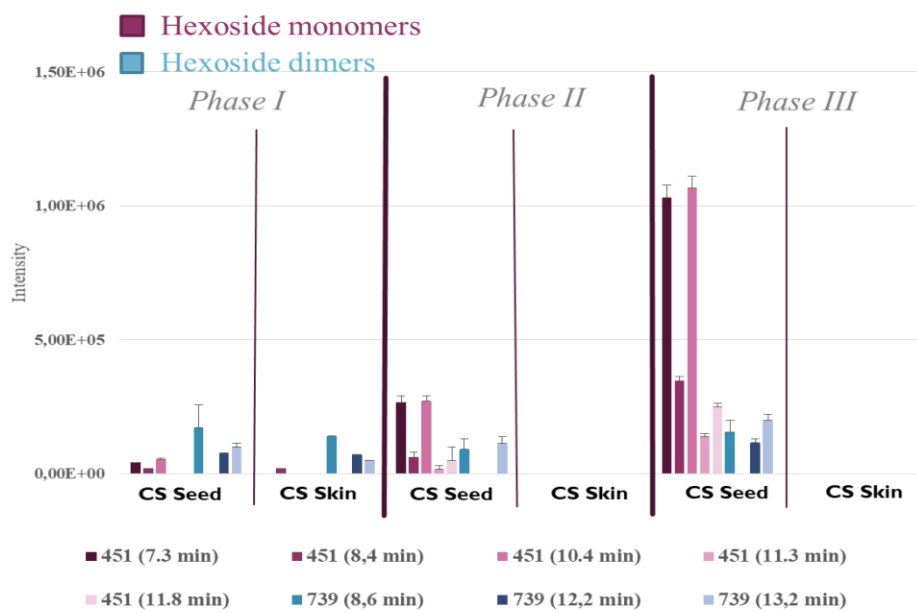
Appendix.4. Flavanol hexosides present in Cabernet Sauvignon (CS) and Syrah in phase I (P-I), phase II (P-II) and phase III (P-III). Average MS intensities and standard deviations correspond to triplicate analyses.

Retention time of flavanol hexoside monomers (min)	MS2/3 fragments
7.3	289
8,4	289
10,4	289
11,3	289
11,8	289

Retention time of flavanol hexoside dimers (min)	MS2/3 fragments	Theoretical hexose position
7,7	289, 577, 587, 721	Upper unit – B ring Upper unit - A ring
8,6	287, 449, 451, 721	Upper unit – A ring
11,8	289, 451, 577, 721	Upper unit – B ring Upper unit - A ring
12,2	287, 289, 341, 339, 435, 449, 451, 569, 587, 721	Upper unit – A ring
13,2	289, 339, 417, 435, 449, 451, 577, 587, 721	Upper unit – A ring
14	287, 289, 271, 339, 449, 451, 587, 721	Upper unit – A ring

Appendix 5. MS$^{2/3}$ of flavanol hexoside monomers and dimers

Chapter 6:

β-Glucosyl Fluoride as Reverse Reaction Donor Substrate and Mechanistic Probe of Inverting Sugar Nucleotide-Dependent Glycosyl-transferases

β-Glucosyl Fluoride as Reverse Reaction Donor Substrate and Mechanistic Probe of Inverting Sugar Nucleotide-Dependent Glycosyltransferases

Alexander Lepak,[¶] Alexander Gutmann,[¶] and Bernd Nidetzky[*,¶,§]

[¶]Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, NAWI Graz, Petersgasse 12, 8010 Graz, Austria

[§]Austrian Centre of Industrial Biotechnology (acib), Petersgasse 14, 8010 Graz, Austria

S *Supporting Information*

ABSTRACT: For a set of flavonoid *O*- and *C*-β-glycosyltransferases, we show that β-glucosyl fluoride can function as substrate for an enzymatic reaction wherein uridine 5′-diphosphate (UDP) α-glucose is synthesized in the presence of UDP. In pH and mutagenesis studies of the *C*-glycosyltransferase from rice, we show that reaction with the β-glucosyl fluoride can serve to identify the acid−base catalytic residue of the enzyme (His24). We also show

Inverting Leloir glycosyltransferase

that β-glucosyl fluoride can rescue activity in an enzyme variant (Ile121Asp) strongly impaired in the canonical reaction wherein flavonoid acceptor is glucosylated from UDP-glucose. Coupling of this variant with the wildtype *C*-glycosyltransferase in a one-pot reaction enabled efficient 3′-β-*C*-glucosylation of phloretin from β-glucosyl fluoride in the presence of substochiometric amounts of UDP.

KEYWORDS: *glycosyltransferase, inverting, retaining, glycosyl fluoride, glycosylation, general acid−base catalysis*
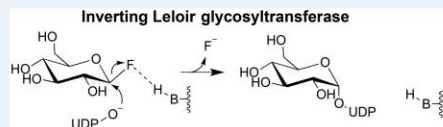
Glycosyl fluorides have gained significant importance as substrates and mechanistic probes of glycosylation reactions performed in water.[1−3] Glycosyl fluorides offer a convenient combination of stability and reactivity under these conditions.[2] This, together with the small steric demand of the fluoride leaving group, makes glycosyl fluorides broadly useful for the study of enzymatic reactions.[2,4−6] Glycosyl fluorides are activated glycosyl donors in a kinetic and thermodynamic sense. With a pK_a of ∼3.2, the fluoride is a very good leaving group. Strong hydration of the fluoride anion in water[7] promotes its expulsion from the glycosyl fluoride. Glycosyl fluorides can therefore drive glycosylations to good synthetic yields.[2,3,5,8,9] Two observations from chemical studies are important mechanistically. The hydrolysis of glycosyl fluorides does not require acid catalysis,[10] and glycosyl fluorides show reactivity with charged nucleophiles in aqueous solution.[11,12] This enabled the assignment of function to active-site residues in glycoside hydrolases through kinetic evaluation of site-directed variants using glycosyl fluorides as substrates.[13,14] Engineered glycoside hydrolases known as "glycosynthases" exploit the advantages of glycosyl fluorides to synthesize product glycosides that these enzymes are unable to hydrolyze.[4,15−17]

Glycosyl fluorides are not well explored as substrates of sugar nucleotide-dependent glycosyltransferases. To our knowledge, the retaining α-galactosyltransferase LgtC from *Neisseria meningitidis*[18] is the only member from this enzyme class for which catalytic reaction with glycosyl fluoride has been reported. LgtC was shown to convert α-galactosyl

fluoride and uridine 5′-diphosphate (UDP) into UDP-α-galactose. In the presence of galactosyl acceptor (e.g., lactose), reaction to the expected product (e.g., α-Gal-1,4-Gal-β-1,4-Glc) took place, but only when catalytic amounts of UDP were also present. Although efficiency of LgtC in terms of k_{cat}/K_m was not high in these reactions (1−2 × $10^{−4}$ mM$^{−1}$ s$^{−1}$), applicability of glycosyl fluorides to the study of glycosyltransferases was strongly suggested. Here, we show that β-glucosyl fluoride can function as donor substrate for inverting β-glucosyltransferases. We also show that the β-glucosyl fluoride thus becomes accessible to mechanistic enzyme characterization along lines established for glycoside hydrolases;[13,14] and that it presents an alternative donor substrate for glycosyltransferases practically used for glycosylation cascade reactions.

Our inquiry started with sucrose synthase which like LgtC is a retaining enzyme.[19] However, when offered α- or β-glucosyl fluoride (2 mM) in the presence of UDP (1 mM) at 30 °C and pH 7.0, the sucrose synthase from soybean (*Glycine max*)[20] proved to be inactive toward forming UDP-glucose. Ion-pairing high-performance liquid chromatography (HPLC) was used to monitor product release (Figure S1, Supporting Information). Addition of fructose (10 mM) to the reactions did not elicit activity. Note: our use of both α- and β-glucosyl fluoride was inspired by two important former observations:

various glycoside hydrolases utilize glycosyl fluorides of the "wrong" anomeric configuration as substrates;[2,21,22] and curiously, LgtC showed slow reaction with an alternative donor (2,4-dinitrophenyl-β-galactoside) having an activated aromatic leaving group of anomeric configuration (β) opposite to that of the natural donor (α).[23] While α-glucosyl fluoride is relatively stable in aqueous solution, the β-glucosyl fluoride hydrolyzes quickly (Figures S2 and S3, Supporting Information). The half-life of β-glucosyl fluoride was determined with proton NMR as ~6 h under the conditions used (30 °C, pD 7.0). Therefore, the stability of especially the β-glucosyl fluoride has to be considered. Note that we typically used the glucosyl fluoride in excess when conversion studies were performed.

We then examined the inverting β-C-glucosyltransferase (GT) from rice (*Oryza sativa*; *Os*CGT) which catalyzes 3′-β-C-glucosylation of the dihydrochalcone phloretin from UDP-glucose.[24] In the presence of UDP (1 mM), *Os*CGT was inactive when offered α-glucosyl fluoride (2 mM) as substrate, but slow enzymatic formation of UDP-glucose occurred when β-glucosyl fluoride (2 mM) was used. When phloretin (1 mM) was added, the α-glucosyl fluoride was still unreactive but the β-glucosyl fluoride was converted gradually to the expected β-C-glucoside (nothofagin). Using ion-pairing HPLC, the acceptor (phloretin) and both intermediary (UDP-glucose) and final reaction product (nothofagin) were quantified (Figure S1, Supporting Information). The initial nothofagin formation rate (0.52 nmol min^{-1} mg^{-1}) under these conditions was about 5 times slower than the UDP-glucose formation rate (2.6 nmol min^{-1} mg^{-1}) measured in the absence of phloretin. Only traces of UDP-glucose were formed in the reaction of β-glucosyl fluoride with UDP when phloretin was present. UMP could not substitute for UDP in reactions of *Os*CGT with β-glucosyl fluoride in the absence and presence of phloretin. Therefore, an overall enzymatic conversion (β-glucosyl fluoride → nothofagin) in two inverting reactions was suggested (Scheme 1). Transformation of β-glucosyl fluoride

Scheme 1. Two-Step Reaction of *Os*CGT with β-Glucosyl Fluoride, Phloretin, and Catalytic Amounts of UDP To Give Nothofagin[a]
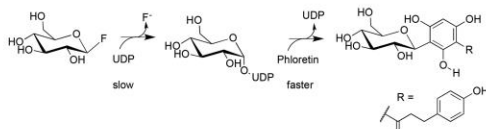
[a]Specific activities measured for the individual reaction steps show that the first step was slow (7.1 nmol min^{-1} mg^{-1}), and the second was comparably fast (2200 nmol min^{-1} mg^{-1}).

to UDP-α-glucose was rate limiting. Phloretin inhibited the enzymatic reaction of β-glucosyl fluoride with UDP. After 24 h, the conversion (based on the phloretin consumed) was 60%.

We expanded the study to five additional *O*- or *C*-β-glucosyltransferases, which like *Os*CGT utilize UDP-glucose for β-glucosylation of phloretin (Table 1). In terms of their sequence, all enzymes were related to *Os*CGT by common membership to family GT-1 of the glycosyltransferase families.[25] Depending on the enzyme used, the relevant glucosylation products were the 3′-*C*-β-glucoside (nothofagin) and the 2′-*O*-β-glucoside (phlorizin) (Table S2, Supporting

Table 1. β-Glucosylation of Phloretin from Different Glucosyl Donors Using a Set of Glycosyltransferases

	specific activity [nmol min^{-1} mg^{-1}]	
enzyme[a]	β-glucosyl fluoride[b]	UDP-glucose[c]
*Os*CGT	4.0	2200
UGT708A6	220	900
UGT71A15	23	2400
UGT708C2	<0.05	2200
*Pc*OGT	<0.05	980
*Rs*AS	<0.05	26

[a]Purified enzyme (1000 μg mL^{-1}) was used. General reaction conditions: pH 8.0, 30 °C. [b]β-Glucosyl fluoride (4 mM), UDP (2 mM), phloretin (1 mM). [c]UDP-glucose (2 mM), phloretin (1 mM). Data are from quadruplicate determinations and have SD of ≤10%.

Information). UGT71A15 additionally formed the 4′-*O*-β-glucoside (trilobatin): Arbutin synthase (*Rs*AS) formed trilobatin exclusively (Table S2, Supporting Information). None of the five glycosyltransferases showed activity with α-glucosyl fluoride. However, when assayed with β-glucosyl fluoride, the bifunctional *O*-/*C*-glycosyltransferase UGT708A6 from maize[26] was active, whereas the *C*-glycosyltransferase UGT708C2 from buckwheat[27] was not. Among three *O*-glycosyltransferases used, only the UGT71A15 from apple[28] showed activity with β-glucosyl fluoride, while enzymes from pear (*Pc*OGT)[20] and Indian snakeroot (*Rs*AS)[29] did not. The use of β-glucosyl fluoride as donor for enzymatic glucosylation of phloretin required the presence of UDP in catalytic amounts. A note of caution is made: traces of UDP copurified with the enzyme (e.g., UGT71A15) in *Strep*-tag affinity purification were already sufficient to elicit activity. Once enzyme-bound UDP was removed completely by ion exchange chromatography, no activity was measurable. Besides demonstrating in initial rate analysis the activity of *Os*CGT, UGT708A6, and UGT71A15 with β-glucosyl fluoride (Table 1), we also confirmed that substrate conversion was 70−80% under the conditions used (24 h, based on phloretin consumed; data not shown). The results summarized in Table 1 imply that β-glucosyltransferases related by both sequence and substrate specificity can differ as regards ability to utilize β-glucosyl fluoride as substrate. This contrasts glycosyltransferases with glycoside hydrolases. For inverting glycoside hydrolases it is common that the glycosyl fluoride corresponding to the canonical substrate is also hydrolyzed.[2]

Inverting glycosyltransferases employ a single displacement-like mechanism involving acid−base catalytic assistance to the departure/attack of the leaving group/nucleophile of the reaction (Scheme 2).[24,30] In flavonoid β-glycosyltransferases of family GT-1 a conserved histidine likely fulfils the function of the catalytic base.[31,32]

*Os*CGT has been well-characterized biochemically[20,24] and has already been used for mechanistic study by site-directed mutagenesis.[24,30] This enzyme was therefore also used here. The relevant His24 of *Os*CGT (Figure 1) was substituted by an alanine, which is a residue incompetent in the catalysis considered. To determine whether indeed the acid−base catalytic function was destroyed by the site-directed replacement, we compared the relative reaction rates of wildtype enzyme and H24A variant for β-glucosylation of phloretin from UDP-glucose and formation of UDP-glucose from β-glucosyl fluoride. Phloretin glucosylation is expected to benefit from catalytic facilitation by a general base as follows. The

Scheme 2. Inverting C-Glycosylation under Base Catalysis (a) and Reaction with β-Glucosyl Fluoride in Absence of Base Catalysis (b)[a]
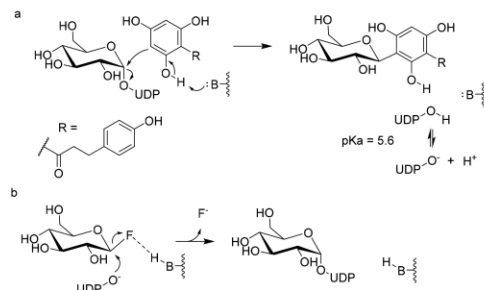
[a]The relevant pK_a of the phloretin acceptor in (a) is 9.4, as discussed in the text. Reaction with β-glucosylfluoride occurs with the catalytic enzyme group protonated or unprotonated, or removed by mutagenesis. Note: tautomerism involved in the C-glycosylation is not shown in (a).
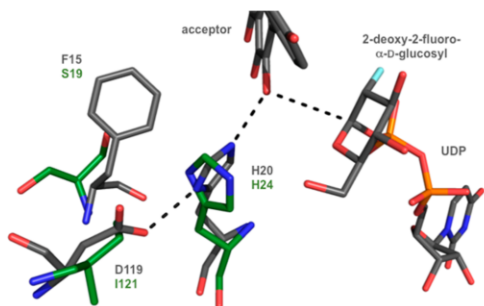

Figure 1. Superimposition of active sites in a structural model of *Os*CGT (green, Phyre2[34]-obtained model from best match) and the experimental structure of *V. vitifera* O-glycosyltransferase (PDB: 2C1Z; gray) bound with the inactive donor analog UDP-2-fluoro-glucose and the kaempferol acceptor. The putative catalytic acid−base is His24 in *Os*CGT, corresponding to His20 in the O-glycosyltransferase.

phloretin molecule shows 3 pK_a values: 7.0, 9.4, and 10.5.[33] Upon 2′-β-O-glucosylation of phloretin the 9.4 pK_a was lost, while the other two pK_a values remained.[33] The phloretin pK_a relevant for 3′-β-C-glucosylation (Scheme 2a) is therefore the one of 9.4. Catalytic deprotonation of the 2′-hydroxy group is thus required for β-glucosylation at C3′. Conversion of β-glucosyl fluoride, by contrast, is expected to proceed readily without assistance from the relevant catalytic residue in the enzyme now acting as a general acid (Scheme 2b).

Compared to wildtype *Os*CGT, there was a large (200-fold) loss of activity in the H24A variant for the canonical reaction dependent on base catalysis while there was only a small (3.2-fold) loss of activity for the reaction with the β-glucosyl fluoride (Table 2). The low basal activity of the wildtype enzyme toward the β-glucosyl fluoride requires one to interpret these results with caution. However, the findings appear nonetheless supportive of the notion that His24 provides base

Table 2. Glucosylation of UDP from β-Glucosyl Fluoride and of Phloretin from UDP-Glucose by *Os*CGT and Variants Thereof

	specific activity [nmol min^{-1} mg^{-1}]	
*Os*CGT	UDP + β-glucosyl fluoride[a]	phloretin + UDP-glucose[b]
wildtype	7.1	2200
H24A	2.2	11
I121D	155	34
S19A	11	390

[a]2 mM UDP, 4 mM β-glucosyl fluoride, 1000 μg mL^{-1} *Os*CGT (variant); [b]1 mM phloretin, 2 mM UDP-glucose, 0.1 mg mL^{-1} *Os*CGT (variant); All reactions were performed at 30 °C and pH 8.0. Specific activities are from quadruplicate determinations and have SD of ≤10%.

catalysis to C-glucosylation of phloretin by *Os*CGT (Scheme 2).

We considered additional experiments to support the notion that conversion of β-glucosyl fluoride into UDP-glucose does not require acid−base catalysis, whereas conversion of phloretin β-glucoside effectively needs it. A pH-rate profile analysis carried out with the I121D variant of *Os*CGT provided that support. Reason to choose the I121D variant was that unlike the wildtype enzyme and H24A variant, which are both highly specific for 3′-β-C-glucosylation of phloretin, the I121D variant is a bifunctional O-/C-glucosyltransferase in the reaction with phloretin, as shown in earlier work[24,33] (see also Table S2, Supporting Information). When offered the 2′-O-β-glucoside phlorizin in the presence of UDP, therefore, the I121D variant catalyzed the reverse reaction to form UDP-glucose. Wildtype *Os*CGT[33] and H24A variant (this study) were completely inactive toward phlorizin under the same conditions. Please note here that the 3′-β-C-glucosylation of phloretin from UDP-glucose is not detectably reversible.[24,33] There is significant thermodynamic difference between nothofagin and phlorizin.[33] These observations implied the use of phlorizin and the I121D variant of *Os*CGT for the pH study considered. The pH-rate profiles obtained with the I121D variant are shown in Figure S4 (Supporting Information). The UDP-glucose formation rate from β-glucosyl fluoride was independent of pH in the pH range 5.5−8.0 whereas the corresponding rate from phlorizin decreased drastically upon increasing the pH from 6.0 to 8.0. These pH profiles suggest the involvement of a molecular group on the enzyme that must be protonated, probably to provide acid catalysis, for activity toward phlorizin. This molecular group, or the protonation state thereof, is not important for activity toward the β-glucosyl fluoride. Mechanistic interpretation of the reactivities of the H24A variant as compared with the reactivities of wildtype *Os*CGT is therefore supported.

In another set of experiments, we sought to demonstrate the use of β-glucosyl fluoride to analyze *Os*CGT specificity. The β-glucosyl fluoride was of interest because it effectively eliminates noncovalent enzyme−substrate interactions at the subsite accommodating the phloretin leaving group/nucleophile. Such noncovalent interactions would be expected to contribute to the precise positioning of the substrates at the enzyme active site and so play an important role in catalysis. They clearly do so in other glycosyltransferases.[30,35,36] To test this idea for *Os*CGT, we characterized two enzyme variants (S19A, I121D) with substitutions that potentially interfere with positioning for
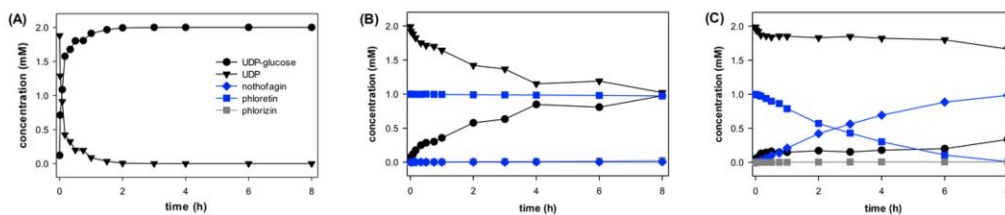
Figure 2. Glycosyltransferase cascade reaction for phloretin 3′-β-C-glucosylation from β-glucosyl fluoride via UDP-glucose is shown. Reactions were done at 30 °C and pH 8.0. (A) Conversion of β-glucosyl fluoride (4 mM) and UDP (2 mM) by I121D variant (500 μg mL^{-1}). (B) Conversion as in (A) in the presence of phloretin (1 mM). (C) Cascade reaction using the substrate concentrations as in (B) in the presence of wildtype OsCGT (50 μg mL^{-1}) and I121D variant (500 μg mL^{-1}). Note: trilobatin was also measured. It was not formed in detectable amounts and is therefore not shown in Figure 2C.

catalysis in the ternary enzyme−substrate complex. As can be inferred from Figure 1, the substitution of Ile121 by Asp is likely to perturb the microenvironment of His24. Substitution of Ser19 by Ala could affect binding of the acceptor substrate (phloretin). Table 2 compares specific activities of I121D and S19A variants for conversion of β-glucosyl fluoride and C-β-glucosylation of phloretin. Contrary to the wildtype enzyme that showed high (310-fold) preference for glucosylation of phloretin, the I121D variant preferred reaction with the β-glucosyl fluoride (4.6-fold) and was 22-fold more active toward β-glucosyl fluoride than wildtype OsCGT. The S19A variant also showed change in reaction preference toward, relatively speaking, better utilization of the β-glucosyl fluoride, although the effect was less pronounced than in the I121D variant.

Observation for the I121D variant, that the β-glucosyl fluoride can elicit activity in a mutated glycosyltransferase strongly impaired in the canonical reaction, carries two major implications. One is, that by enforcing glycosyltransferases to react through an enzyme−substrate complex that leaves the acceptor-binding site unoccupied, glycosyl fluorides represent specific tools for the mechanistic characterization of these enzymes. The other is that conversion of glycosyl fluorides might serve as means of synthesizing sugar nucleotides, with the interesting option of generating sugar nucleotide in situ for use by other enzymes. We show realization of this concept[18,23,37,38] in a glycosyltransferase cascade reaction (cf. Scheme 1) wherein the I121D variant was coupled with wildtype OsCGT for synthesis of nothofagin from β-glucosyl fluoride in the presence of UDP (Figure 2). By combining enzymes suitable for UDP-glucose formation from β-glucosyl fluoride (I121D variant; Table 2; Figure 2A) and phloretin glucosylation from UDP-glucose (OsCGT wildtype; Table 2), respectively, the nothofagin production became efficient (Figure 2C), which it was not when only a single glycosyltransferase, I121D (Figure 2B) or wildtype (data not shown), was used. In Figure 2C, the two enzymes were used to have wildtype activity for glucosylation of phloretin (85 nmol min^{-1} mL^{-1}) balanced with I121D activity for formation of UDP-glucose (78 nmol min^{-1} mL^{-1}). However, the UDP-glucose formation by I121D variant was inhibited about 11.5-fold by the 1 mM phloretin present, as seen from a comparison of the reaction time courses in Figure 2A and 2B (see also Figure S5, Supporting Information). Therefore, only little UDP-glucose accumulated in the reaction of Figure 2C. Nonetheless, the UDP-glucose released was sufficient to promote efficient conversion of phloretin into nothofagin. Note that in the case of a glycosyltransferase showing suitable

activities for both reactions (e.g., UGT708A6), the two-step one-pot conversion of β-glucosyl fluoride into nothofagin might be even carried out by a single enzyme.

In summary, we show that in several flavonoid β-glucosyltransferases of family GT-1, β-glucosyl fluoride can function as alternative donor for inverting glucosyl transfer to UDP. Using β-glucosyl fluoride as substrate for the kinetic characterization of OsCGT variants, we show that mechanistic tools originally developed for glycoside hydrolases[13,14,18] are also applicable in principle to glycosyltransferases. Enzymatic glycosylation of phloretin from β-glucosyl fluoride via intermediary UDP-glucose represents a new variation of the seminal concept of glycosyltransferase reverse reactions for synthetic application.[23,37,38] In terms of scope of use, glycosyl fluorides complement glycosides with aromatic leaving groups as non-natural substrates for glycosyltransferases. Practical application of the relatively unstable β-glucosyl fluoride must however take into account its fast spontaneous hydrolysis. Glycosyltransferases showing good activity toward the β-glucosyl fluoride (e.g., I121D variant of OsCGT, UGT708A6) are therefore essential for glycosylation cascade efficiency. More research is needed to understand requirements of glycosyltransferase structure for utilization of glycosyl fluorides as substrates. Evidence for OsCGT, that a single residue substitution (I121D) caused substantial enhancement of activity with the β-glucosyl fluoride, fuels the hope that glycosyltransferases naturally unable to use glycosyl fluorides might be engineered to do so.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acscatal.8b02685.

> Methods describing glycosyltransferases (cloning, preparation, purification), glucosyl fluorides (preparation, stability evaluation) and analytical procedures used; representative HPLC traces; solvolysis of α- and β-glucosyl fluoride; relative abundance of phloretin glucosides formed by the glycosyltransferases used; pH dependence of UDP-glucose formation from phlorizin and β-glucose fluoride; and inhibition of OsCGT and OsCGT I121D by phloretin (PDF)

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: bernd.nidetzky@tugraz.at.

ORCID ⊙

Bernd Nidetzky: 0000-0002-5030-2643

■ **REFERENCES**

(1) Toshima, K. Glycosyl Fluorides in Glycosidations. *Carbohydr. Res.* **2000**, *327*, 15−26.

(2) Williams, S. J.; Withers, S. G. Glycosyl Fluorides in Enzymatic Reactions. *Carbohydr. Res.* **2000**, *327*, 27−46.

(3) Pelletier, G.; Zwicker, A.; Allen, C. L.; Schepartz, A.; Miller, S. J. Aqueous Glycosylation of Unprotected Sucrose Employing Glycosyl Fluorides in the Presence of Calcium Ion and Trimethylamine. *J. Am. Chem. Soc.* **2016**, *138*, 3175−3182.

(4) Hancock, S. M.; Vaughan, M. D.; Withers, S. G. Engineering of Glycosidases and Glycosyltransferases. *Curr. Opin. Chem. Biol.* **2006**, *10*, 509−519.

(5) Pérez, X.; Faijes, M.; Planas, A. Artificial Mixed-Linked β-Glucans Produced by Glycosynthase-Catalyzed Polymerization: Tuning Morphology and Degree of Polymerization. *Biomacromolecules* **2011**, *12*, 494−501.

(6) Shoda, S. I.; Uyama, H.; Kadokawa, J. I.; Kimura, S.; Kobayashi, S. Enzymes as Green Catalysts for Precision Macromolecular synthesis. *Chem. Rev.* **2016**, *116*, 2307−2413.

(7) Zhan, C. G.; Dixon, D. A. Hydration of the Fluoride Anion: Structures and Absolute Hydration Free Energy from First-Principles Electronic Structure Calculations. *J. Phys. Chem. A* **2004**, *108*, 2020−2029.

(8) Nidetzky, B.; Griessler, R.; Schwarz, A.; Splechtna, B. Cellobiose Phosphorylase from *Cellulomonas uda*: Gene Cloning and Expression in *Escherichia coli*, and Application of the Recombinant Enzyme in a 'Glycosynthase-Type' Reaction. *J. Mol. Catal. B: Enzym.* **2004**, *29*, 241−248.

(9) Nakai, H.; Hachem, M. A.; Petersen, B. O.; Westphal, Y.; Mannerstedt, K.; Baumann, M. J.; Dilokpimol, A.; Schols, H. A.; Duus, J.; Svensson, B. Efficient Chemoenzymatic Oligosaccharide Synthesis by Reverse Phosphorolysis Using Cellobiose Phosphorylase and Cellodextrin Phosphorylase from *Clostridium thermocellum*. *Biochimie* **2010**, *92*, 1818−1826.

(10) Zhang, Y.; Bommuswamy, J.; Sinnott, M. L. Kinetic Isotope Effect Study of Transition States for the Hydrolyses of α- and β-Glucopyranosyl Fluorides. *J. Am. Chem. Soc.* **1994**, *116*, 7557−7563.

(11) Banait, N. S.; Jencks, W. P. Reactions of Anionic Nucleophiles with α-D-Glucopyranosyl Fluoride in Aqueous Solution through a Concerted, A_ND_N (S_N2) Mechanism. *J. Am. Chem. Soc.* **1991**, *113*, 7951−7958.

(12) Chan, J.; Sannikova, N.; Tang, A.; Bennet, A. J. Transition-State Structure for the Quintessential S_N2 Reaction of a Carbohydrate: Reaction of α-Glucopyranosyl Fluoride with Azide Ion in Water. *J. Am. Chem. Soc.* **2014**, *136*, 12225−12228.

(13) Ly, H.-D.; Withers, S.-G. Mutagenesis of Glycosidases. *Annu. Rev. Biochem.* **1999**, *68*, 487−522.

(14) Zechel, D. L.; Withers, S.-G. Glycosidase Mechanisms: Anatomy of a Finely Tuned Catalyst. *Acc. Chem. Res.* **2000**, *33*, 11−18.

(15) Mackenzie, L. F.; Wang, Q.; Warren, R. A. J.; Withers, S. G. Glycosynthases: Mutant Glycosidases for Oligosaccharide Synthesis. *J. Am. Chem. Soc.* **1998**, *120*, 5583−5584.

(16) Honda, Y.; Kitaoka, M. The First Glycosynthase Derived from an Inverting Glycoside Hydrolase. *J. Biol. Chem.* **2006**, *281*, 1426−1431.

(17) Malet, C.; Planas, A. From β-Glucanase to β-Glucansynthase: Glycosyl Transfer to α-Glycosyl Fluorides Catalyzed by a Mutant Endoglucanase Lacking its Catalytic Nucleophile. *FEBS Lett.* **1998**, *440*, 208−212.

(18) Lougheed, B.; Ly, H. D.; Wakarchuk, W. W.; Withers, S. G. Glycosyl Fluorides Can Function as Substrates for Nucleotide Phosphosugar-Dependent Glycosyltransferases. *J. Biol. Chem.* **1999**, *274*, 37717−37722.

(19) Schmölzer, K.; Gutmann, A.; Diricks, M.; Desmet, T.; Nidetzky, B. Sucrose Synthase: A Unique Glycosyltransferase for Biocatalytic Glycosylation Process Development. *Biotechnol. Adv.* **2016**, *34*, 88−111.

(20) Bungaruang, L.; Gutmann, A.; Nidetzky, B. Leloir Glycosyltransferases and Natural Product Glycosylation: Biocatalytic Synthesis of the C-Glucoside Nothofagin, a Major Antioxidant of Redbush Herbal Tea. *Adv. Synth. Catal.* **2013**, *355*, 2757−2763.

(21) Kitahata, S.; Brewer, C. F.; Genghof, D. S.; Sawai, T.; Hehre, E. J. Scope and Mechanism of Carbohydrase Action. Stereocomplementary Hydrolytic and Glucosyl-transferring Actions of Glucoamylase and Glucodextranase with α- and β-D-Glucosyl Fluoride. *J. Biol. Chem.* **1981**, *256*, 6017−6026.

(22) Hehre, E. J.; Brewer, C. F.; Genghof, D. S. Scope and Mechanism of Carbohydrase Action. Hydrolytic and Nonhydrolytic Actions of β-Amylase on α- and β-Maltosyl Fluoride. *J. Biol. Chem.* **1979**, *254*, 5942−5950.

(23) Lairson, L. L.; Wakarchuk, W. W.; Withers, S. G. Alternative Donor Substrates for Inverting and Retaining Glycosyltransferases. *Chem. Commun.* **2007**, *43*, 365−367.

(24) Gutmann, A.; Nidetzky, B. Switching between O- and C-Glycosyltransferase through Exchange of Active-Site Motifs. *Angew. Chem., Int. Ed.* **2012**, *51*, 12879−12883.

(25) Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P. M.; Henrissat, B. The Carbohydrate-Active Enzymes Database (CAZy) in 2013. *Nucleic Acids Res.* **2014**, *42*, D490−D495.

(26) Ferreyra, M. L. F.; Rodriguez, E.; Casas, M. I.; Labadie, G.; Grotewold, E.; Casati, P. Identification of a Bifunctional Maize C- and O-Glucosyltransferase. *J. Biol. Chem.* **2013**, *288*, 31678−31688.

(27) Nagatomo, Y.; Usui, S.; Ito, T.; Kato, A.; Shimosaka, M.; Taguchi, G. Purification, Molecular Cloning and Functional Characterization of Flavonoid C-Glucosyltransferases from *Fagopyrum eculentum* M. (Buckwheat) Cotyledon. *Plant J.* **2014**, *80*, 437−448.

(28) Lepak, A.; Gutmann, A.; Kulmer, S. T.; Nidetzky, B. Creating a Water-Soluble Resveratrol-Based Antioxidant by Site-Selective Enzymatic Glucosylation. *ChemBioChem* **2015**, *16*, 1870−1874.

(29) Hefner, T.; Arend, J.; Warzecha, H.; Siems, K.; Stöckigt, J. Arbutin Synthase a Novel Member of the NRD1ß Glycosyltransferase Family, is a Unique Multifunctional Enzyme Converting Various Natural Products and Xenobiotics. *Bioorg. Med. Chem.* **2002**, *10*, 1731−1741.

(30) Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G. Glycosyltransferases: Structures, Functions, and Mechanisms. *Annu. Rev. Biochem.* **2008**, *77*, 521−55.

(31) Liang, D. M.; Liu, J. H.; Wu, H.; Wang, B. B.; Zhu, H. J.; Qiao, J. J. Glycosyltransferases: Mechanisms and Applications in Natural Product Development. *Chem. Soc. Rev.* **2015**, *44*, 8350−8374.

(32) Offen, W.; Martinez-Fleites, C.; Yang, M.; Kiat-Lim, E.; Davis, B. G.; Tarling, C. A.; Ford, C. M.; Bowles, D. J.; Davies, G. J. Structure of a Flavonoid Glucosyltransferase Reveals the Basis for Plant Natural Product Modification. *EMBO J.* **2006**, *25*, 1396−1405.

(33) Gutmann, A.; Krump, C.; Bungaruang, L.; Nidetzky, B. A Two-Step O-to C-Glycosidic Bond Rearrangement Using Complementary Glycosyltransferase Activities. *Chem. Commun.* **2014**, *50*, 5465−5468.

(34) Kelley, L. A.; Sternberg, M. J. E. Protein Structure Prediction on the Web: A Case Study Using the Phyre Server. *Nat. Protoc.* **2009**, *4*, 363−371.

(35) Chang, A.; Singh, S.; Helmich, K. E.; Goff, R. D.; Bingman, C. A.; Thorson, J. S.; Phillips, G. N. Complete Set of Glycosyltransferase Structures in the Calicheamicin Biosynthetic Pathway Reveals the Origin of Regiospecificity. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 17649−17654.

(36) Albesa-Jové, D.; Guerin, M. E. The Conformational Plasticity of Glycosyltransferases. *Curr. Opin. Struct. Biol.* **2016**, *40*, 23−32.

(37) Zhang, C.; Griffith, B. R.; Fu, Q.; Albermann, C.; Fu, X.; Lee, I. K.; Li, L.; Thorson, J. S. Exploiting the Reversibility of Natural Product Glycosyltransferase-Catalyzed Reactions. *Science* **2006**, *313*, 1291−1294.

(38) Gantt, R. W.; Peltier-Pain, P.; Cournoyer, W. J.; Thorson, J. S. Using Simple Donors to Drive the Equilibria of Glycosyltransferase-Catalyzed Reactions. *Nat. Chem. Biol.* **2011**, *7*, 685.

Supporting Information

β-Glucosyl fluoride as Reverse-Reaction Donor Substrate and Mechanistic Probe of Inverting Sugar Nucleotide-Dependent Glycosyltransferases

Alexander Lepak[¶], Alexander Gutmann,[¶] Bernd Nidetzky[¶][§] *

¶ Institute of Biotechnology and Chemical Engineering, Graz University of Technology, NAWI Graz, Petersgasse 12, 8010 Graz, Austria

§ Austrian Centre of Industrial Biotechnology (acib), Petersgasse 14, 8010 Graz, Austria

* Corresponding Author: Bernd Nidetzky

Institute of Biotechnology and Biochemical Engineering, TU Graz; Petersgasse 12, A-8010 Graz, Austria;

e-mail: bernd.nidetzky@tugraz.at; phone: +433168738400; fax: +433168738434

1. Methods

1.1. Chemicals and reagents

Unless otherwise indicated, all chemicals were from Sigma-Aldrich in the highest purity available. Phloretin, phlorizin, UMP, UDP and UDP-g (97% purity) and UDP-glucose were from Carbosynth. *Strep*-Tactin Sepharose and desthiobiotin were from IBA. BCA assay kit and DNA modifying enzymes were from Thermo Scientific.

1.2. Glucosyl fluoride substrates

The α- and β-glucosyl fluoride were prepared by deprotection of 2,3,4,6-tetra-O-acetyl-α- or β-D-glucopyranosyl fluoride using sodium methoxide in methanol.[1] Fresh 1 M sodium methoxide was prepared by dissolving 23.5 mg sodium in 1 mL methanol. The Zemplen deprotection was performed at 0°C under nitrogen atmosphere by treating one equivalent of substrate dissolved in methanol with 0.4 equivalents of sodium methoxide. Reaction progress was monitored with TLC. A mixture of ethylacetate and methanol (9:1, by volume) was used as eluent. Once the reaction was complete, the solution was neutralized using an ion-exchange resin (Dowex 50W, H[+]-form), filtered, aliquoted and the solvent was removed under vacuum. The glucosyl fluorides were dissolved in D_2O and checked for purity by NMR before each experiment. Enzymatic reactions were started exactly 30 min after dissolution of the glucosyl fluoride in D_2O.

1.3. Construction of expression strains

Codon optimized genes for UGT706A6 (GenBank: ACF81582.1)[2], UGT708C2 (GenBank: BAP90361.1)[3] and arbutin synthase (*Rs*As, GenBank: CAC35167.1)[4] were from GenScript and received in pUC57 vectors. For expression in *E. coli*, the respective gene was integrated into pET-STRP3, which is a custom-made derivate of pET-24d that enables protein expression with an N-terminally fused *Strep*-tag II. The vector was received from the group of Prof. Robert Edwards (Centre for Bioactive Chemistry, Durham University, UK).[5] To clone the glycosyltransferase genes into pET-STRP3, the two flanking restriction sites *Nde*I and *Xho*I were introduced by PCR using forward and reverse primers, respectively (Table S1). The resulting fragment was purified on agarose gel, digested with *Nde*I and *Xho*I fast digest restriction enzymes and cloned into the respective sites of the pET-STRP3 vector. The correct sequence was verified by sequencing of the complete gene. The expression strains were prepared by transformation of electro-competent *E. coli* BL21-Gold (DE3) cells.

Construction of the *E. coli* BL21 Gold (DE3) expression strains for sucrose synthase (*Gm*SuSy; GenBank: AF030231)[6], UGT71A15 (GenBank: DQ103712)[7], *Pc*OGT (UGT88F2; GenBank: FJ854496)[8], *Os*CGT (GenBank: FM179712)[8] and *Os*CGT I121D[8] was described elsewhere. Further active site mutants of *Os*CGT were created by two-stage PCR using the *Os*CGT wildtype or I121D mutant gene as template. The required primers are listed in Table S1.

S2

111

Table S1. Primers used to amplify glycosyltransferase genes or to introduce point mutations.

Enzyme	Primer	Sequence
UGT708A6	UGT708A6_fwd	TAAC<u>CATATG</u>GCTGCTAATGGTGGCG
	UGT708A6_rev	GGTG<u>CTCGAG</u>TTATTTACGTTCGGCGTCACGACAGCG
UGT708C2	UGT708C2_fwd	TAAC<u>CATATG</u>ATGGGTGATCTGACGACGAG
	UGT708C2_rev	GGTG<u>CTCGAG</u>TTAGCGTTTCAGGGAGCCGATAATGTTCATCAG
RsAs	RsAs_fwd	TAAC<u>CATATG</u>GAACACACGCCGCACATTG
	RsAs_rev	GGTG<u>CTCGAG</u>TTAGGTGGAGGAGATTTTATTTTCCCATTTACACG
*Os*CGT H24A / *Os*CGT H24A/I121D	OsCGT_H24A_fw	GGGGACGAGCGCGCCCATGCCGGCGCTC
	OsCGT_H24A_rv	GCCGGCATGGGCGCGCTCGTCCCCTTCG
*Os*CGT S19A	OsCGT_S19A_fw	CCATGCCGGCCGCCGGGATGAGCACCACATGC
	OsCGT_S19A_rv	GCTCATCCCGGCGGCCGGCATGGGCCACCTC

1.4. Enzyme expression and purification

For expression of the glycosyltransferases, the respective strain was grown at 37°C and 120 rpm in 1 L baffled shake flasks in 250 mL LB-medium containing 50 µg mL^{-1} kanamycin. The cells were grown until an optical density at 600 nm of around 0.8 was reached. The temperature was then decreased to 18°C and expression was induced by addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were harvested after 18 h by centrifugation (30 min, 4°C, 5000 rpm), resuspended in water and frozen at -20°C. Cell lysis was done by sonication (10 s on, 20 s off, 6 min total on-time, 400 watt). The cell lysate was cleared of debris (20,000 × g, 60 min) and filtered through a 1.2 µm filter. It was loaded onto two connected 5 mL StrepTrap-HP columns and purification was done according to the protocol supplied by the manufacturer. Fractions containing the enzyme were pooled, concentrated using centrifugal concentrators and purified for a second time by anion exchange chromatography on a 5 mL HiTrapQ HP column. A gradient from 0 - 1 M NaCl (100 mM Tris/HCl buffer, pH 7.5) over 10 column volumes was used. Fractions containing the glycosyltransferase were concentrated and buffer exchanged to 25 mM HEPES (pH 7.5) using centrifugal concentrators. Expression and purification of *Gm*SuSy was done as reported previously.[6] Enzyme purity was checked by SDS-polyacrylamide gel electrophoresis. Protein concentrations were determined by BCA assay using BSA as standard. Enzyme stocks were stored at -20°C.

1.5. HPLC based quantification of phloretin (glucosides) and nucleotides

Phloretin, phloretin glucosides, UMP, UDP and UDP-glucose were quantified by a reversed-phase ion-pairing HPLC method.[9] Tetra-*n*-butylammonium bromide (TBAB) served as ion-pairing reagent. HPLC analysis was performed on a reversed phase C18 column (Kinetex^{TM} 5 µm C18 100 Å column 50 × 4.6 mm, Phenomenex) at 35°C. A 20 mM potassium phosphate buffer (pH 5.9) containing 40 mM TBAB was used as mobile phase A. It was prepared by dissolving TBAB in a 5-fold diluted 100 mM potassium phosphate buffer (pH 5.9) stock. Acetonitrile was used as mobile phase B. Separation was achieved using the following method at a constant flow rate of 2 mL min^{-1}: 10% B (1 min), 10 - 40% B (3.2 min), 40 - 90% B (0.01 min), 90% B (0.79 min), 90 - 10% B (0.01 min), 10% B (1.99 min). UMP, UDP and UDP-glucose were monitored at 262 nm. Phloretin and its glucosides were monitored at 288 nm. Authentic standards of all compounds were used for calibration and concentrations were calculated from peak areas.

The limit of detection for nucleotides was 200 nM.[9] It allowed identification of residual UDP and UDP-glucose in purified enzyme. Measuring conversions of acceptors with glucosyl fluorides in the absence of UDP made the use of UDP-free enzyme preparations essential. To validate nucleotide free preparation of the glycosyltransferases, 500 µL samples containing 2 mg mL^{-1} (~ 40 µM) of enzyme were denatured by the addition of 500 µL of methanol. Precipitated protein was pelleted by centrifugation (20,000 × g, 60 min) and

S3

112

the remaining supernatant tested for presence of nucleotides. This control guaranteed that less than 2% of protein contained nucleotides.

1.6. Enzymatic glycosylations

Enzymatic conversions were monitored using the above described HPLC assay. Unless mentioned otherwise, reactions were performed at 30°C in 50 mM HEPES buffer (pH 8) containing 13 mM $MgCl_2$, 50 mM KCl, 0.13% (w/v) BSA and 20% (v/v) DMSO. Reactions were started by enzyme addition. Withdrawn samples were immediately mixed with an equal volume of acetonitrile to stop enzymatic conversions. Precipitated proteins were removed before HPLC analysis by centrifugation (15 min at 13,200 rpm). Specific activities were calculated from linear initial reaction rates determined from samples withdrawn within the first 15 to 60 min of conversion.

UDP-glucose formation by sucrose synthase from *Glycine max* (*Gm*SuSy) was studied at pH 7.0. Reaction mixtures contained 2 mM α- or β-glucosyl fluoride, 1 mM UDP and 100 µg mL^{-1} *Gm*SuSy. Conversions were followed in absence and presence of 10 mM fructose.

To study glycosyltransferase catalyzed synthesis of UDP-glucose 1 mg mL^{-1} of the respective enzyme was incubated with 2 mM UDP and 4 mM α- or β-glucosyl-fluoride. Replacement of UDP with 2 mM UMP was tested with 1 mg mL^{-1} *Os*CGT and 4 mM α- or β-glucosyl-fluoride.

Glucosylation of 1 mM phloretin was coupled to in-situ synthesis of UDP-glucose from 2 mM UDP and 4 mM α- or β-glucosyl-fluoride. 1 mg mL^{-1} of glycosyltransferase were used. Alternatively, 1 mM phloretin was directly glucosylated from 2 mM UDP-glucose with 0.1 mg mL^{-1} of enzyme. Using 1 mg mL^{-1} of *Os*CGT, also glucosylation of 1 mM phloretin from 2 mM UMP and 4 mM β-glucosyl-fluoride was tested.

The pH dependence of UDP-glucose formation by *Os*CGT I121D (50 µg mL^{-1}) was studied. Reactions containing 4 mM β-glucosyl fluoride and 2 mM UDP were buffered by 50 mM MES (pH 5.5 -7.0) or 50 mM HEPES (pH 6.0 – 8.0). Additionally the pH dependency on UDP-glucose formation from 0.5 mM phlorizin and 5 mM UDP was studied. Reactions were catalyzed by 1 mg mL^{-1} of *Os*CGT I121D and buffered by 50 mM HEPES (pH 6.0-8.0).

Time dependent UDP-glucose formation and phloretin glucosylation by *Os*CGT and *Os*CGT I121D were followed for 8 h. UDP-glucose was synthesized by 500 µg mL^{-1} *Os*CGT I121D from 4 mM β-glucosyl-fluoride and 2 mM UDP. Conversion of in-situ formed UDP-glucose was studied by adding also 1 of mM phloretin. 500 µg mL^{-1} *Os*CGT I121D or 50 µg mL^{-1} *Os*CGT were used. Either separately or in combination.

Inhibition of UDP-glucose formation by phloretin was tested with 50 µg mL^{-1} *Os*CGT I121D and 500 µg mL^{-1} *Os*CGT WT. Conversion of 2 mM UDP with 4 mM β-glucosyl fluoride was followed for 60 min in absence and presence of 1 mM phloretin.

1.7. NMR spectroscopy (including in-situ analysis of solvolysis)

A Varian (Agilent) INOVA 500-MHz NMR spectrometer (Agilent Technologies, Santa Clara, California, USA) and the VNMRJ 2.2D software were used for all measurements. ^{1}H NMR spectra (499.98 MHz) were measured on a 5 mm indirect detection PFG-probe. NMR data were recorded from purified glucosyl-fluorides. In-situ analysis of solvolysis was performed at 30°C in a total volume of 600 µL potassium phosphate buffer (25 mM; pD 7.0) in D_2O containing 50 mM of either α-glucosyl-fluoride or 4 mM β-glucosyl-fluoride. Mestrenova 10 was used for evaluation of spectra. To measure the solvolysis of glucosyl fluorides the integrals of C-1 proton signals were compared to integrals of C-1 proton signals from released α- and β-glucose.

S4

2. Results
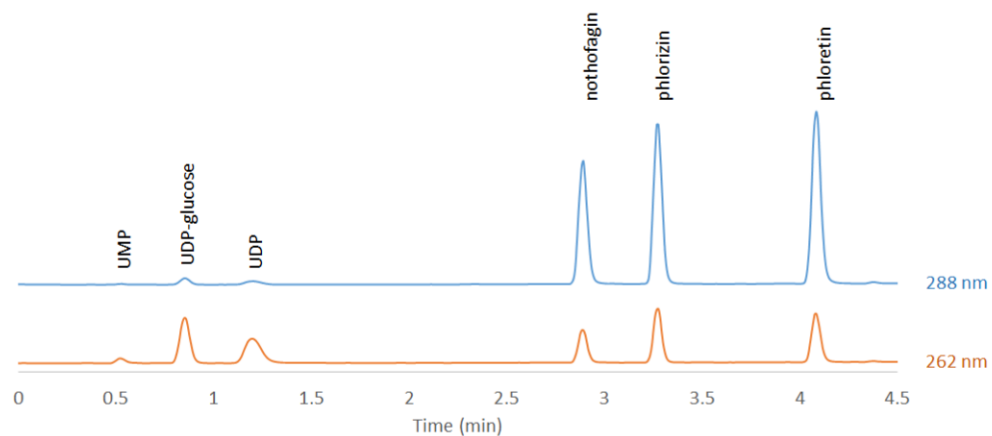
Figure S1. UMP, UDP-glucose, UDP, nothofagin, phlorizin and phloretin were quantified by reversed-phase C-18 HPLC using a TBAB based ion pairing protocol. Nucleotides and phloretin derivatives were monitored at 262 and 288 nm, respectively.
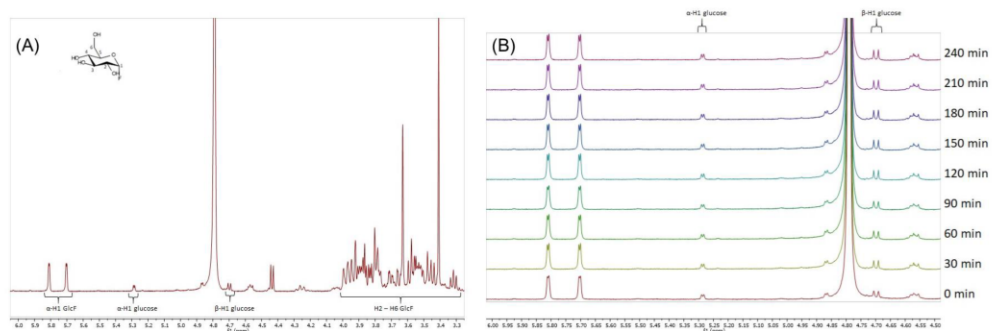
Figure S2. A) The ^{1}H NMR spectrum of α-glucosyl fluoride in 50 mM potassium phosphate buffer (pD 7.0) is shown. B) In-situ ^{1}H NMR measurement of α-glucosyl fluoride solvolysis in 50 mM potassium phosphate buffer (pD 7.0) is shown.
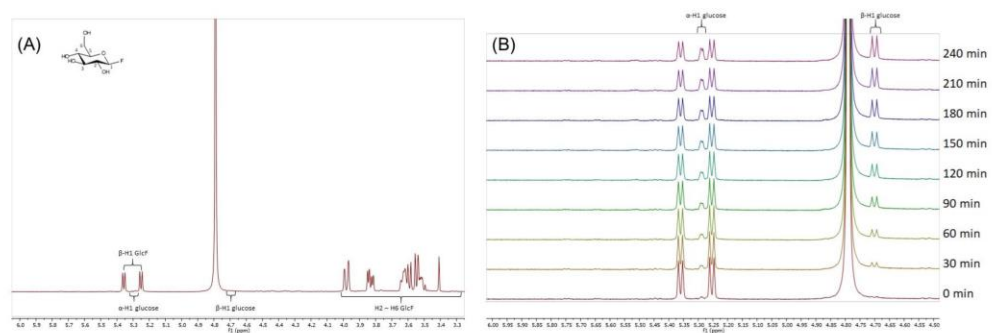
Figure S3. A) The ^{1}H NMR spectrum of β-glucosyl fluoride in 50 mM potassium phosphate buffer (pD 7.0) is shown. B) In-situ ^{1}H NMR measurement of β-glucosyl fluoride solvolysis in 50 mM potassium phosphate buffer (pD 7.0) is shown.

Table S2: Relative abundance of products of enzymatic phloretin glucosylation displayed in %.[a]

Enzyme	Glucoside		
	3'-*C* (nothofagin)	2'-*O* (phlorizin)	4'-*O* (trilobatin)
*Os*CGT	100	0	0
UGT708A6	100	0	0
UGT708C2	100	0	0
UGT71A15	0	23	77
*Pc*OGT	0	100	0
RsAS	0	0	100
*Os*CGT H24A	100	0	0
*Os*CGT I121D	12	17	71
*Os*CGT S19A	100	0	0
*Os*CGT H24A/I121D	10	0	90

[a] 2 mM UDP-glucose, 1 mM phloretin, 1 mg mL^{-1} purified enzyme, pH 8.0, 30°C

S6

115
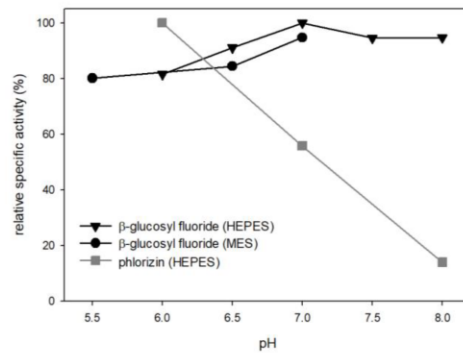
Figure S4. The pH-dependence of UDP-glucose synthesis by *Os*CGT I121D with β-glucosyl fluoride or phlorizin as sugar donor is shown. Reaction conditions: 2 mM UDP, 4 mM β-glucosyl fluoride, 50 µg mL^{-1} *Os*CGT I121D or 5 mM UDP, 0.5 mM phlorizin, 1000 µg mL^{-1} *Os*CGT I121D. Maximum activities of *Os*CGT I121D (100%) with β-glucosyl fluoride and phlorizin were 130 mU mg^{-1} and 1.38 mU mg^{-1}, respectively.
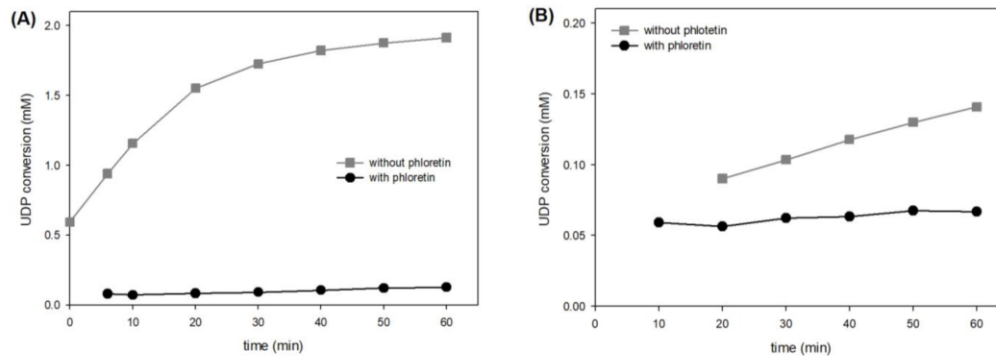
Figure S5. Time courses show the inhibition of UDP-glucose synthesis from β-glucosyl fluoride in presence of phloretin. 50 µg mL^{-1} *Os*CGT I121D and 500 µg mL^{-1} *Os*CGT WT were incubated at pH 8 with 4 mM β-glucosyl fluoride in presence of 2 mM UDP. For inhibition 1 mM phloretin was added.

S7

116

3. References

1. Steinmann, A.; Thimm, J.; Matwiejuk, M.; Thiem, J. Formation of Homooligosaccharides Using Base-Promoted Glycosylation of Unprotected Glycosyl Fluorides. *Macromolecules* **2010**, 43, 3606-3612.

2. Ferreyra, M. L. F.; Rodriguez, E.; Casas, M. I.; Labadie, G.; Grotewold, E.; Casati, P. Identification of a Bifunctional Maize C-and O-Glucosyltransferase. *J. Biol. Chem.* **2013**, 288, 31678-31688.

3. Nagatomo, Y.; Usui, S.; Ito, T.; Kato, A.; Shimosaka, M.; Taguchi, G. Purification, Molecular Cloning and Functional Characterization of Flavonoid C-Glucosyltransferases from Fagopyrum Eculentum M. (Buckwheat) Cotyledon. *Plant J.* **2014**, 80, 437-448.

4. Hefner, T.; Arend, J.; Warzecha, H.; Siems, K.; Stöckigt, J. Arbutin Synthase, a Novel Member of the NRD1ß Glycosyltransferase Family, is a Unique Multifunctional Enzyme Converting Various Natural Products and Xenobiotics. *Bioorg. Med. Chem.* **2002**, 10, 1731-1741.

5. Dixon, D. P., Hawkins, T., Hussey, P. J., & Edwards, R. (2009). Enzyme activities and subcellular localization of members of the Arabidopsis glutathione transferase superfamily. Journal of Experimental Botany, 60(4), 1207-1218.

6. Bungaruang, L.; Gutmann, A.; Nidetzky, B. Leloir Glycosyltransferases and Natural Product Glycosylation: Biocatalytic Synthesis of the C-Glucoside Nothofagin, a Major Antioxidant of Redbush Herbal Tea. *Adv. Synth. Catal.* **2013**, 355, 2757-2763.

7. Lepak, A.; Gutmann, A.; Kulmer, S. T.; Nidetzky, B. Creating a Water-Soluble Resveratrol-Based Antioxidant by Site-Selective Enzymatic Glucosylation. *ChemBioChem* 2015, 16, 1870-1874.

8. Gutmann, A.; Nidetzky, B. Switching between O-and C-Glycosyltransferase through Exchange of Active-Site Motifs. *Angew. Chem., Int. Ed.* **2012**, 51, 12879-12883.

9. Gutmann, A.; Lepak, A.; Diricks, M.; Desmet, T.; Nidetzky, B. Glycosyltransferase Cascades for Natural Product Glycosylation: Use of Plant Instead of Bacterial Sucrose Synthases Improves the UDP-Glucose Recycling from Sucrose and UDP. *Biotechnology J.* **2017**, 12.

Chapter 7:

Decoupling of recombinant protein production from Escherichia coli cell growth enhances functional expression of plant Leloir glycosyltransferases

ARTICLE

WILEY BIOTECHNOLOGY BIOENGINEERING

Decoupling of recombinant protein production from *Escherichia coli* cell growth enhances functional expression of plant Leloir glycosyltransferases

Martin Lemmerer[1] | Juergen Mairhofer[2] | Alexander Lepak[3] | Karin Longus[3] | Rainer Hahn[4] | Bernd Nidetzky[1,3]

[1]Austrian Centre of Industrial Biotechnology, Graz, Austria

[2]enGenes Biotech GmbH, Vienna, Austria

[3]Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, NAWI Graz, Graz, Austria

[4]Department of Biotechnology, University of Natural Resources and Life Sciences Vienna, Vienna, Austria

Correspondence
Bernd Nidetzky, Austrian Centre of Industrial Biotechnology, Petersgasse 14, 8010 Graz, Austria.
Email: bernd.nidetzky@tugraz.at

Abstract

Sugar nucleotide-dependent (Leloir) glycosyltransferases from plants are important catalysts for the glycosylation of small molecules and natural products. Limitations on their applicability for biocatalytic synthesis arise because of low protein expression (≤10 mg/L culture) in standard microbial hosts. Here, we showed two representative glycosyltransferases: sucrose synthase from soybean and UGT71A15 from apple. A synthetic biology-based strategy of decoupling the enzyme expression from the *Escherichia coli* BL21(DE3) cell growth was effective in enhancing their individual (approximately fivefold) or combined (approximately twofold) production as correctly folded, biologically active proteins. The approach entails a synthetic host cell, which is able to shut down the production of host messenger RNA by inhibition of the *E. coli* RNA polymerase. Overexpression of the enzyme(s) of interest is induced by the orthogonal T7 RNA polymerase. Shutting down of the host RNA polymerase is achieved by L-arabinose-inducible expression of the T7 phage-derived Gp2 protein from a genome-integrated site. The glycosyltransferase genes are encoded on conventional pET-based expression plasmids that allow T7 RNA polymerase-driven inducible expression by isopropyl-β-D-galactoside. Laboratory batch and scaled-up (20 L) fed-batch bioreactor cultivations demonstrated improvements in an overall yield of active enzyme by up to 12-fold as a result of production under growth-decoupled conditions. In batch culture, sucrose synthase and UGT71A15 were obtained, respectively, at 115 and 2.30 U/g cell dry weight, corresponding to ~5 and ~1% of total intracellular protein. Fed-batch production gave sucrose synthase in a yield of 2,300 U/L of culture (830 mg protein/L). Analyzing the isolated glycosyltransferase, we showed that the improvement in the enzyme production was due to the enhancement of both yield (5.3-fold) and quality (2.3-fold) of the soluble sucrose synthase. Enzyme preparation from the decoupled production comprised an increased portion (61% compared with 26%) of the active sucrose synthase homotetramer. In summary,

therefore, we showed that the expression in growth-arrested *E. coli* is promising for recombinant production of plant Leloir glycosyltransferases.

1 | INTRODUCTION

Plant metabolisms for natural product biosynthesis and detoxification involve an elaborate enzymatic machinery for attaching sugars onto noncarbohydrate small-molecule structures (Bowles., Lim, Poppenberger, & Vaistij, 2006). This glycosylation machinery comprises a large set of sugar nucleotide glycosyltransferases (GTs). The enzymes also referred to collectively as Leloir GTs, transfer glycosyl residues from sugar nucleotide donors to specific positions on acceptor substrates (Liang et al., 2015). Compared with other enzymes (e.g., glycoside hydrolases; glycoside phosphorylases) that are also able to glycosylate small molecules, the GTs often present a unique combination of high chemo/regioselectivity and relatively flexible substrate specificity (R. Chen, 2018; Desmet et al., 2012; Thuan & Sohng, 2013). The plant GTs are, therefore, promising catalysts for glycoside production (for reviews, see: Lim, 2005; Nidetzky, Gutmann, & Zhong, 2018). Glycosylated derivatives of small molecules (e.g., flavonoids, terpenoids, peptides) with important applications in the food, fragrance, cosmetic, and chemical industries are synthesized efficiently using GTs (De Bruyn, Maertens, Beauprez, Soetaert, & De Mey, 2015; Desmet et al., 2012; Hofer, 2016; Hsu et al., 2018; Kim, Yang, Kim, Cha, & Ahn, 2015; Nidetzky et al., 2018; Olsson et al., 2016; Schmölzer, Lemmerer, & Nidetzky, 2018; Schwab, Fischer, & Wüst, 2015; Schwab, Fischer, Giri, & Wüst, 2015; Xiao, Muzashvili, & Georgiev, 2014).

Limitation on the applicability of plant GTs arises from the fact that these enzymes are difficult to express in standard microbial hosts (Desmet et al., 2012; Lim, 2005; Nidetzky et al., 2018). *Escherichia coli* is most often used. In general, expression is low (≤10 mg/L; e.g., Arend, Warzecha, Hefner, & Stöckigt, 2001; Schmölzer, Gutmann, Diricks, Desmet, & Nidetzky, 2016) and poor (≤1 mg/L) in various instances (e.g., Cai et al., 2017; Welner et al., 2017). With notable exceptions (Arend et al., 2001; Dewitte et al., 2016; Priebe, Daschner, Schwab, & Weuster-Botz, 2018; Schmieder et al., 2016), the enzyme production has received relatively little attention for systematic process development. The main bottlenecks on the production efficiency thus remain largely unknown. Besides specific requirements an individual GT may have, it seems probable that there are also important factors of a more general, if not universal relevance. Discovery of such factors and process optimization along the lines thus suggested would present important advances in the biocatalytic application of GTs.

In a recent study of a bacterial Leloir GT (sucrose synthase from *Acidithiobacillus caldus*; Diricks, De Bruyn, Van Daele, Walmagh, & Desmet, 2015) we showed that the constitutive expression in *E. coli* BL21 shifted the production of recombinant protein mainly to the stationary growth phase (Schmölzer, Lemmerer, Gutmann, & Nidetzky, 2017). Once the glucose carbon source had been depleted, the active enzyme was accumulated gradually to a substantial titer of ∼350 mg/L of culture. This result gave rise to the working hypothesis of this study, namely, the expression in growth-arrested *E. coli* might constitute a general strategy for efficiency-enhanced production of (plant) GTs.

Here, we used a synthetic biology-based approach to decouple *E. coli* BL21(DE3) cell growth from the target gene overexpression (Mairhofer, Striedner, Grabherr, & Wilde, 2016). The underlying concept is built upon the Gp2 protein from the bacteriophage T7. Gp2 inhibits the *E. coli* endogenous RNA polymerase (Mekler, Minakhin, Sheppard, Wigneshweraraj, & Severinov, 2011) whereas it leaves T7 RNA polymerase unaffected (Mairhofer, Striedner et al., 2016). L-Arabinose-inducible expression of the Gp2 gene, thus, allows for the host RNA polymerase to be shut off and, hence, the cell growth to be arrested in a controllable fashion. Under the conditions managed by Gp2, the *E. coli* protein synthesis machinery is taken over for recombinant production of the target protein(s). The practical design embodied in enGenes technology involves genome integration of the Gp2 coding gene under control of the *araB* promoter inducible by L-arabinose. Within this strain background, a pET plasmid vector is used that contains the gene(s) of interest inducible by isopropyl-β-D-galactoside. This design provides flexibility and temporal control for the cell proliferation to be switched off and protein production to be induced (Mairhofer, Striedner et al., 2016; Mairhofer, Stargardt et al., 2016). It presents a new approach toward quiescent *E. coli* cells applied to recombinant protein production (for alternative approaches, see: C. Chen, Walia, Mukherjee, Mahalik, & Summers, 2015; Ghosh, Gupta, & Mukherjee, 2012; Mahalik, Sharma, & Mukherjee, 2014).

Here, we demonstrate the application of the outlined approach to individual and combined production of two representative GTs, the sucrose synthase from soybean (*Glycine max*; *Gm*SuSy; Bungaruang, Gutmann, & Nidetzky, 2013) and the flavonoid GT UGT71A15 from apple (*Malus domestica*; Lepak, Gutmann, Kulmer, & Nidetzky, 2015). The GT reactions are shown in Scheme 1. Enzyme coexpression reflects the idea of a glycosylation cascade in which the sugar nucleotide donor substrate is formed in situ and continuously regenerated (Scheme 1; Nidetzky et al., 2018). We showed that the enzyme production in growth-arrested *E. coli* gives significant improvements in the amount and quality of the recombinant GT as compared with the exactly comparable production in the growing
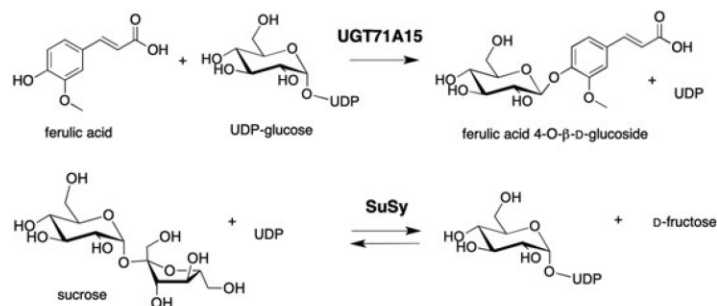
SCHEME 1 The reaction of the glycosyltransferase UGT71A15 as used in this study and reaction of sucrose synthase (SuSy). Coupling of the SuSy reaction to the glycosyltransferase reaction allows for a glycoside synthesis with in situ regeneration of the uridine diphosphate-glucose as donor substrate (for review, see Nidetzky et al., 2018; Schmölzer et al., 2016; for coupled use of UGT71A15 and SuSy, see Lepak et al., 2015)

E. coli reference. We also showed transferability of the production strategy to a high-cell-density fed-batch culture of *E. coli* at 20 L operating scale and obtained up to 830 mg GT protein/L of culture in that way.

2 | MATERIALS AND METHODS

2.1 | Microbial strains

The *E. coli* strain referred to as enGenes-X-press is a BL21(DE3) derivate with the genotype *E. coli* B F⁻*ompT gal dcm lon hsdS_B*(r_B⁻m_B⁻) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*⁺]_K-12(λS) attTn7:: <araC-p_araB-gp2> Δ*araABCD*::CAT^R (Mairhofer, Striedner et al., 2016). The enGenes-X-press has the Gp2 coding sequence integrated into the genome at the Tn7 attachment (attTn7) site. The Gp2 gene expression is controlled by the L-arabinose inducible *araB* promoter and involves transcription by T7 RNA polymerase. The inserted sequence additionally involves a regulator (*araC*) and a chloramphenicol resistance gene. Furthermore, the strain features knockout of the complete L-arabinose degrading operon *araABCD*. The strain enGenes-X-press is proprietary material of enGenes Biotech GmbH. Its construction was reported elsewhere (Mairhofer, Striedner et al., 2016). The reference *E. coli* strain BL21(DE3) was from the Coli Genetic Stock Center (CGSC#: 12504). To compare the performance of the two *E. coli* strains, both were transformed with the relevant expression plasmids, as described below. It was shown (Mairhofer, Striedner et al., 2016) that in the absence of induction by L-arabinose, enGenes-X-press exhibits a growth behavior not different from that of *E. coli* BL21(DE3) and that the presence of L-arabinose does not affect the growth of the host strains as such.

2.2 | Enzyme expression

The previously described pET-STRP3 expression vector was used (Lepak et al., 2015). Enzymes are thus produced as fusion proteins

harboring N-terminal Strep–Tag II. For single enzyme expression, the coding genes were inserted via *NdeI* and *XhoI* restriction sites. The GmSuSy gene (GenBank: AF030231) was codon optimized for expression in *E. coli* (Bungaruang et al., 2013). The UGT71A15 gene (GenBank: DQ103712) was used.

For enzyme coexpression, the vector referred to as pETduo was derived from pET-STRP3 by inserting the UGT71A15 gene into the plasmid harboring the GmSusy gene. Suitable oligonucleotide primers (pDUO_Ins_fwd, pDUO_Ins_rev; Supporting Information) were used to amplify the introduced expression cassette by specific binding upstream of the promoter region and downstream of the terminator sequence. The receiving vector was linearized by amplification using pETduo_BB_fwd and pETduo_BB_rev (Supporting Information) as forward and reverse primer, respectively. The F1 ORI was eliminated in this step (Figures S1 and S2). The resulting DNA fragments were purified on agarose gel and fused by homologous recombination. The sequences of both GT coding genes were verified in the final expression vector.

Expression strains were obtained by transforming the respective expression vector into electrocompetent cells of *the E. coli* BL21(DE3) reference strain and the strain enGenes-X-press.

2.3 | Batch bioreactor cultivations

With exceptions noted, the strains *E. coli* BL21(DE3) and enGenes-X-press were cultivated under exactly comparable conditions. Each strain harbored the pET expression vector for production of GmSuSy, UGT71A15, or both. Precultures (250 ml in 1,000 ml baffled-shaken flasks) were inoculated from the glycerol stocks (100 μl) and incubated overnight at 37°C and 130 rpm (Certomat® BS-1; Sartorius, Göttingen, Germany). LB medium supplemented with sterile filtered antibiotics (*E. coli* BL21(DE3): 50 μg/ml kanamycin; enGenes-X-press: 50 μg/ml kanamycin; 34 μg/ml chloramphenicol) was used.

Bioreactor cultivations were performed, parallelly, in two Labfors III 3.6 L bioreactors from Infors HT Multitron (Bottmingen, Switzerland). A semisynthetic medium (Table S1) prepared from separately

autoclaved components was used. To this, 0.4 ml/L of a sterile filtered trace element solution was added (Table S2). Note: Glycerol was used as the carbon source. Glucose was not used in the batch cultivations because of the well-known effect of carbon catabolite repression by glucose on the induction of gene expression under control of the *araB* promoter and induced by L-arabinose (Brückner & Titgemeyer, 2002; Cagnon, Valverde, & Masson, 1991; Lee & Jung, 2007). A value of 40% air saturation was maintained using agitation and airflow cascade. Polypropylene glycol (PPG; 10%) was used for foam control. The pH was maintained at 7.0 using automated addition of 2 M KOH and 1 M H_3PO_4. Bioreactors were inoculated to an optical density $(OD)_{600nm}$ of 0.5 at a temperature of 37°C. The temperature was lowered as per Table 1 before induction and initiated at an OD_{600nm} of 5. The *E. coli* BL21(DE3) strain was induced with filter-sterilized isopropyl β-D-1-thiogalactopyranoside (IPTG) solution to a final concentration of 0.1 mM. The enGenes-X-press strain was induced with IPTG (0.1 mM) and L-arabinose (100 mM). Samples (10 ml) were taken at certain times and analyzed for cell dry mass and glycerol.

Cell dry mass was determined from 10 ml samples. The cells were centrifuged, washed with 20 ml distilled water, resuspended, and transferred to a dried and weighed beaker, which was then dried at 105°C for 24 hr and reweighed. Cells were disrupted by ultra-sonification (Supporting Information) to obtain the cell extract for measurement and purification.

2.4 | Fed-batch bioreactor cultivations at a larger scale

The strains *E. coli* BL21(DE3) and enGenes-X-press harboring the pET expression vector for production of GmSuSy were used. Fed-batch cultivations were performed in a Bioengineering AG bioreactor (Type NLF22; Wald, Switzerland) with 20 L total working volume (10 L batch volume). The bioreactor was equipped with standard computer-controlled units (Siemens Simatic S7; WinCC; Siemens AG, Munich, Germany). Semisynthetic medium was used (see the Supporting Information). Its composition was adjusted for a total production of 1,580 g dry cell mass. Note: As the cultivation was

done under carbon source-limited growth conditions, glucose could be used. This was not feasible in the batch cultivations.

Precultures (250 ml in 2,000 ml baffled-shaken flasks) were inoculated from cell bank vials (500 μl) and incubated at 37°C and 180 rpm (Infors HT Multitron) until an OD_{600nm} of 3.0–3.5 was reached. An amount of 1,000 OD_{600nm} units was transferred into 400 ml of 0.9% (by weight) NaCl solution and added to the bioreactor. Air saturation was kept at 30% through stirrer speed and aeration rate control. The O_2 and CO_2 contents in the outlet air were measured with a BlueSens (Herten, Germany) BlueInOne Gas Analyzer. The pH was controlled at 7.0 (±0.05) using 25% (by weight) NH_4OH solution. During the batch phase, the temperature was 37°C ± 0.5°C. PPG-2000 was added as antifoam (1.5 ml/L batch culture). Before starting the feeding, when the culture had reached stationary phase (after ∼10 hr; ∼8 g dry cell mass/L), the temperature was lowered to 30°C. Initially, an exponential substrate feed was used to provide a constant specific growth rate of 0.17 hr^{-1} over 11 hr. Then, a linear substrate feed (9.54 g glucose/min) over 4 hr and another linear substrate feed (4.35 g glucose/min) over 15 hr were applied (see the Supporting Information). Substrate feed involved superimposed feedback control of weight loss in the substrate tank. Protein expression was induced after 15 hr of substrate feed. It included the addition of 31.6 mmol IPTG (=20 μmol/g cell dry mass) and 0.1 M L-arabinose based on the final volume of ∼20 L. The induction solution contained 79 mM of $(NH_4)_2SO_4$ based on the end volume. PPG-2000 (500 ml in total) was used for foam control. Samples were taken at certain times and the cell growth was recorded as an increase in OD_{600nm}. The cells were harvested after 40 hr by centrifugation (30 min at 4°C and 4,420 g; Sorvall RC-5B; ThermoFisher Scientific, Waltham, MA) and stored at −20°C. They were disrupted by high-pressure homogenization (Supporting Information) to obtain cell extract for measurement and purification.

2.5 | Protein purification and characterization

2.5.1 | Strep–Tag purification

Prepacked 1 or 5 ml Strep–Tactin Sepharose columns (IBA Life Sciences, Göttingen, Germany) was used on an ÄKTA Explorer 100

TABLE 1 Glycosyltransferase activities from batch bioreactor cultivations of enGenes-X-press and the *Escherichia coli* BL21(DE3) reference

Glycosyltransferase	Construct	Induction temperature (°C)	Reference (U/L)[a]	enGenes-X-press (U/L)[a]	Reference (U/g)[a]	enGenes-X-press (U/g)[a]	Fold increase[c]
GmSuSy	single	30	285	700	25	115	4.6 (2.4)
UGT71A15	single	30	0.3	1.4	0.0	0.2	8.9 (4.7)
GmSuSy	single	25	8.7	16.7	1.1	4.2	3.9 (1.9)
UGT71A15	single	25	3.9	10.6	0.5	2.3	4.4 (2.7)
GmSuSy[b]	double	25	3.5	3.7	0.5	0.7	1.6 (1.1)
UGT71A15[b]	double	25	1.8	3.3	0.2	0.6	2.6 (1.8)

Note. GmSuSy: sucrose synthase from soybean (*Glycine max*); UGT71A15: UDP-glycosyltransferase 71A15.
[a]Recorded at the end of the bioreactor cultivation; results are from biological duplicates or triplicates and agree within less than 10% relative *SD*.
[b]Activities were measured individually.
[c]Refers to U/g data and (in brackets) U/L data.

system (GE Healthcare, Pasching, Austria) and operated at a flow rate of 1 and 5 ml/min, respectively. Tris/HCl buffer (100 mM, pH 8.0, 150 mM NaCl, 1 mM EDTA) was used. Alternatively, a column (0.5 cm inner diameter; 10 cm length) of Streptactin Sepharose High Performance (GE Healthcare) was used. The flow rate was 1 ml/min. Cell extract filtered through 1.2-μm cellulose–acetate syringe filter was loaded onto the column. After loading, the column was flushed with buffer until the UV signal (280 nm) signal reached the baseline level. Elution was done with two column volumes of desthiobiotin (2.5 mM). The columns were regenerated with hydroxy-azo phenyl-benzoic acid (1 mM) and equilibrated again with buffer.

2.5.2 | Size-exclusion chromatography

This was performed using a Superdex 200 increase 10/300 GL column (GE Healthcare) operated on ÄKTA Explorer 100 system. All runs were performed in 10 mM phosphate buffer (pH 7.4) containing 300 mM NaCl. The flow rate was 0.5 ml/min and UV detection (280 nm) was used. Strep–Tactin eluate (200 μl) was applied to the column. For analytical SEC, a G3000SWXL column (300 × 7.8 mm, inner diameter 5 μm; Tosoh Bioscience, Tokyo, Japan) was used.

2.6 | Enzymatic activity measurements

Assays were performed at 30°C in 1.5 ml tubes on an Eppendorf (Vienna, Austria) Thermomixer comfort at 300 rpm. The enzyme was added as E. coli cell extract or in purified form to start the reaction. Product release was measured in samples (20 μl) taken at four different times. Samples were diluted into acetonitrile (180 μl) to stop the reaction. Precipitated protein was removed by centrifugation for 20 min at room temperature and 13,200 rpm. About 5–10 μl of supernatant was used for further analysis. Enzymatic rates were determined from the linear time courses of product formation. One Unit (U) of enzyme activity is the amount of enzyme producing 1 μmol product/min under the assay conditions.

2.6.1 | Sucrose synthase from soybean (*Glycine max*)

Sucrose (500 mM) and uridine 5′-diphosphate (10 mM) were used as substrates in 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer (100 mM; pH 7.5) containing 13 mM MgCl$_2$ and 0.13% bovine serum albumin (BSA; Bungaruang et al., 2013). The released uridine diphosphate (UDP)-glucose was measured by high-performance liquid chromatography (HPLC).

2.6.2 | UDP-glycosyltransferase 71A15

Ferulic acid (1 mM) and UDP-glucose (2 mM) were used as substrates in the HEPES buffer (50 mM; pH 7.5) containing 50 mM KCl, 13 mM MgCl$_2$, 0.13% BSA, and 2% dimethyl sulfoxide. The released ferulic acid-4-O-β-D-glucoside (Scheme 1) was measured by HPLC.

2.7 | Analytical HPLC methods

2.7.1 | Glycerol

This was analyzed on a Merck–Hitachi LaChrome HPLC System equipped with an Aminex HPX-87H column (BioRad, Richmond, CA), a Merck–Hitachi LaChrome L-7250 autosampler, and a Merck L-7490 RI detector. The system was operated at 65°C, using a flow rate of 0.6 ml/min with 5 mM sulfuric acid as the eluent.

2.7.2 | Reaction of GmSuSy

UDP-glucose and UDP were analyzed by reversed phase HPLC using a Kinetex® C18 column (5 μm, 100 Å, 50 × 4.6 mm; Phenomenex, Torrance, CA) in reversed phase ion-pairing mode. The analysis was performed at 35°C with a mobile phase of 87.5% 20 mM potassium phosphate buffer (pH 5.9) containing 40 mM tetra-n-butylammonium bromide and 12.5% acetonitrile. An isocratic flow rate of 2 ml/min was used and the detection was at 262 nm.

2.7.3 | Reaction of UGT71A15

Ferulic acid and ferulic acid-4-O-β-glucoside were analyzed by reversed phase HPLC using the above described Kinetex® C18 column in the reversed phase mode. The analysis was performed at 35°C using a gradient separation (mobile phase A: H$_2$O + 0.1% formic acid and mobile phase B: acetonitrile + 0.1% formic acid) at a flow rate of 1 ml/min. Gradient conditions were as follows: 0.00 min 10% B; 0.50 min 10% B; 4.00 min 70% B; 4.30 min 70% B; 4.31 min 10% B; 7.00 min 10% B. Detection was at 320 nm. Authentic standards were used for calibration.

3 | RESULTS AND DISCUSSION

The enzymes used are representative of plant GTs applied to biocatalytic glycosylation of small molecules (Nidetzky et al., 2018). UGT71A15 shows broad acceptor substrate specificity and has previously been used for the glycosylation of flavonoids (Lepak et al., 2015). It is a member of the GT family GT-1, a large enzyme family comprising numerous plant GTs involved in small-molecule glycosylation (Liang et al., 2015). UGT71A15 is a 50.2 kDa protein that functions as a monomer. GmSuSy belongs to the GT family GT-4. It is a functional homotetramer composed of 92.2 kDa subunits. GmSuSy has previously been used for UDP-glucose recycling in glycosylation reactions by coupled GTs, including UGT71A15 (Scheme 1; Lepak et al., 2015; Schmölzer et al., 2016; Schmölzer et al., 2018).

3.1 | Single gene expression for production of GmSuSy

We performed controlled batch bioreactor cultivations at 25 and 30°C to compare enzyme production in enGenes-X-press under

growth arrest to enzyme production in the normal growing *E. coli* BL21(DE3). The different temperatures were chosen to examine their effect on the enzyme production and cell growth. The results are shown in Figure 1 (panels A and B; 30°C) and in Figure S3 (25°C). At 30°C, enGenes-X-press and *E. coli* BL21(DE3) showed similar growth until the time of induction (8 hr). Whereas *E. coli* BL21(DE3) continued exponential growth afterward (Figure 1b), the growth of enGenes-X-press was reduced to a very small amount, hence, effectively switched off at this point (Figure 1a). Its growth arrest notwithstanding, the strain enGenes-X-press continued consumption of the glycerol carbon source similarly as the growing *E. coli* BL21(DE3) did (Figure 1a,b). The GmSusy production, measured as unit enzyme activity/g dry cell mass was far superior in enGenes-X-press as compared with *E. coli* BL21(DE3). Both strains started from a similar specific activity of 20 U/g at the time of induction. Whereas in *E. coli* BL21(DE3) the specific activity was constant in the induction phase, it increased almost linearly with time in enGenes-X-press and reached 115 U/g after 24 hr. This represents an approximately fivefold enhancement of specific activity compared with the reference. Cultivations at 25°C (Figure S3a,b) gave only poor enzyme production, with an activity of just ~4 U/g in enGenes-X-press and ~1 U/g in *E. coli* BL21(DE3). The strain growth was still switched off reliably in enGenes-X-press through induction but glycerol consumption was slow compared with *E. coli* BL21(DE3) in the postinduction phase. However, 25°C was not an option for GmSusy production, we showed later that it was one for production of UGT71A15.

In Table 1 we summarize the production parameters of the different bioreactor cultivations performed. In terms of enzyme activity/culture volume, the production in enGenes-X-press was effective in improving (~2.5-fold) the reference production. However, decreased biomass formation by enGenes-X-press as compared with *E. coli* BL21(DE3) mitigated the effect of growth-arrested production on the cell mass-based specific enzyme activity. We showed later that enzyme production in high-cell density fed-batch culture could overcome this problem. Using the specific activity of the purified GmSuSy of 4.5 U/mg protein, comparable to the earlier studies (Bungaruang et al., 2013), we calculated that production in enGenes-X-press at 30°C gave an overall enzyme yield of 156 mg/L culture. The corresponding production in *E. coli* BL21(DE3) gave 63 mg/L. These values are used together with measurements of total protein content/g cell mass and activity/g cell mass to estimate that the recombinant GmSuSy accounted for roughly ~5% and ~1% of the total intracellular protein in enGenes-X-press and *E. coli* BL21(DE3), respectively.

3.2 | Single gene expression for the production of UGT71A15

UGT71A15 was previously noted to be difficult to produce. In a shake-flask culture of *E. coli* BL21(DE3) about 1–5 mg protein/L was
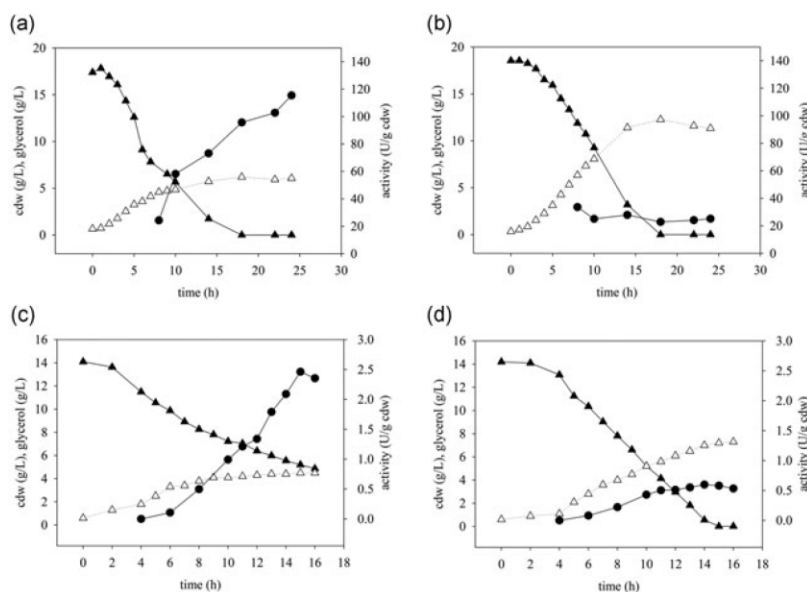
FIGURE 1 Time courses of growth, glycerol consumption, and enzyme formation in batch bioreactor cultivations of enGenes-X-press (a,c) and *Escherichia coli* BL21(DE3) (b,d) producing GmSuSy (a,b) and the glycosyltransferase UGT71A15 (c,d). The induction temperature was 30°C for GmSuSy production (a,b) and 25°C for UGT71A15 production (c,d). The symbols show cell dry mass concentration, open triangles; glycerol concentration, full triangles; volumetric enzyme activity, full circles. GmSuSy: sucrose synthase from soybean (*Glycine max*); UGT71A15: UDP-glycosyltransferase 71A15

obtained (Lepak et al., 2015). The same series of bioreactor cultivations described above for GmSuSy were carried out with UGT71A15. The results are shown in Figure 1 (panels C and D; 25°C) and Figure S4 (30°C). Overall trends in biomass growth and glycerol consumption were similar as noted before. Production at 25°C was strongly preferred overproduction at 30°C, probably because UGT71A15 was not stable at the higher temperature. As observed for GmSuSy, the UGT71A15 was produced poorly in the growing *E. coli* BL21(DE3) (Figure 1d). There was only a small increase in specific enzyme activity in the phase after the induction. By contrast, UGT71A15 was formed efficiently in enGenes-X-press after the induction (Figure 1c). We noted that the arrest of growth of enGenes-X-press was less clear-cut at 25°C (Figure 1c) than it was at 30°C (Figure 1a). A marked slow-down of growth was, however, observed after the induction (Figure 1c). Nonetheless, the effect of growth reduction on enzyme production was pronounced, as evident from comparing panels C and D in Figure 1. Approximately, a fourfold increase in the specific activity was thus achieved. The activity of 2.3 U/g cell mass obtained for UGT71A15 was 50-fold lower than that obtained for GmSuSy, reflecting roughly the differences in specific activity of the two enzymes as purified proteins. The parameters of the enzyme production are listed in Table 1. The volumetric enzyme titer was 11 U/L culture when using enGenes-X-press, a 2.5-fold improvement was observed when compared with the reference strain. The specific activity of purified UGT71A15 is 0.5 U/mg. Translated into functionally expressed recombinant protein the production yield was 21 mg/L of enGenes-X-press culture, which can be compared with 8 mg/L of *E. coli* BL21(DE3) culture. In terms of abundance relative to total intracellular protein, values of ~1.0 and ~0.2% are calculated for enGenes-X-press and *E. coli* BL21(DE3), respectively.

3.3 | Gene coexpression for production of UGT71A15 and GmSuSy

Considering the temperature requirements of UGT71A15 revealed in the single gene expression studies, coexpression of the GmSuSy and UGT71A15 genes was performed at 25°C. Time courses from bioreactor cultivations of enGenes-X-press and *E. coli* BL21(DE3) are shown in Figure S5. The profiles of growth and glycerol consumption of both strains producing the two GTs (Figure S5) were highly similar as compared to their production of only UGT71A15 (Figure 1c,d). Upon induction, enGenes-X-press showed slow accumulation of both enzyme activities. The final specific activity (Table 1) was substantially lower (UGT71A15: fivefold; GmSuSy: eightfold) than the single gene expression cultures. Interestingly, the production of GmSuSy next to UGT71A15 was hardly detectable in *E. coli* BL21(DE3). Gene coexpression interfered with functional production of the individual enzymes and it did so in both enGenes-X-press and *E. coli* BL21(DE3). The volumetric titer of GmSuSy was decreased from 3.7 to 0.8 mg/L for enGenes-X-press and from 1.9 to 0.8 mg/L in the reference. For UGT71A15, titers dropped from 21.2 to 6.6 mg/L in enGenes-X-press and from 7.8 to

3.7 mg/L in the reference. Therefore, at this stage, enzyme production through single gene expression would be preferable. Irrespective of the possible incompatibility of GmSuSy and UGT71A15 for combined production in a single host, production under growth arrest in enGenes-X-press proved superior to production in growing *E. coli* BL21(DE3) under all conditions used (Table 1).

3.4 | Downstream processing and characterization of GmSuSy

Protein quality is key in recombinant protein production. Although activity is the quintessential parameter of functional expression of the enzyme, it alone is not sufficient to inform about the overall quality of the recombinant GTs production. We, therefore, isolated with Strep–Tactin affinity purification the GmSuSy produced in enGenes-X-press and *E. coli* BL21(DE3). The cell mass (~1 g dry matter each) from bioreactor cultures that run for 24 hr at 30°C (Figure 1a,b) was used. Target protein was recovered in a single sharp protein peak (Figure S6). The protein yield after purification was 26.7 mg/g for enGenes-X-press and 5.6 mg/g for *E. coli* BL21(DE3). The specific activity of isolated GmSuSy (enGenes-X-press: 5.1 U/mg; *E. coli* BL21(DE3): 4.8 U/mg) was in line with the previous report for GmSuSy produced in a shaken-flask culture of *E. coli* BL21(DE3; Bungaruang et al., 2013). Analysis with sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE; Figure S6) showed a low abundance of GmSuSy in the soluble fraction from cell disruption. On the basis of the semiquantitative densitometry, ~10% of GmSuSy was present in the soluble fraction in enGenes-X-press whereas it was only ~2% in the reference.

3.5 | Fed-batch cultivation for the production of GmSuSy at a larger scale

It was mentioned that if volumetric enzyme titer is the parameter used for evaluation, approximately, twofold lower biomass yield in enGenes-X-press cultivations as compared with *E. coli* BL21(DE3) cultivations reduces the benefit of enzyme production under conditions of arrested growth (cf. Table 1). We addressed this problem for the case of GmSuSy concomitantly with assessing the translatability of the enzyme production in enGenes-X-press to a high-cell-density fed-batch cultivation at one magnitude order larger scale (20 L). The important question was, whether under these conditions the advantage of approximately fivefold higher enzyme activity/g cell mass would be retained without compromising the total biomass formation. We recall that glucose was the carbon source used in the fed-batch cultivation. Glycerol was used in the batch cultivation to avoid the effect of carbon catabolite repression by glucose (Brückner & Titgemeyer, 2002; Cagnon et al., 1991; Lee & Jung, 2007).

The fed-batch protocol (Figure S7) gave a final biomass concentration of 56.0 g dry mass/L for enGenes-X-press and 67.5 g dry mass/L for *E. coli* BL21(DE3), a difference of just ~20% (Table 2). On the basis of the protein purified from cell material (~1 g dry

TABLE 2 Scaled-up production of GmSuSy in fed-batch bioreactor cultivation and recovery of the enzyme

E. colistrain	Biomass yield (g dry cells/L)[a]	Protein content(mg GmSuSy/g dry cells)	Protein recovery (mAU × ml)[b]	SEC active fraction/total active enzyme(%)[c] / -fold[d]
enGenes-X-press	56	14.8	673[e] (80)[f]	61/12.5
BL21(DE3) reference	68	5.0	226[e] (15)[f]	26/1

Note. GmSuSy: sucrose synthase from soybean (*Glycine max*); SEC: size-exclusion chromatography
[a]From the end of the bioreactor cultivation. The corresponding volumetric yields of GmSuSy are 0.83 g/L using enGenes-X-press and 0.34 g/L using the BL21(DE3) reference
[b]Relevant peak area (absorbance detection at 280 nm) in chromatography times the volume collected.
[c]Percentage of enzymatically active protein in the eluate from the analytical SEC.
[d]Enhanced production of the total active enzyme in enGenes-X-press as compared with the reference; that is: 80×0.61/(15×0.26), from the table.
[e]Strep–Tactin eluate, and
[f]eluate from the preparative SEC.

matter) thus received, we obtained 14.8 mg/g for enGenes-X-press and 5.0 mg/g for *E. coli* BL21(DE3), also qualitatively in line with SDS-PAGE (Figure S8). These numbers allow one to calculate a volumetric titer of GmSuSy, which was 830 mg/L for the enGenes-X-press culture and 337 mg/L for the *E. coli* BL21(DE3) culture. The recombinant GmSuSy thus accounted for 1.5% of the total intracellular protein in enGenes-X-press and only 0.5% in the reference. The specific activity of the GmSuSy isolated from fed-batch production was in the range 0.5–2.1 U/mg, between three and 10-fold lower than the specific activities of enzyme isolated from the batch bioreactor cultures.

We, therefore, subjected the enzyme obtained from Strep–Tactin purification to an additional step of size-exclusion chromatography (SEC) on a Sephadex 200 column, as shown in Figure 2. Samples from both productions showed a protein peak corresponding in molecular size (350–400 kDa) to the functional GmSuSy tetramer. However, there was also the material of a larger size that represented protein agglomerates of an undefined degree. The two samples analyzed differed considerably regarding the relative content of such agglomerates. Whereas in the sample from the *E. coli* BL21(DE3) culture, the agglomerated protein exceeded in abundance by far the protein in its native size. The

sample from the enGenes-X-press culture contained the GmSuSy tetramer as its main constituent.

As the protein peaks in the SEC of the *E. coli* BL21(DE3) sample were baseline separated, they were collected as fraction I and II and were further analyzed by analytical SEC, as shown in Figure S9. The fraction I did not elute from the column, probably because the agglomerates had already reached size limit for the SEC matrix. The analytical SEC of fraction II showed the expected tetrameric GmSuSy. We also determined the specific activity of the protein in each fraction and found none in fraction I and 4.7 U/mg in fraction II. By integrating the absorbance traces from the preparative SEC (Figure 2) we calculated, after normalization, the portion of native tetramer in the enGenes-X-press sample and found it to be 65% whereas that in the *E. coli* BL21(DE3) it was only 25% (Table 2).

Considering both the specific protein production/g cell mass and the quality of the protein thus made, we found that GmSuSy production in growth-arrested enGenes-X-press outperforms the reference production in growing *E. coli* BL21(DE3) by more than one magnitude order (12.5-fold; Table 2). In terms of improvement from batch to fed-batch culture, GmSuSy activity was increased from 285 to 390 U/L for *E. coli* BL21(DE3). In enGenes-X-press, however, the increase was far more significant (3.3-fold), 700 to 2,300 U/L, in
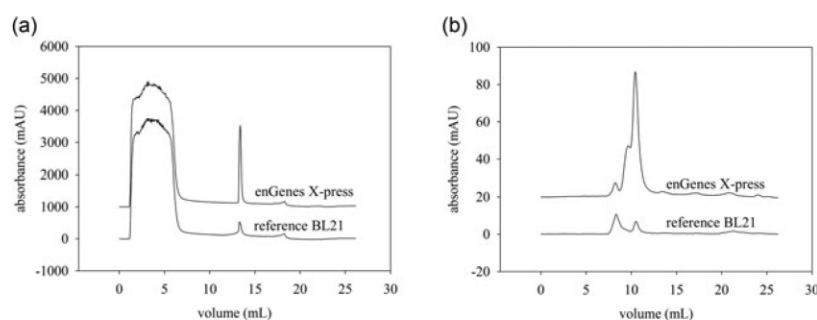
FIGURE 2 Protein quality analysis in GmSuSy produced in enGenes-X-press and *Escherichia coli* BL1(DE3) by high-cell density fed-batch cultivation. Absorbance traces of protein elution from Strep–Tactin affinity chromatography (a) and subsequent preparative SEC (b). The SEC trace reveals heterogeneity in both enzyme preparations, however, much less so in the preparation from production in enGenes-X-press. The peak at around 11 ml elution volume corresponds to the size expected for the native GmSuSy tetramer. GmSuSy: sucrose synthase from soybean (*Glycine max*); SEC: size-exclusion chromatography; UGT71A15: UDP-glycosyltransferase 71A15

consequence of the enhanced biomass formation in the fed-batch as compared with the batch culture.

4 | CONCLUSION

To advance biocatalysis for small-molecule glycosylation with plant GTs, better process technologies for the recombinant production of these enzymes are required (e.g., Priebe et al., 2018; Schmideder et al., 2016). The problem is difficult due to the involvement of multiple process variables, and the complex interrelationship these variables have with each other. A modularized approach interconnecting molecular engineering strategies at the levels of the gene, the protein, and the host organism is promising to make the development more predictable and faster. Focusing on *E. coli* as the production host, we showed here that gene expression under arrested cell growth was highly effective for GT synthesis as it enhanced yield and improved quality. Compared with the growing *E. coli* reference, the overall boost of production of the functional, high-quality enzyme in enGenes-X-press exceeded one order of magnitude. The total amount of functional enzyme (GmSuSy) produced in high-cell density fed-batch culture at 20 L scale was 830 mg/L, surpassing common titers of plant GTs in *E. coli* productions by 100-fold or more. The outstanding titer of 5.5 g/L for the glucosyltransferase from *Vitis vinifera* expressed in *E. coli* BL21(DE3) is clearly noted at this point (Priebe et al., 2018). With its effectiveness and scalability demonstrated in principle, the approach of enzyme production decoupled from cell growth might find further uses with different GTs and potentially other difficult-to-express proteins (see also: C. Chen et al., 2015; Gosh et al., 2012; Mahalik et al., 2014). For such proteins, HIV-1 protease can serve as a representative example (Mairhofer, Striedner et al., 2016). Practical realization of the approach in the strain enGenes-X-press thus represents a validated platform technology that offers flexible interconnection with gene and protein design strategies for enhanced recombinant protein production.

CONFLICT OF INTEREST

J. M. is the CEO of enGenes Biotech GmbH and has an interest in the commercial exploitation of enGene-X-press technology for recombinant protein production.

ORCID

Juergen Mairhofer http://orcid.org/0000-0002-2686-1971
Rainer Hahn http://orcid.org/0000-0001-5654-5032
Bernd Nidetzky http://orcid.org/0000-0002-5030-2643

REFERENCES

Arend, J., Warzecha, H., Hefner, T., & Stöckigt, J. (2001). Utilizing genetically engineered bacteria to produce plant-specific glucosides. *Biotechnology and Bioengineering*, *76*, 126–131. https://doi.org/10.1002/bit.1152.

Bowles., D., Lim, E.-K., Poppenberger, B., & Vaistij, F. E. (2006). Glycosyltransferases of lipophilic molecules. *Annual Review of Plant Biology*, *57*, 567–597. https://doi.org/10.1146/annurev.arplant.57.032905.105429.

Brückner, R., & Titgemeyer, F. (2002). Carbon catabolite repression in bacteria: Choice of the carbon source and autoregulatory limitation of sugar utilization. *FEMS Microbiology Letters*, *209*, 141–148. https://doi.org/10.1111/j.1574-6968.2002.tb11123.x.

Bungaruang, L., Gutmann, A., & Nidetzky, B. (2013). Leloir glycosyltransferases and natural product glycosylation: Biocatalytic synthesis of the C-glucoside nothofagin, a major antioxidant of redbush herbal tea. *Advanced Synthesis and Catalysis*, *355*, 2757–2763. https://doi.org/10.1002/adsc.201300251.

Cagnon, C., Valverde, V., & Masson, J.-M. (1991). A new family of sugar-inducible expression vectors for *Escherichia coli*. *Protein Engineering*, *4*, 843–847. https://doi.org/10.1093/protein/4.7.843.

Cai, R., Chen, C., Li, Y., Sun, K., Zhou, F., Chen, K., & Jia, H. (2017). Improved soluble bacterial expression and properties of the recombinant flavonoid glucosyltransferase UGT73G1 from *Allium cepa*. *Journal of Biotechnology*, *255*, 9–15. https://doi.org/10.1016/j.jbiotec.2017.06.011.

Chen, C.-C., Walia, R., Mukherjee, K. J., Mahalik, S., & Summers, D. K. (2015). Indole generates quiescent and metabolically active *Escherichia coli* cultures. *Biotechnology Journal*, *10*, 636–646. https://doi.org/10.1002/biot.201400381.

Chen, R. (2018). Enzyme and microbial technology for synthesis of bioactive oligosaccharides: An update. *Applied Microbiology and Biotechnology*, *102*, 3017–3026. https://doi.org/10.1007/s00253-018-8839-2.

De Bruyn, F., Maertens, J., Beauprez, J., Soetaert, W., & De Mey, M. (2015). Biotechnological advances in UDP-sugar based glycosylation of small molecules. *Biotechnology Advances*, *33*, 288–302. https://doi.org/10.1016/j.biotechadv.2015.02.005.

Desmet, T., Soetaert, W., Bojarová, P., Křen, V., Dijkhuizen, L., Eastwick-Field, V., & Schiller, A. (2012). Enzymatic glycosylation of small molecules: Challenging substrates require tailored catalysts. *Chemistry*, *18*, 10786–10801. https://doi.org/10.1002/chem.201103069.

Dewitte, G., Walmagh, M., Diricks, M., Lepak, A., Gutmann, A., Nidetzky, B., & Desmet, T. (2016). Screening of recombinant glycosyltransferases reveals the broad acceptor specificity of stevia UGT-76G1. *Journal of Biotechnology*, *233*, 49–55. https://doi.org/10.1016/j.jbiotec.2016.06.034.

Diricks, M., De Bruyn, F., Van Daele, P., Walmagh, M., & Desmet, T. (2015). Identification of sucrose synthase in non-photosynthetic bacteria and characterization of the recombinant enzymes. *Applied Microbiology and Biotechnology*, *99*, 8465–8474. https://doi.org/10.1007/s00253-015-6548-7.

Ghosh, C., Gupta, R., & Mukherjee, K. (2012). An inverse metabolic engineering approach for the design of an improved host platform for over-expression of recombinant proteins in *Escherichia coli*. *Microbial Cell Factories*, *11*, 93. https://doi.org/10.1186/1475-2859-11-93.

Hofer, B. (2016). Recent developments in the enzymatic O-glycosylation of flavonoids. *Applied Microbiology and Biotechnology*, 100, 4269–4281. https://doi.org/10.1007/s00253-016-7465-0.

Hsu, T. M., Welner, D. H., Russ, Z. N., Cervantes, B., Prathuri, R. L., Adams, P. D., & Dueber, J. E. (2018). Employing a biochemical protecting group for a sustainable indigo dyeing strategy. *Nature Chemical Biology*, 14, 256–261. https://doi.org/10.1038/NCHEMBIO.2552.

Kim, B. G., Yang, S. M., Kim, S. Y., Cha, M. N., & Ahn, J. H. (2015). Biosynthesis and production of glycosylated flavonoids in *Escherichia coli*: Current state and perspectives. *Applied Microbiology and Biotechnology*, 99, 2979–2988. https://doi.org/10.1007/s00253-015-6504-6.

Lee, Y.-J., & Jung, K.-H. (2007). Modulation of the tendency towards inclusion body formation of recombinant protein by the addition of glucose in the araBAD promoter system of *Escherichia coli*. *Journal of Microbiology and Biotechnology*, 17, 1898–1903. doi:not available/not found.

Lepak, A., Gutmann, A., Kulmer, S. T., & Nidetzky, B. (2015). Creating a water-soluble resveratrol-based antioxidant by site-selective enzymatic glucosylation. *ChemBioChem*, 16, 1870–1874. https://doi.org/10.1002/cbic.201500284.

Liang, D.-M., Liu, J.-H., Wu, H., Wang, B.-B., Zhu, H.-J., & Qiao, J.-J. (2015). Glycosyltransferases: Mechanisms and applications in natural product development. *Chemical Society Reviews*, 44, 8350–8374. https://doi.org/10.1039/c5cs00600g.

Lim, E.-K. (2005). Plant glycosyltransferases: Their potential as novel biocatalysts. *Chemistry. An European Journal*, 11, 5486–5494. https://doi.org/10.1002/chem.200500115.

Mahalik, S., Sharma, A. K., & Mukherjee, K. J. (2014). Genome engineering for improved recombinant protein expression in *Escherichia coli*. *Microbial Cell Factories*, 13, 177. https://doi.org/10.1186/s12934-014-0177-1.

Mairhofer, J, Striedner, G, Grabherr, R, & Wilde, M (2016). Patent No. WO2016/174195A1. *Uncoupling growth and protein production*.

Mairhofer, J., Stargardt, P., Feuchtenhofer, L., Pontiller, J., Cserjan-Puschmann, M., Grabherr, R., & Striedner, G. (2016). Innovation without growth: Non-growth associated recombinant protein production in *Escherichia coli*. *New Biotechnology*, 33, S26.

Mekler, V., Minakhin, L., Sheppard, C., Wigneshwaraj, S., & Severinov, K. (2011). Molecular mechanism of transcription inhibition by phage T7 gp2 protein. *Journal of Molecular Biology*, 413, 1016–1027. https://doi.org/10.1016/j.jmb.2011.09.029.

Nidetzky, B., Gutmann, A., & Zhong, C. (2018). Leloir glycosyltransferases as biocatalysts for chemical production. *ACS Catalysis*, 8, 6283–6300. https://doi.org/10.1021/acscatal.8b00710.

Olsson, K., Carlsen, S., Semmler, A., Simón, E., Mikkelsen, M. D., & Møller, B. L. (2016). Microbial production of next-generation stevia sweeteners. *Microbial Cell Factories*, 15, 207. https://doi.org/10.1186/s12934-016-0609-1.

Priebe, X., Daschner, M., Schwab, W., & Weuster-Botz, D. (2018). Rational selection of biphasic reaction systems for geranyl glucoside production by *Escherichia coli* whole-cell biocatalysts. *Enzyme and Microbial Technology*, 112, 79–87. https://doi.org/10.1016/j.enzmictec.2017.11.003.

Schmieder, A., Priebe, X., Rubenbauer, M., Hoffmann, T., Huang, F.-C., Schwab, W., & Weuster-Botz, D. (2016). Non-water miscible ionic liquid improves biocatalytic production of geranyl glucoside with *Escherichia coli* overexpressing a glucosyltransferase. *Bioprocess and Biosystems Engineering*, 39, 1409–1414. https://doi.org/10.1007/s00449-016-1617-6.

Schmölzer, K., Lemmerer, M., & Nidetzky, B. (2018). Glycosyltransferase cascades made fit for chemical production: Integrated biocatalytic process for the natural polyphenol C-glucoside nothofagin. *Biotechnology and Bioengineering*, 115, 545–556. https://doi.org/10.1002/bit.26491.

Schmölzer, K., Lemmerer, M., Gutmann, A., & Nidetzky, B. (2017). Integrated process design for biocatalytic synthesis by a Leloir glycosyltransferase: UDP-glucose production with sucrose synthase. *Biotechnology and Bioengineering*, 114, 924–928. https://doi.org/10.1002/bit.26204.

Schmölzer, K., Gutmann, A., Diricks, M., Desmet, T., & Nidetzky, B. (2016). Sucrose synthase: A unique glycosyltransferase for biocatalytic glycosylation process development. *Biotechnology Advances*, 34, 88–111. https://doi.org/10.1016/j.biotechadv.2015.11.003.

Schwab, W., Fischer, T., & Wüst, M. (2015). Terpene glucoside production: Improved biocatalytic processes using glycosyltransferases. *Engineering in Life Sciences*, 15, 376–386. https://doi.org/10.1002/elsc.201400156.

Schwab, W., Fischer, T. C., Giri, A., & Wüst, M. (2015). Potential applications of glucosyltransferases in terpene glucoside production: Impacts on the use of aroma and fragrance. *Applied Microbiology and Biotechnology*, 99, 165–174. https://doi.org/10.1007/s00253-014-6229-y.

Thuan, N. H., & Sohng, J. K. (2013). Recent biotechnological progress in enzymatic synthesis of glycosides. *Journal of Industrial Microbiology and Biotechnology*, 40, 1329–1356. https://doi.org/10.1007/s10295-013-1332-0.

Welner, D. H., Shin, D., Tomaleri, G. P., DeGiovanni, A. M., Tsai, A. Y.-L., Tran, H. M., …, Adams, P. D. (2017). Plant cell wall glycosyltransferases: High-throughput recombinant expression screening and general requirements for these challenging enzymes. *PLOS One*, 12, e0177591. https://doi.org/10.1371/journal.pone.0177591.

Xiao, J., Muzashvili, T. S., & Georgiev, M. I. (2014). Advances in the biotechnological glycosylation of valuable flavonoids. *Biotechnology Advances*, 32, 1145–1156. https://doi.org/10.1016/j.biotechadv.2014.04.006.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

SUPPORTING INFORMATION

Decoupling of recombinant protein production from *Escherichia coli* cell growth enhances functional expression of plant Leloir glycosyltransferases

Martin Lemmerer[1], Juergen Mairhofer[2], Alexander Lepak[3], Karin Longus[3], Rainer Hahn[4], Bernd Nidetzky[1,3,*]

[1]Austrian Centre of Industrial Biotechnology (acib), Petersgasse 14, 8010 Graz, Austria

[2]enGenes Biotech GmbH, Mooslackengasse 17, 1190 Vienna, Austria

[3]Institute of Biotechnology and Biochemical Engineering, Graz University of Technology, NAWI Graz, Petersgasse 12/I, 8010 Graz, Austria

[4]Department of Biotechnology, University of Natural Resources and Life Sciences Vienna, Muthgasse 18, 1190 Vienna, Austria

*Corresponding author: Bernd Nidetzky; e-mail: bernd.nidetzky@tugraz.at; phone: +43 316 873 8400; FAX: +43 316 873 8434

S1

Supporting methods

Construction of the co-expression plasmid vector

The vector pETduo was derived from pET-STRP3 by inserting the UGT71A15 gene into the plasmid harboring the GmSusy gene. Two oligonucleotide primers (pDUO_Ins_fwd, pDUO_Ins_rev) were used.

pDUO_Ins_fwd: gatctcgatcccgcgaaattaatacg

pDUO_Ins_rev: cagggcgcgtcccattcg

The receiving vector was linearized by amplification using two oligonucleotide primers, pETduo_BB_fwd and pETduo_BB_rev. Their use eliminated the F1 ORI in the vector.

pETduo_BB_fwd: ggacgcgccctgtggcacttttcggggaaatgt and

pETduo_BB_rev: gggatcgagatcttcctcctttcagcaaaaaacccct

Fed-batch media and feed composition

Semi-synthetic medium. Media were prepared to enable production of 1580 g dry cells. The composition of the feed is given in relation to the unit mass (g) of cell dry mass: 94.1 mg KH_2PO_4, 31.8 mg H_3PO_4, 41.2 mg $C_6H_5Na_3O_7 \times 2H_2O$ (citric acid, trisodium salt), 45.3 mg $(NH_4)_2SO_4$. Compounds were dissolved in an 8 L feed volume. The other substrate components were added in relation to the grams of cell dry mass to be produced in the batch-phase (80 g): 46.0 mg $MgCl_2 \times 6H_2O$, 20.2 mg $CaCl_2 \times 2H_2O$, 50 μL trace element solution (Table S2) and 3.3 g glucose $\times H_2O$. Initially, the batch medium (10 L) also contained yeast extract (0.15 g/g cell dry mass). Filter-sterilized kanamycin (50 μg/mL) and chloramphenicol (34 μg/mL) were added.

The substrate feed was controlled by increasing pump speed according to an exponential growth algorithm with superimposed feedback control of weight loss in the substrate tank. A

specific growth rate (μ) of 0.111 h^{-1} was held for 15 h. After the first 15 h of feeding, protein expression was induced (see the main text).

Sampling, cell disruption, protein purity and concentration

Sampling. The cell growth was recorded as increase in OD_{600nm}. Samples (2×1 mL; 2×10 mL) were collected by centrifugation at 4°C for 30 min. Pellets were used for determination of cell dry mass. The glycerol concentration in the supernatant was determined by ion exchange HPLC and refractive index detection (see Methods in main text). Cells were harvested after 16 h and 24 h respectively by 30 min centrifugation (Sorvall RC-5B) at 4°C, 4420 g, re-suspended in deionized water to a concentration of ~100 g dry cells /L and stored at -20°C.

Cell disruption in batch cultivation and preparation of sample for Strep-Tactin purification. Cell disruption was performed at 50% amplitude for 1 min (1 s pulse on and 2 s pulse off) using a Sonic dismembrator (Ultrasonic Processor FB-505; Fisher Scientific, Vienna, Austria) equipped with a 3.1 mm micro-tip. Supernatant (cell-free extract) was collected after centrifugation at 21,130 *g* and 4°C for 30 min (Centrifuge 5424 R; Eppendorf, Vienna, Austria) and used for activity assay. For Strep-Tactin purification the same procedure was used and cell extract was additionally filtered through a 1.2 μm filter prior to loading onto the column.

Cell disruption in fed-batch cultivation. Cell disruption was performed by re-suspending cells at a concentration of 50 g_{cdw}/L in 20 mM Tris/HCl, 100 mM NaCl, pH 8.0 buffer and processing the suspension on a Panda 2000 high pressure homogenizer (GEA Westfalia, Oelde, Germany) for two passages at 700 bar. Cell debris was removed by centrifugation at 10,000 *g* at 4°C for 30 min in an Avanti JXN-26 centrifuge (Beckman Coulter, Vienna, Austria).

S3

Protein concentration and purity. The BCA assay was used with bovine serum albumin (BSA) as the standard. Protein purity was determined by SDS-PAGE. Protein bands were stained with Coomassie Brilliant Blue and analyzed by densitometry.

S4

Supporting Tables

Table S1. Composition of the medium used in batch bioreactor cultivations

Component	Concentration (g/L)
K_2HPO_4	4.58
KH_2PO_4	3
tryptone	0.8
yeast extract	0.4
Na_3-citrate·$2H_2O$	2
$MgSO_4$·$7H_2O$	0.8
$CaCl_2$·$2H_2O$	0.08
Trace element solution	0.4[a]
$(NH_4)_2SO_4$	3.6
NH_4Cl	2.96
glycerol 98%	17.36

mL/L

S5

Table S2. Composition of the trace element solution

Component [a]	Concentration (g/L)
$FeSO_4 \cdot 7H_2O$	40
$MnSO_4 \cdot H_2O$	10
$AlCl_3 \cdot 6H_2O$	10
$CoCl_2$	4
H_3BO_3	0.5
$CuCl_2 \cdot 2H_2O$	1
$ZnSO_4 \cdot 7H_2O$	2
$Na_2MoO_2 \cdot 2H_2O$	2

[a] All components were dissolved in 5 M HCl.

S6
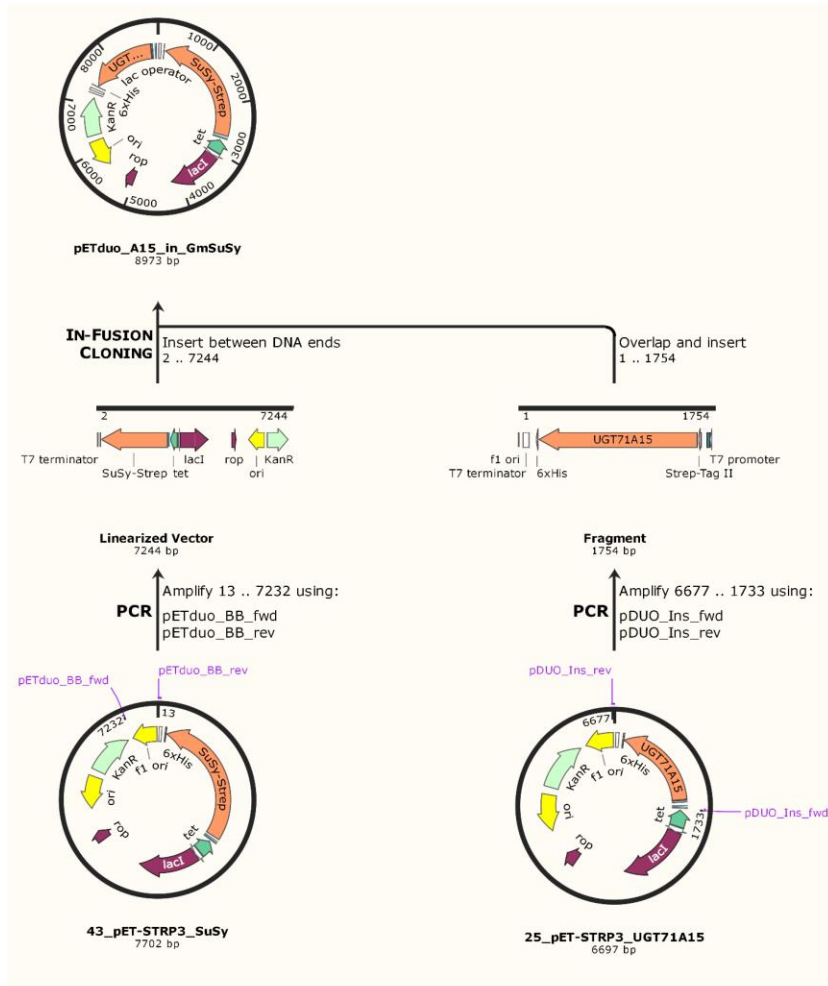
Figure S1. Construction of the enzyme co-expression vector by insertion of UGT71A15 gene into the pET plasmid vector harboring the GmSuSy gene.

Figure S2. Vector map of the pET co-expression vector for production of UGT71A15 and
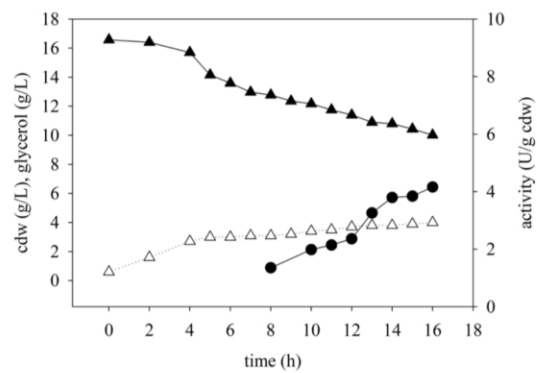
GmSuSy.

Figure S3. Time courses of growth, glycerol consumption and enzyme formation in batch bioreactor cultivations of enGenes-X-press (A) and *E. coli* BL21(DE3) (B) producing GmSuSy. The induction temperature was 25°C. The symbols show: cell dry mass concentration, open triangles; glycerol concentration, full triangles; volumetric enzyme activity, full circles.

S9

A

B
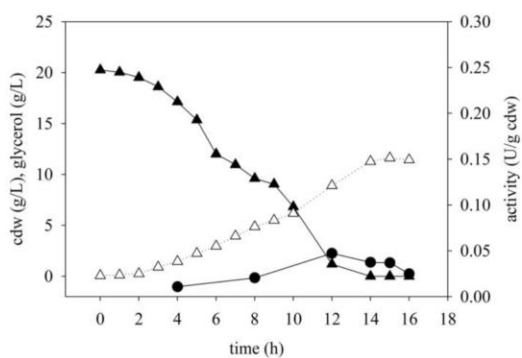
Figure S4. Time courses of growth, glycerol consumption and enzyme formation in batch bioreactor cultivations of enGenes-X-press (A) and *E. coli* BL21(DE3) (B) producing UGT71A15. The induction temperature was 30°C. The symbols show: cell dry mass concentration, open triangles; glycerol concentration, full triangles; volumetric enzyme activity, full circles.
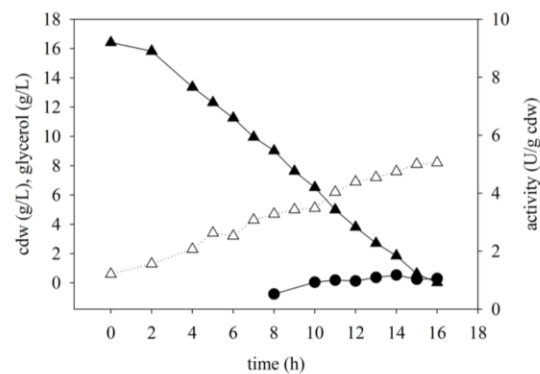
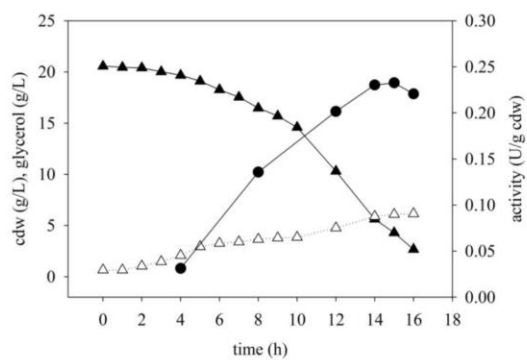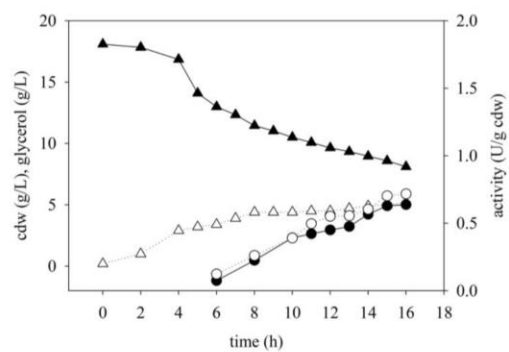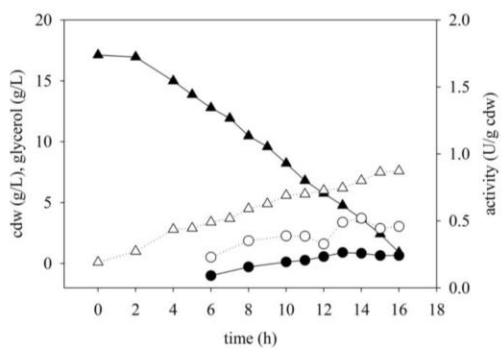**Figure S5**. Time courses of growth, glycerol consumption and enzyme formation in batch bioreactor cultivations of enGenes-X-press (A) and *E. coli* BL21(DE3) (B) producing GmSuSy and UGT71A15. The induction temperature was 25°C. The symbols show: cell dry mass concentration, open triangles; glycerol concentration, full triangles; volumetric enzyme activity, full circles - UGT71A15, open circles - GmSuSy.
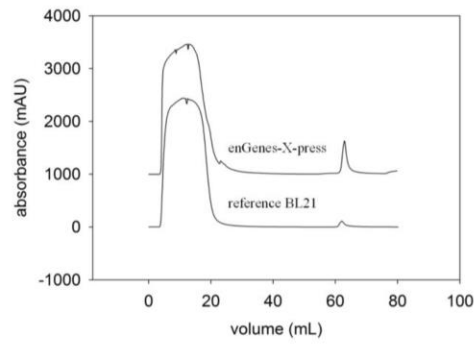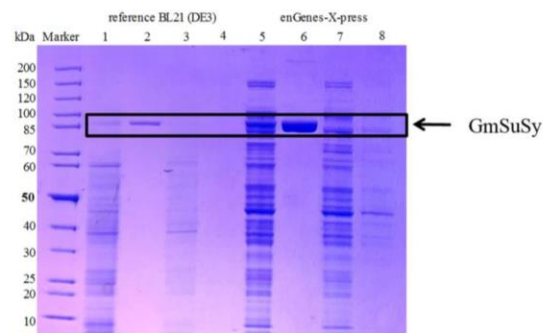
Figure S6. Purification of GmSuSy from cell material produced in batch bioreactor cultivations of enGenes-X-press and the *E. coli* BL21(DE3) reference. A: Absorbance traces of protein elution from Strep-Tactin affinity chromatography. B: Analysis by SDS PAGE; lanes 1 - 4 are fractions from purification of BL21(DE3) material, lanes 5 – 8 are from purification of enGenes-X-press material. Lanes 1 and 5, loading fraction; lanes 2 and 6, elution fraction; lanes 3 - 4 and 7 - 8 are flow through fractions.

S12
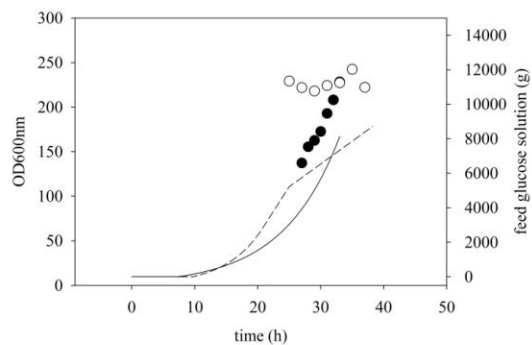
Figure S7. Time courses of fed-batch bioreactor cultivations of enGenes-X-press and *E. coli* BL21(DE3) for production of GmSuSy. Symbols show OD_{600nm} measurements for enGenes-X-press (empty circles) and the BL21(DE3) reference (full circles) after induction. The dashed and the solid lines show the glucose feed for enGenes-X-press and the BL21(DE3) reference, respectively.

S13
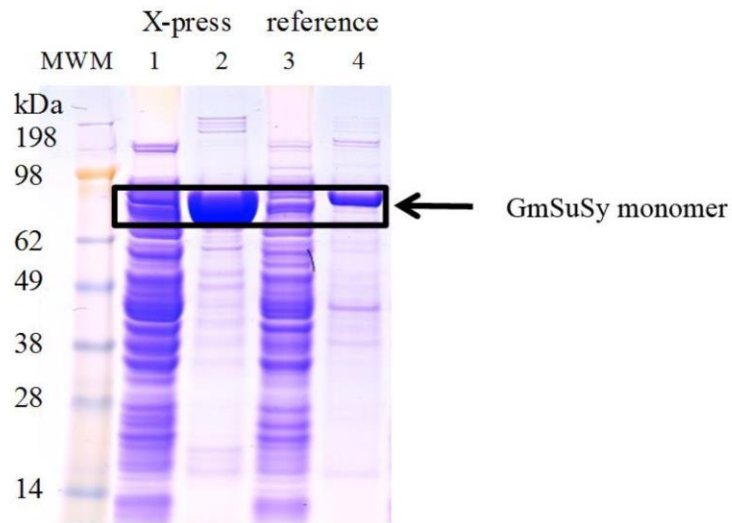
Figure S8. Analysis by SDS-PAGE of the Strep-Tactin column load and the eluting fraction derived from material produced with enGene-X-press and *E. coli* BL21(DE3) in fed-batch cultivation. Lanes 1 and 3 are the column loads; lanes 2 and 4 are eluting fractions. The arrow indicates the expected position of the GmSuSy monomer.
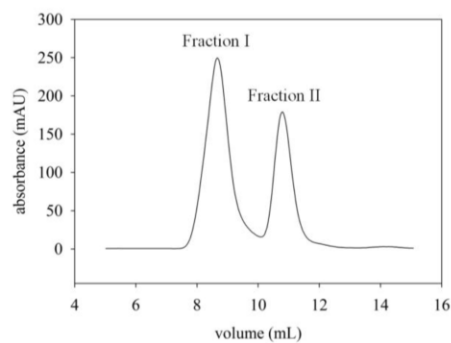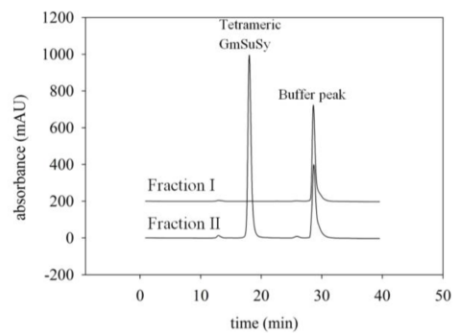
Figure S9. Fractionation of the GmSuSy eluate from Strep-Tactin purification using preparative SEC on a SD-200 column (A) and analysis of the SEC fractions thus obtained on an analytical size exclusion column (B). A: fractions I and fraction II represent agglomerated oligomeric and native (tetrameric) species of GmSuSy, respectively. B: Analytical SEC analysis of fractions I and II.

S15

Chapter 8:

Whole Mass Molecule Spectroscopy Measurement of Kinetic Isotope Effect in *C*- and *O*-glycosyltransferases

Abstract

Glycosylation of natural products is a key mechanism of functionalizing small molecules. For many compounds this reaction is mediated by inverting UDP dependent glycosyltransferases (UGT) which are able to confer a sugar moiety to form either a *O, N, S* or *C* bond. While the formation of the first three is performed through a direct nucleophilic S_N2 mechanism the C-glycosylation mechanism was under debate until the mechanism was also characterized as an S_N2 through direct nucleophilic displacement on the C-1 position of the sugar moiety by a carbanion of the aglycon. To study the reaction mechanisms of glycosyltransferases we measured kinetic isotope effects on the reaction. Labelled UDP-glucose synthesized through a one-pot cascade reaction from labelled glucose allowed measurement of α-secondary deuterium kinetic isotope effect (KIE) on the glycosylation reaction. To implement an easy approach a NMR based and mass spectroscopy based approach was compared. The obtained KIE in flavonoid GTs *Os*CGT and *Pc*OGT from rice and pear, respectively, was compared and showed no significant difference confirming the identical reaction mechanism.

Introduction

Glycoconjugates and glycans are central compounds in biological systems as they play key roles in signaling, recognition and bacterial cell wall synthesis. Many of these important compounds are synthesized by glycosyltransferases (GT) which attach glycosyl moieties by catalyzing the nucleophilic acceptor attack on the anomeric carbon of the sugar (see scheme. 1). GTs can be distinguished depending on product formation as either inverting or retaining enzymes[1]. While retaining GTs perform their reaction by an S_Ni-type reaction[2], inverting GTs utilize nucleoside diphosphate sugars (e.g. UDP-glucose and UDP-galactose) and perform their reaction via a single displacement mechanism with an oxocarbenium ion-like transition state and an asynchronous S_N2 mechanism [3a, b].

Inverting GTs gained interest in biotechnological applications due to their high selectivity, which implies an advantage over unspecific chemical synthesis of glycosylated compounds. Furthermore they play a vital role in functionalization of small molecules [4a, b, c] and proteins, only achieving functionality upon specific glycosylation. The capability of inverting GTs to form *O*-, *N*-, *S*- or *C*- bonds, enables them to form a wide variety of possible compounds which could be target of drug formulation and derivatization[5]. The formations of *C*-glycosides is outstanding in that the *C-C* linkage is highly resilient against hydrolysis.
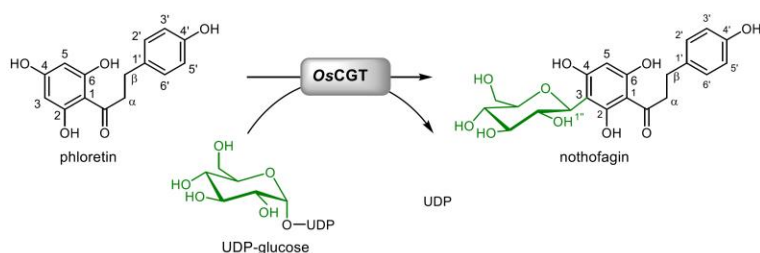
Scheme 1. Glycosylation of phloretin by *Os*CGT. UDP-dependent inverting glycosyltransferases transfer a glycosyl moiety from an activated nucleotide-sugar onto a receiving nucleophile.

In recent studies an *O*-glycosyltransferase from pear (P*yrus communis, Pc*OGT) and a *C*-glycosyltransferase from rice (*Oryza sativa, Os*CGT) were reported to either form phlorizin (2-*O*-glycoside) or nothofagin (3-*C*-glycoside) from phloretin [6]. The study demonstrated the possibility to change the reaction outcome for *Os*CGT to form the *O*-glycoside instead of the *C*-glycoside by introducing the single point mutation I121D. From a mechanistic stand point this is of special interest as it can help to understand if the reaction of both enzymes is performed via a direct nucleophilic attack of the activated carbon at the anomeric C-atom of the sugar or by a rearrangement reaction leading to the *C*-glycoside after initial O-glycosylation (see scheme 2).
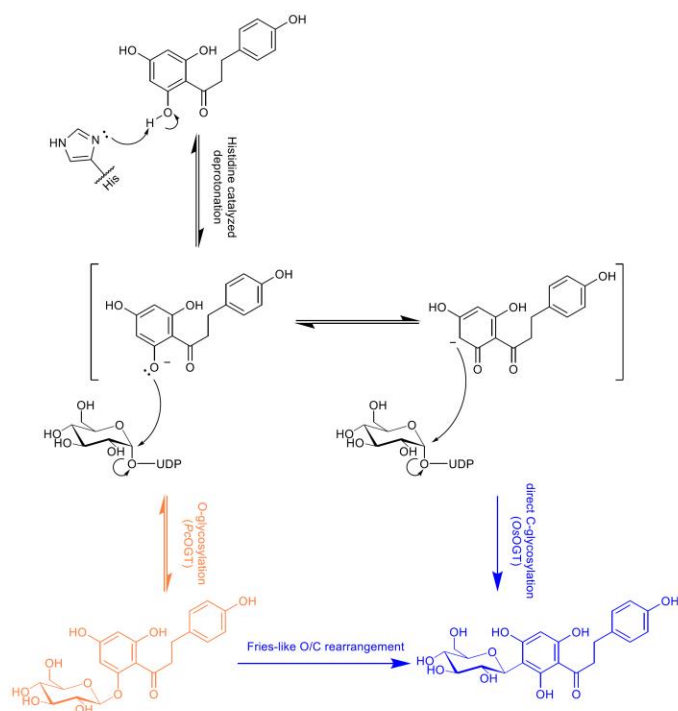
Scheme 2. Proposed mechanism of C-glycosylation. Activation of the aryl acceptor in phloretin is achieved by a conserved Histidine in both CGT (blue) and OGT (orange). Earlier mechanisms proposed an *O-/C-* rearrangement after initial O-glycosylation as performed by OGT.

A technique developed by enzymologists to study transition states of enzymes is the measurement of kinetic isotope effects (KIE) caused by incorporation of heavier isotopes. Enrichments of isotopically labelled substrate or product provide information regarding the rate determining-step by differences in the energy levels of each step[7]. We therefore aimed to measure the KIE of the rate-limiting step for both inverting GTs. Assuming the reaction is achieved by a direct glycosylation, the expected transition state would be similar in both cases and the resulting KIE on V_{max}/K_m identical. Measurement of kinetic isotope effects allow the discrimination of early and late transition states. The kinetic isotope effect we expected to observe for the two investigated inverting GTs is based on a S_N2 reaction in which the catalytic

histidine activates the hydroxyl group in position 2 of phloretin. The charge after proton subtraction is delocalized and the carbanion can act as additional nucleophile in competition with the hydroxyl group. Through different positioning in the active site the glycosylation is favored on position C-3 in case of the *Os*CGT. In the transition state an inversion at the anomeric carbon atom takes place through which the β-glucoside is created. As structures of *Os*CGT and *Pc*OGT are lacking the exact nature of the transition state is unclear.

Calculation of kinetic isotope effects

The following mathematical derivations are from Paul Cook and William Cleland[8]. In brief, isotope effects determined by internal competition on V/K for the labeled substrate the rate equation for the unlabeled substrate is:

$$v_{\mathrm{U}} = \frac{V_{\mathrm{U}}\mathbf{U}}{K_{\mathrm{U}}\left(1 + \frac{\mathbf{L}}{K_{\mathrm{L}}}\right) + \mathbf{U}}$$

(7-1)

Where concentrations of labeled and unlabeled substrate are represented by **U** and **L**. V_U and K_U are kinetic parameters and K_U is the K_M of the unlabeled substrate, which can be interpreted as competitive inhibitor. For the labeled substrate the equation is:

$$v_{\mathrm{L}} = \frac{V_L\mathbf{L}}{K_{\mathrm{L}}\left(1 + \frac{\mathbf{U}}{K_{U}}\right) + \mathbf{L}}$$

(7-2)

Where the unlabeled substrate acts as competitive inhibitor. These reactions can be rearranged by dividing each by the respective K_M value:

$$-\frac{d\mathbf{U}}{dt} = v_{\mathrm{U}} = \frac{(V/K)_{\mathrm{U}}\mathbf{U}}{\left(1 + \frac{\mathbf{L}}{K_{\mathrm{L}}} + \frac{\mathbf{U}}{K_{\mathrm{U}}}\right)}$$

(7-3)

$$-\frac{d\mathbf{L}}{dt} = v_{\mathrm{L}} = \frac{(V/K)_{\mathrm{L}}\mathbf{L}}{\left(1 + \frac{\mathbf{U}}{K_{\mathrm{U}}} + \frac{\mathbf{L}}{K_{\mathrm{L}}}\right)}$$

(7-4)

If equation 7-3 is divided by equation 7-4:

$$\frac{d\mathbf{U}}{d\mathbf{L}} = \frac{(V/K)_{\mathrm{U}}\mathbf{U}}{(V/K)_{\mathrm{L}}\mathbf{L}}$$

(7-5)

Or

$$\left(\frac{d\mathbf{L}}{\mathbf{L}}\right){}^{\mathrm{L}}(V/K) = \frac{d\mathbf{U}}{\mathbf{U}}$$

(7-6)

Which integrates to:

$$\ln\left(\frac{\mathbf{L}}{\mathbf{L}_0}\right){}^{\mathrm{L}}(V/K) = \ln\left(\frac{\mathbf{U}}{\mathbf{U}_0}\right)$$

(7-7)

Or

$${}^{\mathrm{L}}(V/K) = \frac{\ln\left(\dfrac{\mathbf{U}}{\mathbf{U}_0}\right)}{\ln\left(\dfrac{\mathbf{L}}{\mathbf{L}_0}\right)} = \frac{\log\left(\dfrac{\mathbf{U}}{\mathbf{U}_0}\right)}{\log\left(\dfrac{\mathbf{L}}{\mathbf{L}_0}\right)}$$

(7-8)

The fractional reaction is defined as:

$$f = \frac{(\mathbf{U}_0 + \mathbf{L}_0) - (\mathbf{U} + \mathbf{L})}{(\mathbf{U}_0 + \mathbf{L}_0)} = 1 - \frac{(\mathbf{U} + \mathbf{L})}{(\mathbf{U}_0 + \mathbf{L}_0)}$$

(7-9)

The isotope ratio in the product at fractional reaction f is:

$$R_P = \frac{\mathbf{L}_0 - \mathbf{L}}{\mathbf{U}_0 - \mathbf{U}}$$

(7-10)

While the ratio in residual substrate at f is:

$$R_S = \frac{\mathbf{L}}{\mathbf{U}}$$

(7-11)

X_{U} and X_{L} represent initial mole fractions:

$$R_0 = \frac{\mathbf{L}_0}{\mathbf{U}_0} = \frac{X_{\mathrm{L}}}{X_{\mathrm{U}}}$$

(7-12)

And it follows:

$$\text{KIE} = {}^{\text{L}}(V/K) = \frac{\log\left[1 - \dfrac{f}{(X_u + \frac{X_\text{L} R_\text{P}}{R_0})}\right]}{\log\left[1 - \dfrac{f}{\left(X_\text{L} + \frac{X_\text{U} R_0}{R_\text{P}}\right)}\right]} = \frac{\log\left[\dfrac{(1-f)}{(X_\text{U} + \frac{X_\text{L} R_\text{S}}{R_0})}\right]}{\log\left[\dfrac{(1-f)}{\left(X_\text{L} + \frac{X_\text{U} R_0}{R_\text{S}}\right)}\right]} \tag{7-13}$$

From these equations we can simulate the variation of isotope ratios with f values over the course of the reaction for different KIE values (see fig. 1). To determine ${}^{\text{L}}(V/K)$ from R_S and R_P, f values of $0.5 - 0.7$ and $0.2 - 0.3$ give the lowest error margins, respectively. In this study we did not rely on single measurements at a given f value, but instead took several samples over the course of the reaction to minimize the error derived from determination of f.

7-13 can now be solved as X_U and X_L are defined by the initial amounts of substrate and R_0 is defined by X_U and X_L itself. The only values which have to be determined by measurements are therefore R_P, R_S and f. To simulate curves for different KIEs R_P and R_S were fitted and plotted against a given f to give raise to the KIE of interest. For the given plot R_P was varied and R_S resulted from the inverse calculated amounts of used substrate assuming an initial equimolar ratio of labelled and unlabelled substrate. Plots are normally using R_P / R_0 or R_S / R_0 so that independent of the chosen ratio of substrates at the start of the reaction the lack of a KIE would always result in a value of 1.
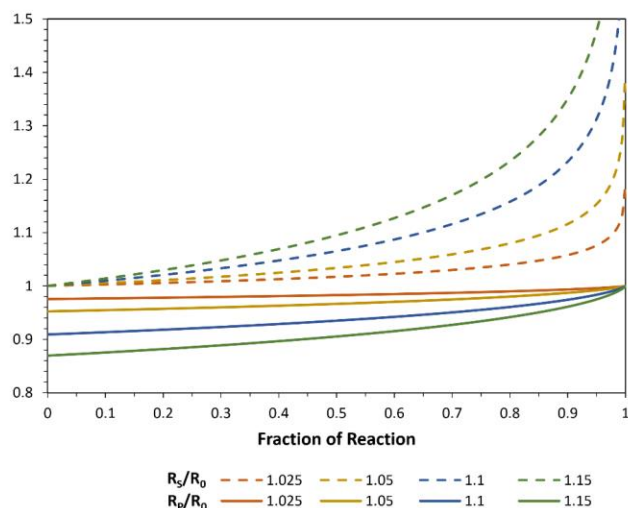
Figure 1. Predicted isotope ratios in substrate (R_S) and product (R_P) as a function of fractional reaction f and the initial isotope ratio in the substrate (R_0). Four different $^1(V/K)$ values of 1.025, 1.05, 1.1 and 1.15 were assumed for calculation.

Methods

Materials

Isotopically labelled [1-^{2}H], [6,6-^{2}H,^{2}H] and unlabeled D-glucoses were obtained from Omicron Biochemicals (South Bend, Indiana, USA) and had isotopic purities of 98%, 95% and 98%, respectively. UTP (> 99% purity) was bought from Carbosynth (Compton,UK). Bovine serum albumin (> 98% purity) were bought from SigmaAldrich (Vienna, Austria). Deuterium oxide for NMR measurements (99.96% ^{2}H) was aquired from Euriso-Top (Saint-Aubin Cedex, France). *Strep*-Tactin Sepharose and desthiobiotin were bought from IBA. All other chemicals were either from SigmaAldrich or Roth and were of the highest purity available. *S. cerevisiae* hexokinase (25 U mg^{-1}), rabbit muscle phosphoglucomutase (≥ 100 U mg^{-1}) and S. cerevisiae inorganic pyrophosphatase (≥ 500 U mg^{-1}) were purchased from Sigma-Aldrich. Calf intestine alkaline

phosphatase was from New England Biolabs (Ipswich, MA, USA). Gene BLLJ_1074 coding for UTP-glucose-1-phosphate uridylyltransferase from *Bifidobacterium longum subsp. longum* JCM 1217 (UniProt:E8MIY8), cloned into a pET-30a vector using *Nde*I and *Xho*I restriction sites for expression with a C-terminally fused His$_6$-tag, was kindly provided by Dr. Motomitsu Kitaoka (National Agriculture and Food Research Organization, Japan). The complete gene was sequenced before creating an expression strain by transformation of electrocompetent *E. coli* BL21 Gold (DE3) cells.

Cloning, Expression and purification and expression

Construction of the *E. coli* BL21 Gold (DE3) strain expressing *Pc*OGT (UGT88F2; GenBank: FJ854496) and *Os*CGT (GenBank: FM179712) was described elsewhere [8]. For expression of glycosyltransferases the respective strain was grown at 37°C and 120 rpm in 1 L baffled shake flasks in 250 mL LB-medium containing 50 µg mL^{-1} kanamycin until an optical density at 600 nm of around 0.8 was reached. The temperature was then decreased to 18°C and expression was induced by addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). Cells were harvested after 18 h by 30 min centrifugation at 4°C, 5000 rpm, resuspended in ddH2O and frozen at -20°C until cell lysis by sonication (10 s on, 20 s off, 6 min total on-time, 400 watt). The cell lysate was cleared of debris (20,000 x g, 60 min), filtered through a 1.2 µm filter and loaded onto 2 x 5 mL StrepTrap-HP columns and purified according to the manufacturer's instructions. Thereafter enzyme containing fractions were pooled, concentrated using a centrifugation concentrator and purified for a second time by anion exchange chromatography on a 5 mL HiTrapQ HP column using a gradient from 0-1 M NaCl over 10 CV to remove residual nucleotides (Buffer A: 100 mM Tris/HCl pH 7.5, Buffer B: 100 mM Tris/HCl pH 7.5, 1 M NaCl). Fractions containing the desired glycosyltransferases were concentrated and buffer exchanged to 25 mM HEPES pH 7.5 using Vivaspin-centrifugal concentrators (Sartorius Stedim, Vienna, Austria) with a molecular cut-off weight of 30 kDa. Expression and purification of UGPase was previously reported in detail[9]. Enzyme purity was checked by SDS-polyacrylamide gel electrophoresis (PAGE) and protein

concentrations were determined by BCA assay using BSA as standard. Enzyme stocks were stored at -20°C. The specific protein activity for *Os*CGT and *Pc*OGT was 2.2 U mg^{-1} and 0.98 U mg^{-1}, respectively.

Synthesis of UDP-glucose

The synthesis is shown schematically in Scheme 3. The method for enzymatic synthesis of UDP-α-D glucose was adapted from literature [9]. The reaction mixture (10 mL) contained 15 mM (isotope-labeled) monosaccharide , 45 mM uridine 5'-triphosphate (UTP), 5 mM MgCl$_2$, 3 µM α-D-glucose 1,6-bisphosphate and 0.13%(w/v) bovine serum albumin (BSA) dissolved in 50 mM Tris/HCl buffer (pH 7.5). 12 U mL^{-1} hexokinase , 6 U mL^{-1} phosphoglucomutase, 1.2 U mL^{-1} UTP-glucose-1-phosphate uridylyltransferase (UGPase) and 1.2 U mL^{-1} inorganic pyrophosphatase (PPase) were added and the reaction was incubated at 30 °C for 24 hrs. Prior to chromatographic purification of UDP-glucose, enzymes were removed by ultrafiltration using Vivaspin-centrifugal concentrators (Sartorius Stedim, Vienna, Austria) with a molecular cut-off weight of 10 kDa. Purification of the produced nucleotide sugars was done using an ÄKTA FPLC liquid chromatograph (GE Healthcare, Vienna, Austria) equipped with a 5 mL SuperQ 650M anion exchange column and a 2 mL sample loop. A gradient over 100 ml at 4 ml min^{-1} from 0 % B to 20% B (buffer A: 20 mM sodium acetate pH 4.3 and buffer B: 1 M sodium acetate pH 4.3) was used for elution of bound compounds. Prior to each run, the column was regenerated flushing it with at least three column volumes of buffer B and buffer A. UV absorption (λ = 254 nm) was used to detect the target compounds, which were collected. All product-containing fractions were pooled and concentrated by freeze drying. Sodium acetate was removed from nucleotide sugar preparations using ÄKTA FPLC with a 2 mL sample loop and a Superdex Peptide 10/300 GL size exclusion column (GE Healthcare). Elution was performed with deionized water at a flow rate of 1 mL min^{-1}. The target compound was detected by UV absorption (λ = 254 nm). Residual H$_2$O was removed by lyophilization on a Christ Alpha 1-4 lyophilizer (B. Braun Biotech International, Melsungen,Germany), after which the sugar nucleotide product was obtained as white powder. Product identity was confirmed by NMR and product purity was confirmed to be about 98% (see fig. 2).
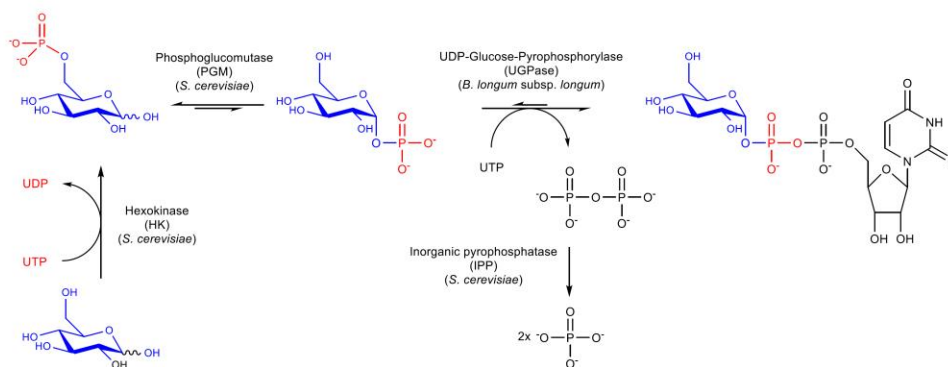
Scheme 3. Preparation of UDP-glucose. UDP-glucose is synthesized from glucose in a one-pot reaction using 4 enzymes consuming 2 molecules of UTP per molecule of UDP-glucose. This reaction scheme was used to prepare labeled UDP-glucose from labeled glucose derivates.
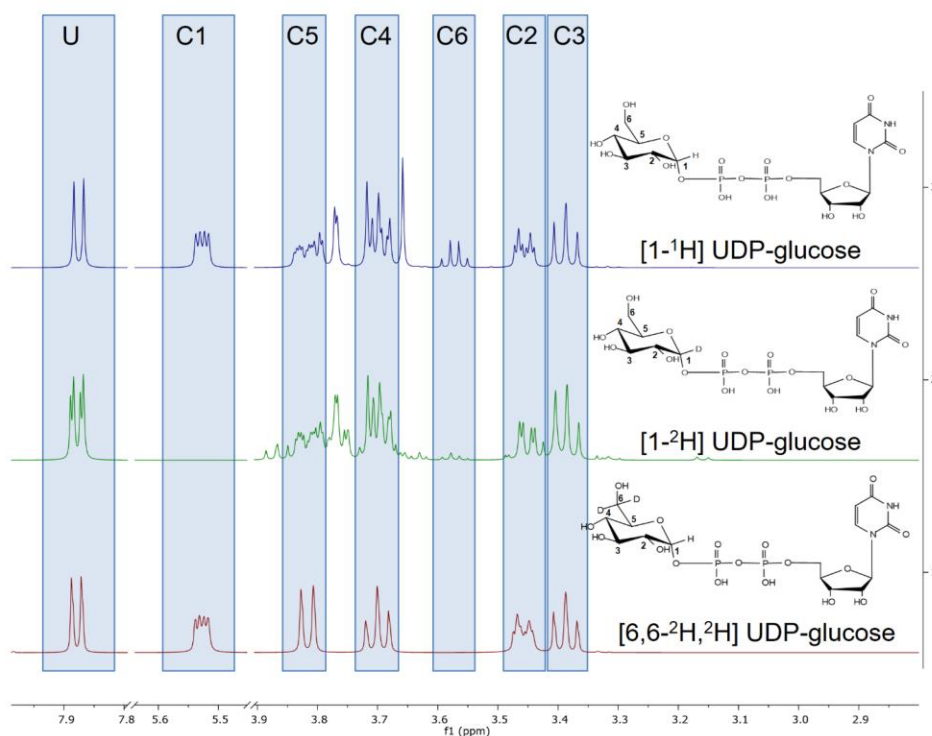

Figure 2. NMR spectra of labeled UDP-glucose. Spectra of each UDP-glucose show incorporation of deuterium atoms. Shown are the characteristic ranges corresponding to the Glucosyl ring. The high purity of product is confirmed by absent signals for position C1 and C6 caused by deuteration. U: uridine, C1 – C6: position 1-6 in the glucose moiety

KIE measurements by in-situ NMR

In a first attempt the formation of labeled an unlabeled product was observed by in-situ NMR. The reaction was performed in 50 mM deuterated phosphate buffer pH 7.5 containing 0.5 mM 1H UDP-glucose, 0.5 mM 1D UDP-glucose, 1 mM Phloretin and 0.05 µM *Os*CGT. The respective enzyme was purified as described above with the exception that the buffer was exchanged towards 50 mM deuterated phosphate buffer pH 7.5 after pooling fractions. All components were prepared as 10-fold stock solutions in 1-fold deuterated phosphate buffer. 60 µl of the stock solutions for all components were mixed with 10 µl of protein solution and filled up with 1-fold deuterated buffer to a final volume of 600 µl. In this way addition of non-deuterated solvent was avoided. The reaction was started by addition of the enzyme and was than shimmed and locked on the deuterium. T_0 is defined by the first recorded spectrum approximately 10 minutes after addition of enzyme. 1 dimensional proton in-situ measurements were recorded at 37°C on a Varian 500 MHz spectrometer every 20 minutes for 15h. 1-H NMR spectra were measured at 499.98 MHz on a 5 mm indirect detection PFG-probe. The water signal was pre-saturated by a shaped pulse, using a standard pre-saturation sequence: relaxation delay 2 s; 90° proton pulse; spectral width 8 kHz; number of points 32 k.

Data was interpreted using Mestrenova 12.0. After automatic identification of solvent peaks, artificial spectra were iterated to remove solvent signals. These reiterated spectra were used for integration of signal peaks. The integration values were later used for calculation of conversions and molar fractions of individual compounds.
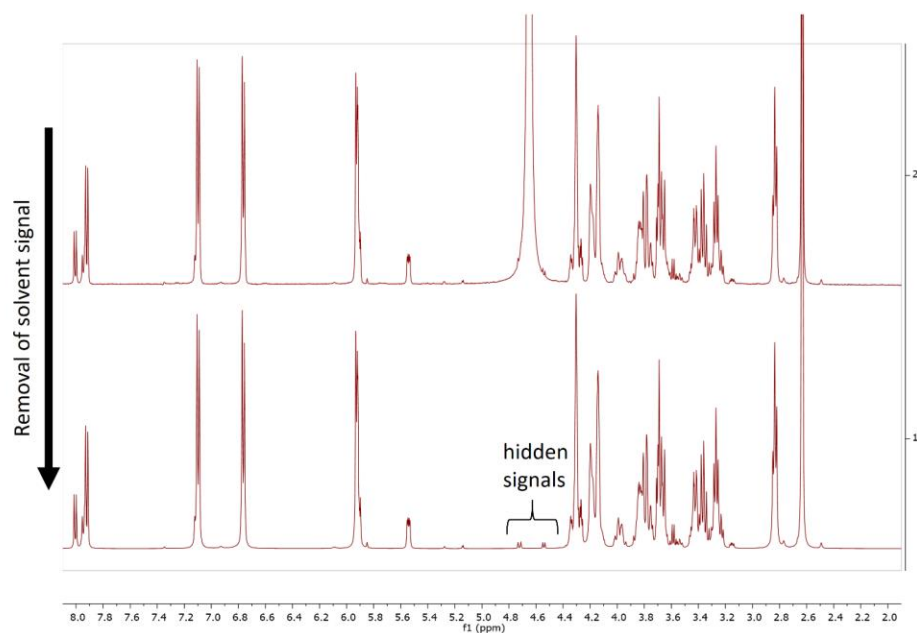
Figure 3. NMR spectra from in-situ measurements. The depicted spectrum at 3h was corrected using Mnova to remove solvent peaks and allow identification of hidden peaks.

Table 1. Chemical shifts of carbon attached protons. Compounds were identified in-situ, based on the given chemical shifts and integrated peak areas were used for calculation of reaction progress.

		Phloretin	Nothofagin
C-1''	Anomeric carbon		4.72 ppm
C-3/C-5	Ring A	5.93 ppm	5.90 ppm
C-2'/C-6'	Ring B	7.10 ppm	7.12 ppm
C-3'/C-5'	Ring B	6.75 ppm	6.76 ppm
		UDP	UDP-glucose
C-5	Base	7.94 ppm	7.92 ppm
C-1''	Anomeric carbon		5.53 ppm
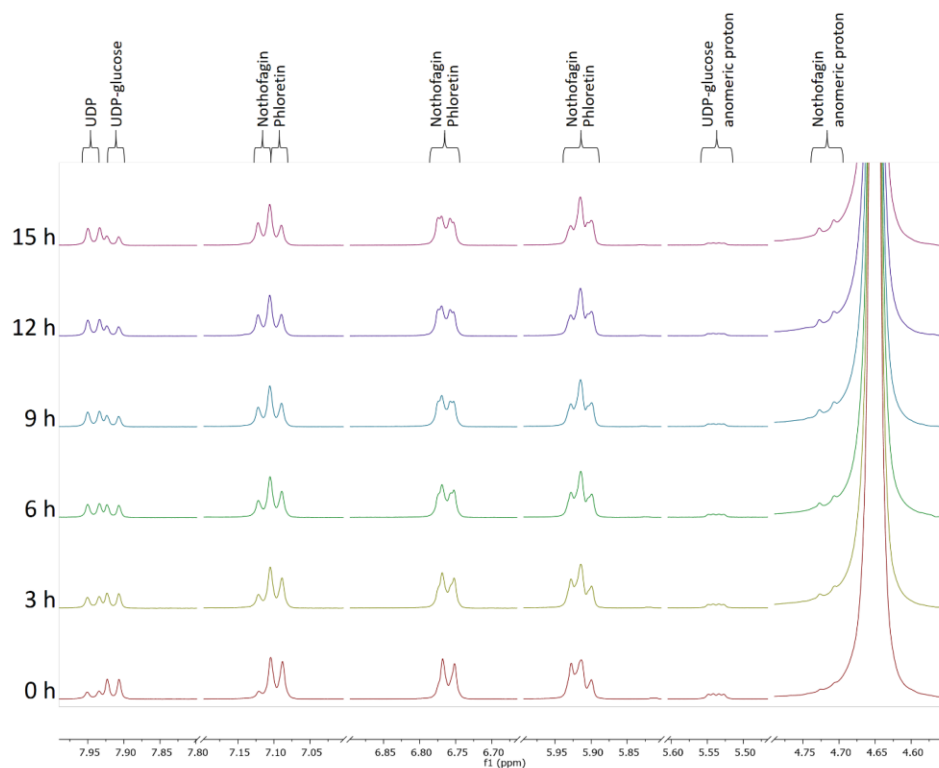
Figure 4. Time resolved in-situ NMR spectra. Formation of UDP and nothofagin was detected by their respective signals. Phloretin (δ=7.10, 7.11) was converted into nothofagin (δ=7.11, 7.12) under consumption of UDP-glucose (δ=5.53/7.92,7.93) thereby releasing UDP (δ=7.93, 7.94). The reaction was performed using 1 mM phloretin, 0.5 mM of 1D and unlabeled UDP-glucose in deuterated phosphate buffer at pH 7.5 in presence of 0.05 µg/ml of *Os*CGT.

Proton signals of glucose 1''H and acceptor 3H were compared to signals of position 2' and 3' of the acceptor to observe conversion of phloretin and calculate 1D/1H ratios. For overlapping peaks a symmetrical splitting was assumed and the two non-overlapping signal areas were each multiplied by two to calculate ratios.

HPLC-based determination of phloretin-(glucosides) and UDP-(glucose)

Phloretin, its glucosides and uridine-derivates (such as UMP, UDP and UDP-glucose) were analyzed in a combined method using reversed-phase ion-pairing HPLC with tetra-*n*-butylammonium bromide (TBAB) as ion pairing reagent. HPLC analysis was performed on a reversed phase C18 column (Kinetex 5 µm C18 100 Å column 50 x 4.6 mm, Phenomenex) at 35°C. 20 mM potassium phosphate, pH 5.9 containing 40 mM TBAB were used as mobile phase A. It was prepared by dissolving TBAB in a 5-fold diluted 100 mM potassium phosphate buffer (pH 5.9) stock. Acetonitrile was used as mobile phase B. Separation was achieved using following method at a constant flow rate of 2 mL min^{-1}: 10% B (1 min), 10-40% B (3.2 min), 40-90% B (0.01 min), 90% B (0.79 min), 90-10% B (0.01 min), 10% B (1.99 min). Uridine and phloretin derivatives were monitored at 262 and 288 nm, respectively. UMP, UDP, UDP-glucose, phloretin, nothofagin and phlorizin were used as authentic standards and concentrations were calculated from peak areas. UMP, UDP, UDP-glucose, phloretin and phlorizin were purchased from Carbosynth, whereas Nothofagin was synthesized in-house and purity confirmed by NMR.[10]

The detection limit for nucleotides was 0.5 µM [10] and allowed identification of residual UDP and UDP-glucose in purified enzyme. Measuring conversions of acceptors with defined ratios of labeled and unlabeled UDP-glucose made a nucleotide free preparation of the proteins essential. To validate nucleotide free preparation of the glycosyltransferases 500 µl samples containing 2 mg ml^{-1} (~ 40µM) of enzyme were denatured by addition of 500 µl Methanol. Precipitated protein was pelleted by centrifugation (20,000 x g, 60 min) and the remaining supernatant tested for presence of nucleotides. This control guaranteed that less than 2.5 % of the purified protein contained nucleotides.

Given a detection limit of 0.5 µM for nucleotides, I can assume that the error for the determination of f is 0.05% as a total of 1 mM of UDP-glucose was used.

For phloretin derivates the detection limit is about 50 nM and the resulting error on the determination of f derived from product measurements is about 10-fold lower and was therefore used for the calculation of f used in the calculation of $^{L}(V/K)$.

Determination of glycosyltransferase catalyzed conversions

The described HPLC assay applying ion pairing was used to quantify phloretin and its glucosides. *Os*CGT and *Pc*OGT reactions were performed in 200 µl scale. The reaction was performed in 50 mM 4-(2-hydroxyethyl)morpholine buffer pH 7.5 containing 0.5 mM labeled UDP-glucose, 0.5 mM unlabeled UDP-glucose, 2 mM Phloretin and 0.05 µM of the corresponding glycosyltransferase and 5% DMSO. For each KIE calculation 10-15 samples were taken. Reactions were started by addition of enzyme at 30°C without shaking. To stop the reaction samples of typically 5 µL were mixed with 25 µl methanol and 20 µl ddH$_2$O and centrifuged for at least 15 min at 13,200 rpm to remove precipitated compounds. Before HPLC analysis samples were split and one sample was used for TBAB based determination of conversion and the other uncontaminated sample was stored at -20°C until analysis by for whole mass spectroscopy.

For *Os*CGT and *Pc*OGT, phloretin glucosylation was performed with 1D and 6,6D UDP-glucose in competition with unlabeled UDP-glucose each, summing up to four experiments. Further each set-up was tested for influence of Calf Intestinal Phosphatase (CIP) to validate if the presence of UDP and possible back reaction of the product is influencing the determined KIE. For this purpose each reaction was performed presence and in absence of CIP. To study the influence of UDP degradation by phosphatase a final concentration of 1 U/ml calf intestinal phosphatase from SigmaAldrich (Vienna, Austria) was added to the reactions.

Whole molecule mass based isotope enrichment measurements

High Resolution-ESI-MS measurements were performed on a Q-Exactive Hybrid Quadrupole-Orbitrap MS with direct injection of samples by a Dionex Ultimate 3000 series UHPLC (Thermo Fisher Sci., Erlangen, Germany) at 0.1 ml/min ddH$_2$O containing 0.1% formic acid. The HR-MS was furnished with an electrospray ionization source (ESI) using nitrogen as nebulizer and drying gas. Measurements were performed in negative mode, with a drying gas temperature of 350 °C, spray voltage of 3.2 kV, and a resolution of 70000 unless otherwise stated. The observed mass range was set to m/z 430-580 with subsequent recording of MS ions. For each injection a single raw file was created.

Data preparation using the Lipid Data Analyzer

Collected data was processed using the Lipid Data Analyzer[11]. In a first step the folder containing the raw files and the quantitation list was loaded using the batch quantitation option. When starting the processing the raw files are automatically converted and for each chromatogram an excel file containing the fully automatic integrated results is created. These results however assume a natural ^{13}C abundance and are not properly deisotoped due to the labeled substrates which do not match with the assumed abundance of natural isotopes as the ^{13}C and ^{2}H isotopes both lead to an increased mass of 1. Therefore the excel files have to be opened individually and the lines containing the integration results have to be deleted. Afterwards each chromatogram is opened manually and the settings from table 2 are applied. Afterwards the automatic batch integration provides correct results as the borders are now set for the labeled substrates and the automatic deisotoping will only remove the excess signals from the natural ^{13}C abundance but not the ^{2}H counts. After this semi-manual treatment an excel file containing all signal counts for each compound, which was assigned with the manual window settings, can be created when loading the raw file containing folder into the statistical data analysis tool and exporting the compound counts of interest.

Table 2. LDA settings for manual chromatogram integration and data processing. Each window was set manually in the LDA to define the range of signal that had to be integrated.

Compound	Retention time	Mass	Mass [-H form]	Window width
1H-Nothofagin	0.05 – 0.9 min	436.1369471	435.12912	0.03 Da
1D-Nothofagin	0.05 – 0.9 min	437.1432239	436.13540	0.03 Da
6,6D-Nothofagin	0.05 – 0.9 min	438.1495006	437.14168	0.03 Da
1H-Phlorizin	0.05 – 0.9 min	436.1369471	435.12912	0.03 Da
1D-Phlorizin	0.05 – 0.9 min	437.1432239	436.13540	0.03 Da
1H-UDP-glucose	0.05 – 0.9 min	566.0550216	565.04720	0.03 Da
1D-UDP-glucose	0.05 – 0.9 min	567.0612983	566.05347	0.03 Da
6,6D-UDP-glucose	0.05 – 0.9 min	568.0706716	567.05975	0.03 Da

Data interpretation and KIE calculation

After integrating data using the LDA, isotope ratios are calculated using excel and KIE's were calculated using equation 7-13 [12a, b]. R_0 is given as 1 because labelled and unlabelled UDP-glucose were mixed in an equimolar ratio and possible pipetting errors were not taken into account.. R_P, R_S, X_U and X_L however were obtained from total ion counts integrated by the LDA. For each f value a KIE was calculated using R_P for product based KIE and R_S for substrate based KIE. f values were obtained by HPLC-UV measurements as described earlier. So the resulting error on the KIE includes uncertainties of the HPLC-UV based determination of f and the integration data from the MS.All KIE's of a reaction were then averaged and the error is the mean deviation from this average.

KIE correction for initial substrate ratio errors by fitting of R_0

Both KIE's calculated from R_P and R_S are depending on R_0 and its proper determination. To improve the quality of KIE calculation we tried to compensate for non-equimolar initial mixing of substrates, which would result in R_0 that is not 1. To compensate this error the measurement of R_0 is necessary but would also be dependent on the measurement quality and introduce new errors. For this reason R_0 was treated as a variable in equation 7-13. For each time point of a data series R_0 has to be identical as it is the initial ratio of the substrates. Assuming that the KIE for each f value is the same, R_0 was solved by averaging the KIE for each f value and iterating R_0 until the sum of deviations of KIE's from all four reactions using either 1D or 6,6D UDP-glucose became minimal. This is possible because each reaction of the two labeled substrates was set up with the same mix of labeled and unlabeled substrate. The KIE was then recalculated using the determined R_0.

Correction of isotope dependent ionization

To eliminate the isotope dependent effect on measured isotope ratios, we determined the discripency between measured and expected isotope ratios for each reaction. As the sum of labeled substrate and labeled product has to be constant over the course of the reaction R_0 is:

$$R_0 = R_{tot} = (1 - f_1) R_S + f_1 R_P \tag{7-14}$$

Using these two assumptions the real R_P and R_S can be determined by correcting the measured ratios. The function of $R_0 = (1 - f_1) R_S + f_1 R_P$ can be described as $R_0 = a * f_1 + c$, where a is the slope of the trend line and c is the intersection with the y axis when plotting measured R_{tot} against R_0.

For $f_1 = 0$:

$$R_0 = R_s \tag{7-15}$$

For R_S follows:

$$R_{S\,real} = \frac{R_{S\,measured}}{(c)} \tag{7-16}$$

And for $f_1 = 1$:

$$R_0 = R_P \tag{7-17}$$

For $R_P\ follows$:

$$R_{P\,real} = \frac{R_{P\,measured}}{(a + c)} \tag{7-18}$$

Using the constants a and c determined from trendlines for each reaction R_P and R_S were corrected and KIE's recalculated using fitted R_0, which were determined as described earlier.

Determination of forward commitment to catalysis

To determine forward commitment to catalysis reactions were performed in 200 µl scale by isotope trapping. The reaction was performed in 50 mM 4-(2-hydroxyethyl)morpholine buffer pH 7.5. 200 µM OsCGT were preincubated with 10 mM [1-^{2}H] UDP-glucose at 30°C for 10 min allowing formation of an equilibrium enzyme–substrate complex (ES). 180 µl chase solution containing 3.3 mM unlabeled UDP-glucose, 5.5 mM phloretin and 5.5% DMSO were rapidly mixed with the 20 µl of the ES solution to start the reaction. Samples were then taken as fast as possible in approximatedly 0.9s time intervals. To stop the reaction samples of typically 5 µL were mixed with 25 µl methanol and 20 µl ddH$_2$O and centrifuged for at least 15 min at 13,200 rpm to remove precipitated compounds. Before HPLC analysis samples were split and one sample was used for TBAB based determination of conversion and the other uncontaminated sample was stored at - 20°C until analysis by for whole mass spectroscopy. The turnover number was calculated from the enzyme concentration and the product concentration, whereas the concentration of Enzyme-Substrate complex was assumed to be identical to the enzyme concentration. The concentration of labeled nothofagin was calculated from the absolute nothofagin concentration and the ratio of labeled nothofagin determined by MS.

Assuming that the equilibration is much faster it is to be expected that no enrichment of labelt product during the first turnovers occurs. We choose therefore to divide the concentration of formed labeled [1D nothofagin] product by the concentration of Enzyme-Substrate complex [ES] which can be assumed to be equal to the concentration of enzyme itself.

$$^{L}(V/K) = \left(k + C_f\right)\big/\left(1 + C_f\right) \tag{7-19}$$

The ratio of labeled nothofagin produced to the initial concentration of ES complex was plotted as a function of turnovers and extrapolated back to turnover 0. C_f was calculated from the intersection of trend line with the Y axis.[13] The distribution of enzyme-bound substrate proceeding to form product before equilibrating with free substrate is expressed as C_f [14]

Results

KIE measurements by NMR

The first approach to measure the kinetic isotope effect on the transfer of a glucose moiety through an inverting glycosyltransferase was performed by in-situ NMR. Initially the sensitivity of this method was tested by studying the formation of nothofagin catalyzed by *Os*CGT from unlabeled UDP-glucose. Phloretin and its derivates contain several protons which would be possible candidates for probing. Ring A contains two protons at position C-3 and C-5 of which one is lost during glycosylation. Ring B contains four proton of which two always share the same chemical shift due to the symmetry of the ring, i.e. C-2'/C-6' and C-3'/C-5'. In solution we observed that protons C-2'/C-6' and C-3'/C-5' in the B-ring showed a slight shift of 0.02 ppm acceptor and glycosylated product. The proton signals for C-3 and C-5 are also identical in phloretin. This symmetry is however broken by the glycosylation at position C-2 (phlorizin) or C-3 (nothofagin) and the C-5 signal is shifted by 0.03 ppm (see table 1 and fig. 4).

Therefore we calculated conversion by comparing product signals of either C-5 or C-1''with the C-2'/C-6' and C-3'/C-5'signals as reference. Furthermore the decrease of UDP-glucose was detected by the signal shift at the signal δ=7.92 for nucleotide base proton. We found that only product detection via C-3 and referencing to C-2/C-6 allowed us to follow the reaction progress in a reliable way (see figure 5). The product measurement through detection of the C-1'' signal had the highest error margin due to the fact that the anomeric carbon signal is buried under the Water peak and cannot be determined with high precision. The reaction was stopped by addition of methanol after recording the last proton spectrum and HPLC measurement confirmed the achieved conversion of ~60% observed by NMR which was in accordance with end point determination. Still this conversion was lower than the expected 100% conversion and was achieved much slower than usual (15h instead of 4h). The phenomenon of decreased enzymatic activity in strong magnetic fields was observed for other enzymes as well and is not fully explained yet. For increased reliability of calculated KIE values higher conversions would be beneficial for the determination of substrate dependent enrichments as explained earlier. For this reason the aim would be to record data from no conversion to full conversion.

Figure 5. In-situ NMR determinations of phloretin C-glycosylation by *Os*CGT. Compared are different reaction curves based on combinations and product signals in Ring A and different reference signals in Ring-B. In addition the reaction progress was also calculated from the UDP-glucose signal and UDP release. The reaction was performed using 1 mM phloretin, 0.5 mM of 1D and unlabeled UDP-glucose in deuterated phosphate buffer at pH 7.5 in presence of 0.05 µg/ml of *Os*CGT.

nothofagin

Reference backbone signal
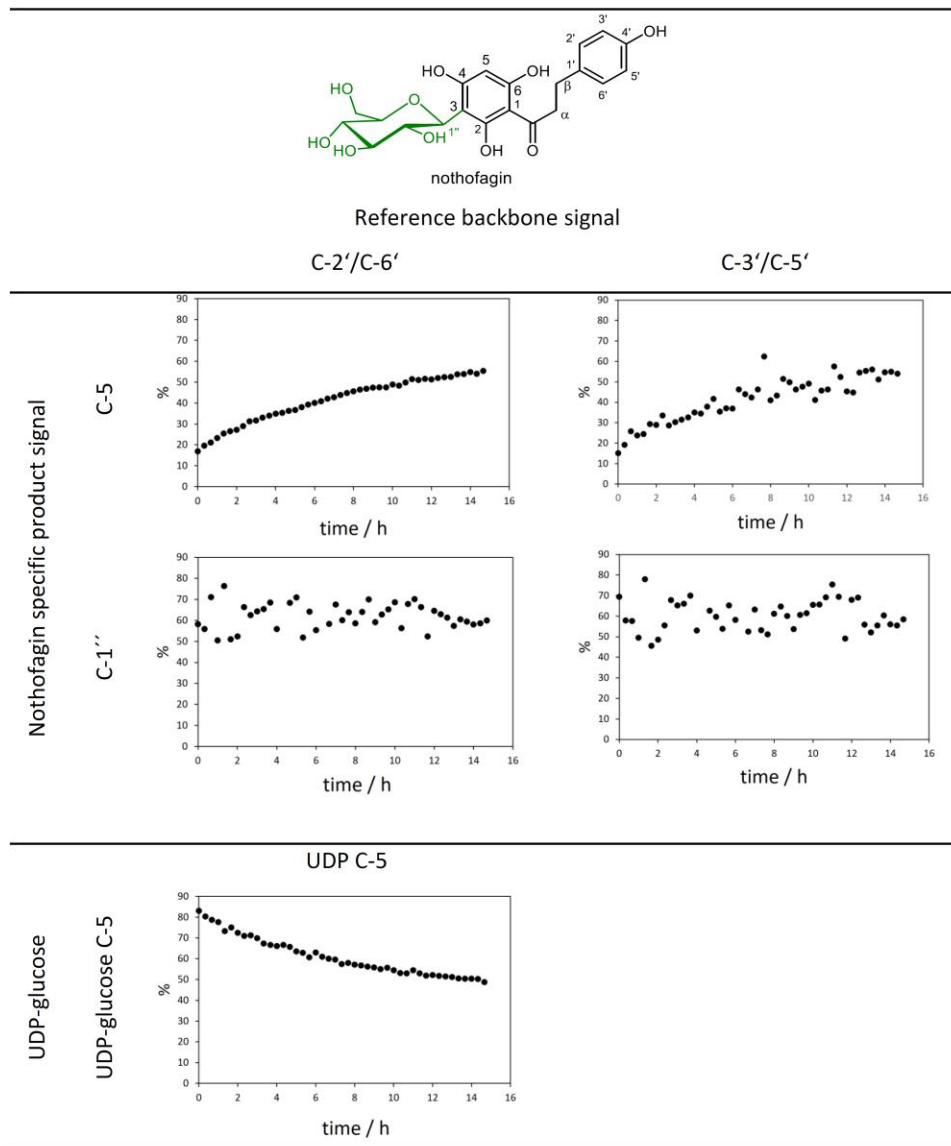
KIE measurements by HPLC-MS

As the implemented NMR based approach does not allow to determine isotopic enrichments we decided to apply a HPLC-MS based method which was established in cooperation with the Institute of Chemistry – Analytical Chemistry, University of Graz. The first tests to separate the reaction compounds by reverse phase chromatography revealed that nucleotide derivates of each injected sample adhere to the column and made discrimination of nucleotide origin impossible. In this way a HPLC method is only able to separate phloretin and product, but makes the measurement of UDP-glucose impossible. Only long wash steps of more than 100 CV with water were able to elute UDP-glucose fully from C18 material.

In order to minimize cross contaminations, we switched towards a direct injection (DI) approach. Quantification of ratios for UDP-glucose, nothofagin, phlorizin and their labeled derivates was performed using the LipidDataAnalyzer tool developed by the Institute of Computational Biotechnology, NAWI Graz, Graz University of Technology [11] which evaluates spectra integrity and performs automated deisotoping. The tool was used as described in the methods section and no specific adjustments to the program settings were performed.

Table 3 shows the initial KIE values that were obtained. The error on each KIE is quite large and the KIE's for substrate or product are not matching. From equation 4-13 it is obvious that R_0 has besides R_S and R_P the greatest influence on the calculated values. It can be assumed that f is obtained in a reasonable quality as it is directly accessible by HPLC-UV. For this reason we implemented a method to correct R_0 and compared the obtained KIE's.

Table 3. Calculated KIE based on whole molecule mass spectroscopy. Given are uncorrected KIE's that were each calculated from data obtained for the enrichment in the substrate and in the formed product.

		*Os*CGT		*Pc*OGT	
		- CIP	+ CIP	- CIP	+ CIP
1D	UDPG based	1.463 ± 0.318	1.517 ± 0.344	1.808 ± 0.928	1.569 ± 0.720
	Product based	0.762 ± 0.125	0.766 ± 0.117	0.579 ± 0.116	0.559 ± 0.188
6,6D	UDPG based	0.852 ± 0.100	0.859 ± 0.127	0.852 ± 0.124	0.864 ± 0.119
	Product based	0.902 ± 0.090	0.912 ± 0.052	0.724 ± 0.053	0.731 ± 0.077

KIE correction for initial substrate ratio errors by fitting of R_0

The iteration of R_0 improved the consistency between product and substrate based KIE calculation dramatically in case of the formation of the C-glycoside nothofagin. As shown in table 4 only a small differences in case of the reaction leading to the C-glycoside nothofagin were observed. The KIE for 1D and 6,6D probes are almost identical and the addition of phosphatase also had no effect on the KIE. Still we observed a difference between substrate and product based determination which was greater than the average error.

The O-glycoside phlorizin forming reaction on the other hand showed still a large discrepancy between substrate and product based KIE.

Table 4. Calculated KIE based on whole molecule mass spectroscopy. The given KIE's were each corrected by a fitted R_0 value in the calculation and derived from data obtained for the enrichment in the substrate and in the formed product and.

		*Os*CGT		*Pc*OGT	
		- CIP	+ CIP	- CIP	+ CIP
1D	UDPG based	0.961 ± 0.057	0.985 ± 0.060	1.017 ± 0.018	1.013 ± 0.024
	Product based	0.927 ± 0.059	0.922 ± 0.076	0.695 ± 0.077	0.689 ± 0.118
6,6D	UDPG based	0.986 ± 0.020	0.986 ± 0.034	0.976 ± 0.014	0.968 ± 0.020
	Product based	0.837 ± 0.112	0.849 ± 0.064	0.674 ± 0.058	0.675 ± 0.089

As it is reported in literature the ionization itself can have an intrinsic isotope effect[15a, b]. To investigate possible isotope dependent ionization effects R_{ges} was plotted against f (see figure 4). As we show the resulting enrichment in the product is not matching the initial R_0. The significant increase of R_P towards the completed reaction indicates a better ionization of the unlabeled product. The substrate on the other hand shows almost no isotope dependent ionization. Furthermore it can be observed that the labeling at position C-1 in an O-glycoside has the strongest effect whereas C-glycosides are affected almost identical by the heavier isotope in position C-1 or C-6 of the glycoside. This is reasonable as the fragmentation of the product occurs normally within the glycosidic ring system for the C-glycoside, whereas the fragmentation of phlorizin would occur within the O-glycosidic bond. And is therefore possibly dependent on a label in position 1.
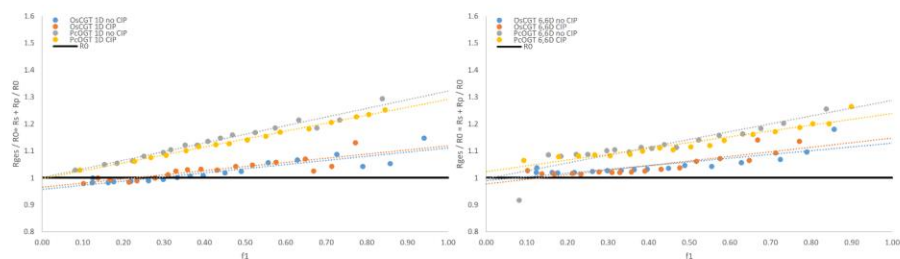
Figure 6. Determination of isotope dependent ionization. R_{ges} is plotted agains f and correction factors for R_P and R_S are calculated from the resulting trend line.

Correction of isotope dependent ionization

As shown in 4 R_{ges} is not equal to R_0 over the course of the reaction and a isotope dependent ionization can be assumed. We corrected R_P and R_S as described in the method section and R_{ges} became almost identical to R_0 (see fig. 5) and constant over the course of the reaction. The corrected R_P and R_S values were then used to recalculate the observed KIE and indicate clearly that the correction improved the product based calculations especially for the O-glycoside (see table 5). Still all KIE's are close to 1, although the KIE observed for the OsCGT is slightly lower.
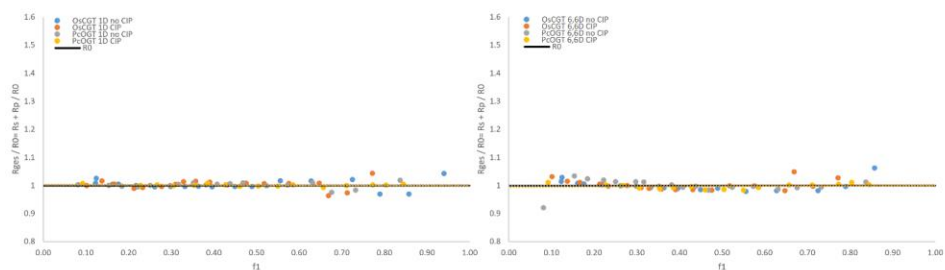

Figure 7. Determination of isotope dependent ionization. After correction of isotope dependent ionization the sum of labled product and substrate match the expected amount which is defined by the initial R_0.

Table 5. Calculated KIE based on whole molecule mass spectroscopy. The KIE are each calculated from data obtained for the enrichment in the substrate and in the formed product and further corrected by a fitted R_0 value in the calculation, additionally isotope dependent ionization was also corrected.

| | | *Os*CGT | | *Pc*OGT | |
		- CIP	+ CIP	- CIP	+ CIP
1D	UDPG based	1.085 ± 0.057	1.087 ± 0.048	1.014 ± 0.016	1.020 ± 0.031
	Product based	1.082 ± 0.034	1.078 ± 0.048	1.010 ± 0.025	1.014 ± 0.013
6,6D	UDPG based	1.017 ± 0.041	1.052 ± 0.057	0.984 ± 0.015	0.908 ± 0.073
	Product based	1.000 ± 0.070	1.030 ± 0.041	0.959 ± 0.067	0.923 ± 0.082

Forward commitment to catalysis

Because observed KIEs on V_{max}/K_M $^L(V/K)$ include contributions from all isotopically sensitive steps preceding the irreversible step, i.e. the glucose transfer, we determined the forward commitment by an ion trapping experiment.
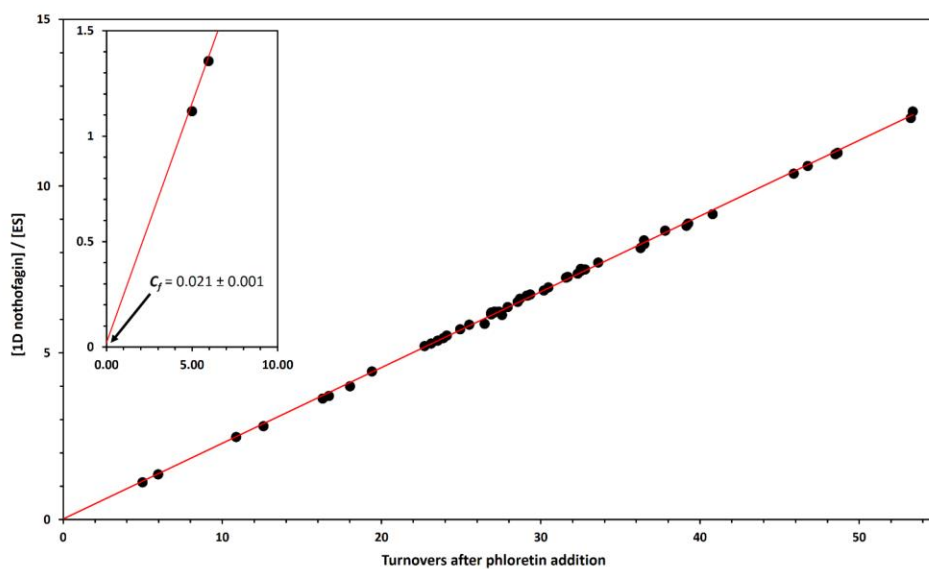

Figure 8. Determination of intrinsic KIEs and correction for forward commitment factor (C_f). Measurement of C_f by isotope trapping for the glycosylation of phloretin by *Os*CGT. Insert shows intersection of c_f trendline with y-axis

As shown no significant enrichment of labelled product can be confirmed during the first turnovers, which leads to the conclusion that C_f is close to 0 and the intrinsic KIE can therefore be calculated by Equation 7-19 where $^L(V/K)$ and k the observed KIE which has to be corrected by C_f. The correction lead to only minor changes in de determined KIE's. After correcting for all possible effects and eliminating these contributions, it can be shown that the KIE is close to 1 in all cases

Table 6. Calculated KIE based on whole molecule mass spectroscopy. The KIE are each calculated from data obtained for the enrichment in the substrate and in the formed product and further corrected by the measured C_f value.

		*Os*CGT		*Pc*OGT	
		- CIP	+ CIP	- CIP	+ CIP
1D	UDPG based	1.087 ± 0.057	1.089 ± 0.048	1.014 ± 0.016	1.020 ± 0.031
	Product based	1.084 ± 0.034	1.080 ± 0.048	1.010 ± 0.025	1.014 ± 0.013
6,6D	UDPG based	1.017 ± 0.041	1.053 ± 0.057	0.984 ± 0.015	0.906 ± 0.073
	Product based	1.000 ± 0.070	1.031 ± 0.041	0.958 ± 0.067	0.921 ± 0.082

As we show R_S/R_0 shows an enrichment of the heavier isomer over the course of the reaction whereas R_P/R_0 is approaching asymptotically the value 1.
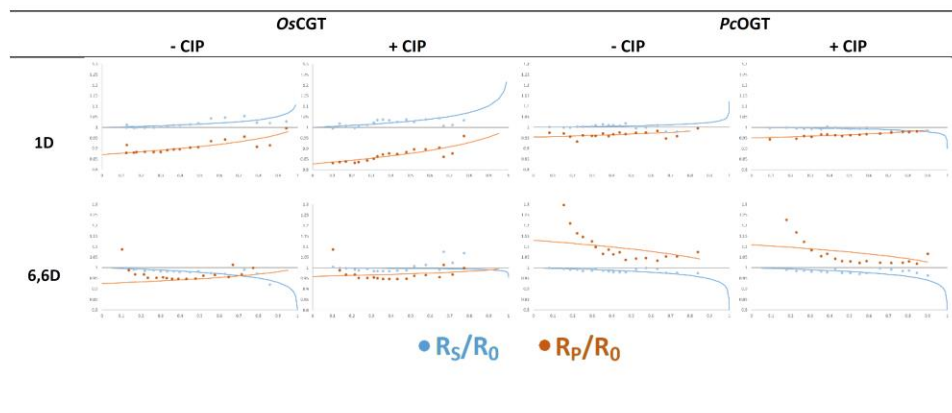

Figure 9. Fitted data for calculated KIEs. Each plot depicts the measured isotopic ratios for product (orange) and substrate (blue). Additionally the fitted curve for the calculated KIEs are depicted.

KIE values for *Os*CGT and *Pc*OGT were independently calculated for UDP-glucose and nothofagin or phlorizin, respectively, by fitting the data as described earlier [16]. As we can show the collected data is not always matching with the predicted fit. While especially the reactions with [1-^{2}H] labeled UDP-glucose allow a good fitting, the data from remotely labeled 6,6D UDP-glucose resulted in values which were not matching with expected behavior although the fitting still allowed the calculation of KIE values. These values were close to 1 which could explain the inversion of some plots when the KIE is lower than 1, which easily happens when the error on the KIE is larger than the distance to 1 and values lower than 1 lead to an inversion of the curvature as shown for *Pc*OGT in presence of 6,6D UDP-glucose. Although a KIE can be calculated for the products of 6,6D UDP-glucose and the correction of the isotopic ratio by R_0 reiteration and correction of isotope dependent ionization, we can show that the data is not matching with a fit which is derived from the calculated KIE (see fig.9). We found that the values are almost identical for *Os*CGT and *Pc*OGT (see table 6). In comparison the remotely labeled 6,6D UDP-glucose resulted in nearly identical enrichments as 1DUDP-glucose. In case of the *Pc*OGT it seems that the remote label at position 6 shows a negative isotope effect, which is unexpected.

Another possible effect negatively influencing the quality of KIE measurements on Vmax/Km are reverse reactions from the product towards UDP-glucose. While nothofagin is not hydrolysable by *Os*CGT, *Pc*OGT is prone to catalyze the reverse reaction at low pH values. To remove these adverse effects, alkaline phosphatase was added to control reactions to remove any released UDP, which would be a possible partner for formation of UDP-glucose. As we found alkaline phosphatase is neither increasing the reaction rate by shifting the equilibrium towards the formation of product nor does it affect the observed KIE values for *Os*CGT or *Pc*OGT beyond the error margin of the method itself. We therefore assume that UDP under the described conditions cannot cause a severe reverse reaction. The values for the intrinsic KIE accounting for the determined C_f value and the neglectable reverse reaction are given in table 6. In addition reverse reaction rate was determined and reported elsewhere [17].

In total the data present a comprehensive outlook for the implementation of a mass spectroscopy based determination of KIEs on glycosyltransferases.

Discussion

In the herein presented work we concluded that NMR determination of α-secondary KIEs were not possible to measure due to lacking capability to distinguish between labeled and unlabeled reaction partners. Our approach was dependent on three signals of the formed product. (1) Signals on the B ring for quantification of total abundance of the acceptor and product, (2) proton signal at position C-5 for determination of conversion progress and (3) the proton signal at the anomeric carbon C-1'' to determine the ratio between labelled and unlabeled product. As shown in figure 5 only signals (1) and (2) were reliable measured, which leaves us a method which can measure the overall conversion but contains no information about isotopic enrichments of the glucosyl moiety. This problem could possibly be solved by implementing an improved method based on HSQC identification and deconvolution of the individual signal. This would allow to measure to measure also the ration of isotopic labelling [18].

As a sensitive measurement of a kinetic isotope effect would have to make use of data obtained for the anomeric proton, we decided to switch to a mass spectroscopy-based approach. We found that the method provided data with a respective error margin that was low enough to actually determine if both enzymes perform the reaction in a similar manner (see table 7). The method using a direct injection approach is still not universally applicable due to the used detection by Orbitrap which is prone to limited quantification reliability in presence of to many ions. Nevertheless detection of product at the beginning of the reaction and detection of UDP-glucose at the end of the reaction require the injection of large Ion bundles with low abundance of the respective analyte due to unfavorable ratios. A possible improvement would be a HPLC based pre-separation which was not achieved due to cross-contaminations of residual nucleosides. Another solution for further improvement would be the use of a triple quad MS due to its better dynamic range. Also the preseparating C-trap of the Orbitrap could be used to enrich the observed analyte for better quantification. Unfortunately this approach was undermined by the capabilities of the used LDA-software package which is handling only single time resolved scanning windows and was not handling data properly with two scanning windows for UDP-glucose and glycosylated product each.

Literature

1. Gloster, T. M. (2014). Advances in understanding glycosyltransferases from a structural perspective. *Current opinion in structural biology*, *28*, 131-141.

2. Lee, S. S., Hong, S. Y., Errey, J. C., Izumi, A., Davies, G. J., & Davis, B. G. (2011). Mechanistic evidence for a front-side, S N i-type reaction in a retaining glycosyltransferase. *Nature chemical biology*, *7*(9), 631.

3. a) Kim, S. C., Singh, A. N., & Raushel, F. M. (1988). Analysis of the galactosyltransferase reaction by positional isotope exchange and secondary deuterium isotope effects. *Archives of biochemistry and biophysics*, *267*(1), 54-59. b) Bruner, M., & Horenstein, B. A. (1998). Isotope trapping and kinetic isotope effect studies of rat liver α-(2→ 6)-sialyltransferase. *Biochemistry*, *37*(1), 289-297.

4. a) Kren, V., & Martínková, L. (2001). Glycosides in medicine:"The role of glycosidic residue in biological activity". *Current medicinal chemistry*, *8*(11), 1303-1328. b) Gantt, R. W., Peltier-Pain, P., Singh, S., Zhou, M., & Thorson, J. S. (2013). Broadening the scope of glycosyltransferase-catalyzed sugar nucleotide synthesis. *Proceedings of the National Academy of Sciences*, 201220220. c) Lepak, A., Gutmann, A., Kulmer, S. T., & Nidetzky, B. (2015). Creating a Water-Soluble Resveratrol-Based Antioxidant by Site-Selective Enzymatic Glucosylation. *ChemBioChem*, *16*(13), 1870-1874.

5. Xie, K., Chen, R., Chen, D., Li, J., Wang, R., Yang, L., & Dai, J. (2017). Enzymatic N-Glycosylation of Diverse Arylamine Aglycones by a Promiscuous Glycosyltransferase from Carthamus tinctorius. *Advanced Synthesis & Catalysis*, *359*(4), 603-608.

6. Gutmann, A., & Nidetzky, B. (2012). Switching between O-and C-glycosyltransferase through exchange of active-site motifs. *Angewandte Chemie International Edition*, *51*(51), 12879-12883.

7. Cleland, W. W. (2005). The use of isotope effects to determine enzyme mechanisms. *Archives of biochemistry and biophysics*, *433*(1), 2-12.

8. Gutmann, A., & Nidetzky, B. (2012). Switching between O-and C-glycosyltransferase through exchange of active-site motifs. *Angewandte Chemie International Edition*, *51*(51), 12879-12883.

9. Eixelsberger, T., Horvat, D., Gutmann, A., Weber, H., & Nidetzky, B. (2017). Isotope Probing of the UDP-Apiose/UDP-Xylose Synthase Reaction: Evidence of a Mechanism via a Coupled Oxidation and Aldol Cleavage. *Angewandte Chemie International Edition*, *56*(9), 2503-2507.

10.Gutmann, A., Lepak, A., Diricks, M., Desmet, T., & Nidetzky, B. (2017). Glycosyltransferase cascades for natural product glycosylation: Use of plant instead of bacterial sucrose synthases improves the UDP-glucose recycling from sucrose and UDP. *Biotechnology journal*, *12*(7), 1600557.

11. Hartler, J., Trötzmüller, M., Chitraju, C., Spener, F., Köfeler, H. C., & Thallinger, G. G. (2010). Lipid Data Analyzer: unattended identification and quantitation of lipids in LC-MS data. *Bioinformatics*, *27*(4), 572-577.

12. a) Dahlquist, F. W., Rand-Meir, T., & Raftery, M. A. (1969). Application of secondary. alpha.-deuterium kinetic isotope effects to studies of enzyme catalysis. Glycoside hydrolysis by lysozyme and. beta.-glucosidase. *Biochemistry*, *8*(10), 4214-4221. b) Cook, P. F., & Cleland, W. W. (2007). *Enzyme kinetics and mechanism*. Garland Science.

13. Poulin, M. B., Schneck, J. L., Matico, R. E., McDevitt, P. J., Huddleston, M. J., Hou, W., ... & Schramm, V. L. (2016). Transition state for the NSD2-catalyzed methylation of histone H3 lysine 36. *Proceedings of the National Academy of Sciences*, *113*(5), 1197-1201.

14. Rose, I. A. (1980). The isotope trapping method: Desorption rates of productive E· S complexes. In *Methods in enzymology* (Vol. 64, pp. 47-59). Academic Press.

15. a) Vestal, M., & Futrell, J. H. (1970). Metastable ions and isotope effects in the mass spectra of propane and deuterated propanes. *The Journal of Chemical Physics*, *52*(2), 978-988. b) Zakett, D., Flynn, R. G. A., & Cooks, R. G. (1978). Chlorine isotope effects in mass spectrometry by multiple reaction monitoring. *The Journal of Physical Chemistry*, *82*(22), 2359-2362.

16. Dahlquist, F. W., Rand-Meir, T., & Raftery, M. A. (1969). Application of secondary. alpha.-deuterium kinetic isotope effects to studies of enzyme catalysis. Glycoside hydrolysis by lysozyme and. beta.-glucosidase. *Biochemistry*, *8*(10), 4214-4221.

17. Lepak, A., Gutmann, A., & Nidetzky, B. (2018). β-Glucosyl Fluoride as Reverse Reaction Donor Substrate and Mechanistic Probe of Inverting Sugar Nucleotide-Dependent Glycosyltransferases. *ACS Catalysis*, *8*(10), 9148-9153.

18. Manning, K. A., Sathyamoorthy, B., Eletsky, A., Szyperski, T., & Murkin, A. S. (2012). Highly precise measurement of kinetic isotope effects using 1H-detected 2D [13C, 1H]-HSQC NMR spectroscopy. *Journal of the American Chemical Society*, *134*(51), 20589-20592.