Two-step *O*- to *C*-glycosidic bond rearrangement using complementary glycosyltransferase activities

Alexander Gutmann, Corinna Krump, Linda Bungaruang and Bernd Nidetzky

Institute of Biotechnology and Biochemical Engineering,
Graz University of Technology, Petersgasse 12/1, A-8010 Graz, Austria

Electronic Supplementary Information

Table of contents

1 Methods

1.1 Chemicals and reagents

Unless otherwise noted, all chemicals were obtained from Sigma-Aldrich (Vienna, Austria) in the highest purity available. Phlorizin dihydrate (≥98%) was purchased from Carl Roth (Karlsruhe, Germany). DNA modifying enzymes were from Thermo Scientific (Waltham, MA, US) and PCR primers were obtained from Life Technologies (Carlsbad, CA, US). Phusion® High-Fidelity DNA Polymerase was purchased from New England Biolabs (Ipswich, MA, US). *Strep*-Tactin® Sepharose® and desthiobiotin were obtained from IBA (Goettingen, Germany).

1.2 Strain construction

The *Os*CGT gene (GenBank: FM179712) was kindly provided from the group of Prof. Robert Edwards (Centre for Bioactive Chemistry, Durham University, UK) in a pET-STRP3 vector which is a custom made derivative of pET-24d that enables protein expression with an N-terminally fused *Strep*-tag II.[1-2] Introduction of the mutation I121D was described elsewhere.[3] The *Pc*OGT gene (UGT88F2; GenBank: FJ854496) was a kind gift from the group of Prof. Karl Stich (Institute of Chemical Engineering, Vienna University of Technology, Austria).[4] Flanking restriction sites for *Nde*I and *Xho*I were added and the respective internal restriction sites were removed by overlap extension PCR as described previously.[3] Subsequently the gene was inserted into *Nde*I and *Xho*I sites of a pET-28a vector for expression with N-terminally fused His-tag. Expression strains were created by transformation of electro-competent *E. coli* BL21-Gold (DE3) cells with the described plasmids.
Correct sequences of the complete genes and the regions around both cloning sites of the pET-Strep3 plasmid were verified by sequencing. Vector maps and complete DNA sequences of *Os*CGT and *Pc*OGT as well as the pET-Strep3 plasmid can be found in the Appendix.

1.3 Enzyme expression and purification

Cultivation of *E. coli* cells for protein expression was described elsewhere in detail.[3] Cells were grown at 37°C and 120 rpm in 1 L baffled shake flasks containing 300 mL LB-media (50 µg mL^{-1} kanamycin). At an optical density at 600 nm of 0.8-1.0, protein expression was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). After overnight expression at 25°C cells were harvested by 30 min centrifugation at 5,000 rpm and 4°C, resuspended in water and stored at -70°C until disruption by repeated passage through a cooled French press at 100 bar. Before purification cell debris was removed by centrifugation (45 min, 13,200 rpm, 4°C) and cell extract was filtrated through a 1.2 µm cellulose-acetate filter.
*Os*CGT and *Os*CGT_I121D were purified by *Strep*-tag affinity chromatography on 3 mL gravity flow *Strep*-Tactin® Sepharose® columns at 4°C as previously described.[3] After loading cell extract the column was washed with 5 column volumes (CV) of washing buffer (100 mM Tris pH 8.0, 150 mM NaCl, 1 mM EDTA). Enzymes were eluted with 3 CVs of elution buffer (100 mM Tris, pH 8.0, 150 mM NaCl, 1 mM EDTA, 2.5 mM desthiobiotin) whereas the first 0.5 CVs were discarded, and the rest was pooled.
*Pc*OGT was purified at 4°C on a 5 mL HiTrap™ Chelating FF column (GE Healthcare) loaded with Ni^{2+}. A flow rate of around 5 $mLmin^{-1}$ was applied manually using a syringe. Cell extract was loaded on the column after equilibration with 5 CVs of buffer W (20 mM Tris pH 7.4, 500 mM NaCl, 20 mM imidazole). After washing with further 5 CVs of buffer W *Pc*OGT was eluted in two steps using 4 CVs buffer E1 (20 mM Tris pH 7.4, 500 mM NaCl, 250 mM imidazole) and 5 CVs of buffer E2 (20 mM Tris pH 7.4, 500 mM NaCl, 500 mM imidazole), respectively.
Fractions containing purified proteins were concentrated and buffer was exchanged to 25 mM HEPES, pH 7.0, using centrifugal concentrators with a Molecular Weight Cut Off of 10,000. Purified enzymes were stored as small aliquots at -20°C to avoid multiple freeze-thaw cycles. Protein concentrations of purified enzymes were determined photometrically on a NanoDrop 2000 system (Thermo Scientific). Molecular weight and molar extinction coefficients were calculated using Peptide Properties Calculator. Enzyme purities were estimated by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) using Coomassie Blue staining.

1.4 Quantification of dihydrochalcones – HPLC based activity assay

Quantification of phloretin and its glycosides was achieved by an HPLC based assay as described elsewhere in detail.[3] It was applied to follow all enzymatic conversions and to determine activities of *Pc*OGT and *Os*CGT (I121D). Briefly, reaction mixtures contained if not mentioned otherwise 0.6 mM UDP-glucose, 1 mM phloretin, 25 mM Tris pH 7, 13 mM $MgCl_2$, 50 mM KCl, 0.13% BSA and 20% DMSO. Conversions were started by glycosyltransferase addition. After certain incubation times at 30°C samples were withdrawn and enzymatic

conversions were stopped by mixing with an equal volume of acetonitrile. Precipitated protein was removed by centrifugation for 15 min at 13,200 rpm. Typically 10 µL of supernatant were applied on an Agilent 1200 HPLC system equipped with a Chromolith® Performance RP-18e endcapped column (100–4.6 mm) from Merck. Separation of phloretin and its glycosides was monitored by UV detection at 288 nm under thermostatic control at 35°C by following method. Solvent A: water with 0.1% TFA; Solvent B: acetonitrile with 0.1% TFA; Gradient: 7.5 min: 20-47.5% B (1 mL min^{-1}), 0.05 min: 47.5-100% B (1 mL min^{-1}), 1.45 min: 100% B (1.5 mL min^{-1}), 0.05 min: 100-20% B (1.5 mL min^{-1}), 2.45 min: 20% B (1.5 mL min^{-1})
One unit of *Os*CGT (I121D) or *Pc*OGT activity was defined as the amount of enzyme glucosylating 1 µmol phloretin per minute under the following conditions: 0.6 mM UDP-glucose, 1 mM (*Os*CGT/*Pc*OGT) or 0.1 (OsCGT_I121D) phloretin, 25 mM Tris (*Os*CGT, *Pc*OGT) or BisTris (OsCGT_I121D), pH 7, 13 mM $MgCl_2$, 50 mM KCl, 0.13% BSA and 20% DMSO at 30°C. At least 4 distinct measurements were used to calculate linear initial rates.

1.5 Quantification of UDP and UDP-glucose - capillary zone electrophoresis

Concentrations of UDP and UDP-glucose were determined by capillary zone electrophoresis. Samples were prepared as for HPLC measurements by mixing with an equal volume of acetonitrile to stop enzymatic conversion. Precipitated protein was removed by centrifugation (15 min; 13,200 rpm). Capillary zone electrophoresis was performed using a 3D capillary electrophoresis system (Hewlett Packard) equipped with a fused-silica capillary (56 cm × 50 µm) with an extended light path. The capillary was preconditioned by a 6 min flush with background electrolyte (20 mM sodium borate, pH 9.3). Samples were loaded by pressure injection (50 mbar, 10 s) and compounds were resolved at 22 kV over 22 min at 18°C. UDP and UDP-glucose were monitored on a diode array detector at 262 nm.

1.6 Characterization of *Pc*OGT reaction reversibility

To test if phloretin formation from phlorizin is a result of reverse glycosylation by *Pc*OGT, 1 mM phlorizin was incubated at 30°C in 50 mM BisTris, pH 7.0, 13 mM $MgCl_2$, 50 mM KCl, 0.13% BSA and 20% DMSO. Two mM UDP, 25 mU mL^{-1} *Pc*OGT, or both were added. Conversion of phlorizin was followed by HPLC.
To obtain pH-profiles of *Pc*OGT in glycosylation and deglycosylation direction, the standard protocol of HPLC based activity assay was modified in following manner. A set of reaction buffers was prepared from pH 3 to 10 in steps of 0.5 (25 mM citrate pH 3-7; 25 mM Tris pH 7-9.5, 25 mM CAPS pH 9.5-10). All reactions contained 13 mM $MgCl_2$, 50 mM KCl, 0.13% BSA and 20% DMSO. In glycosylation direction 0.1 mM phloretin and 0.6 mM UDP-glucose were used as substrates and reactions were started by addition of following amounts of *Pc*OGT: 60 ng mL^{-1} (pH 5-5.5); 30 ng mL^{-1} (pH 6-6.5, 9.5-10); 15 ng mL^{-1} (pH 7.0-9.0); Deglycosylation reactions (pH 3-9.5) contained 1 mM phlorizin and 2 mM UDP and were started by addition of 60 ng mL^{-1} *Pc*OGT. All reactions were incubated at 30°C and samples were withdrawn every 20 min for 1h. The actual pH of reaction mixtures was determined as the average of pH measurements at the beginning and at the end of the observed time span.
To reach equilibrium, another 8.3 µg mL^{-1} *Pc*OGT were added to reactions in deglycosylation direction after samples for initial rate measurements were withdrawn. After overnight incubation at 30°C phloretin and phlorizin concentrations were experimentally determined by HPLC. The UDP-glucose concentration was inferred from reaction stoichiometry. Equilibrium constants were calculated in glycosylation direction according to equation 1.

$$K_{eq} = \frac{c_{phlorizin} \cdot c_{UDP}}{c_{phloretin} \cdot c_{UDP-glucose}} \quad (1)$$

1.7 Potentiometric titration of phlorizin, phloretin, UDP and UDP-glucose

pK_a-values of all compounds involved in phlorizin conversion by *Pc*OGT (phlorizin, phloretin, UDP, UDP-glucose) were determined by potentiometric titration. Of all analytes, solutions of 7.5 mM were made in decarbonated water containing 20% DMSO, 13 mM $MgCl_2$ and 50 mM KCl to reproduce rearrangement conditions. Analyte volumes of 10 (UDP, UDP-glucose) or 20 mL (phlorizin, phloretin) were titrated at 30°C with 150 mM NaOH as titrant using a SenTix® Mic electrode from WTW (Weilheim, Germany) for pH monitoring.

1.8 *O*- to *C*-glucoside rearrangement by *Os*CGT and *Pc*OGT

Unless mentioned otherwise reaction mixtures contained 2 mM UDP and 5 mM phlorizin in 50 mM Tris pH 7 containing 13 mM $MgCl_2$, 50 mM KCl, 0.13% BSA and 20% DMSO. Conversions were started by addition of 100 mU mL^{-1} *Pc*OGT and 50 mU mL^{-1} *Os*CGT. All reactions were incubated at 30°C. Aliquots were withdrawn, stopped by acetonitrile addition and analysed by RP-HPLC as described.
To test the influence of cosolvents on *O*- to *C*-glucoside rearrangement conversions were carried out in presence of different concentrations (5, 10, 15 or 20%) of either ethanol or DMSO.
The effect of pH on the rearrangement was studied by replacing the standard buffer with 50 mM HEPES buffers ranging from pH 6 to 8.5 (in steps of 0.5). The actual pH of reaction mixtures was determined as the average of pH measurements at beginning and end of 5 h long conversions at 30°C.
The impact of substrate concentrations on *O*- to *C*-glycosidic bond rearrangement was studied by varying phlorizin (0.05 - 10 mM; 2 mM UDP) and UDP (0.005 - 1.2 mM; 5 mM phlorizin; 5% DMSO). Linear initial phlorizin consumption and nothofagin production rates were determined during initial 25 (phlorizin variation) or 60 min (UDP variation).

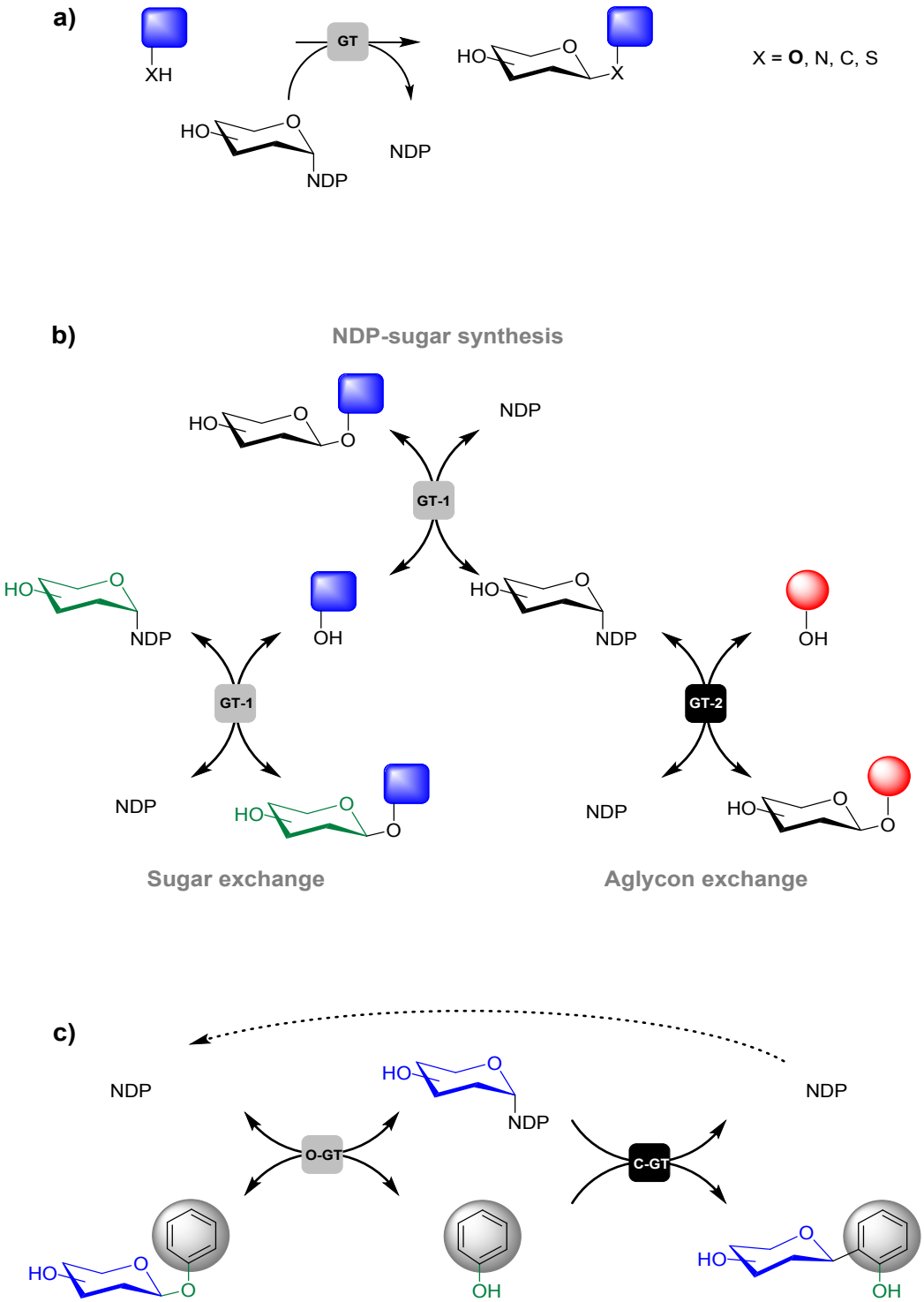
1.9 *O*- to *C*-glucoside rearrangement by *Os*CGT_I121D

Unless mentioned otherwise all conversions were made with 1.25 mU mL^{-1} *Os*CGT_I121D in 50 mM BisTris buffer, pH 7.0, containing 13 mM $MgCl_2$, 50 mM KCl, 0.13% BSA and 20% DMSO. During 24-70 h long time courses at least 7 samples were analysed by RP-HPLC as described previously.

For optimisation of substrate concentrations, 0.05, 0.5 and 5 mM phlorizin were tested in combination with 0.5 and 5 mM UDP. Evaluation of pH dependency was done using 50 mM BisTris buffers of pH 6.0, 7.0 and 8.0 for conversion of 0.5 mM phlorizin with 5 mM UDP. The effect of UDP-glucose addition (0, 0.5, 1, 5 mM) was also investigated using 0.5 mM phlorizin and 5 mM UDP. Conversion of 0.5 mM phlorizin under optimized conditions was performed with 2.5 mU mL^{-1} *Os*CGT_I121D in presence of 5 mM UDP and 1 mM UDP-glucose.

2. Results

2.1 Synthetic use of reverse glycosyltransferase reactions



Scheme S1 Natural product glycosylations through a canonical glycosyltransferase reaction (a), two-step exchange processes exploiting a suitable glycosyltransferase reaction in reverse direction (b), and *O*- to *C*-glycosidic bond rearrangement catalysed by complementary glycosyltransferase activities (c)

2.2 Determination of enzyme purity by SDS-PAGE

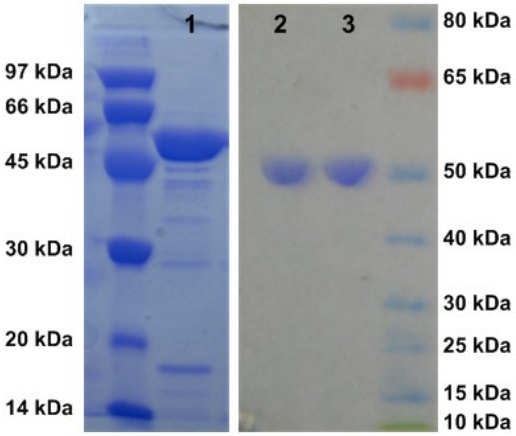


Fig. S1 SDS-PAGE of enzymes purified by affinity chromatography: Lane 1: His-tagged *Pc*OGT (55.7 kDa) purified on Ni Sepharose™; Lane 2 and 3: *Strep*-tagged *Os*CGT and *Os*CGT_I121D, respectively purified by *Strep*-tag affinity chromatography (51.3 kDa each)

2.3 Reverse glycosylation of phlorizin by *Pc*OGT

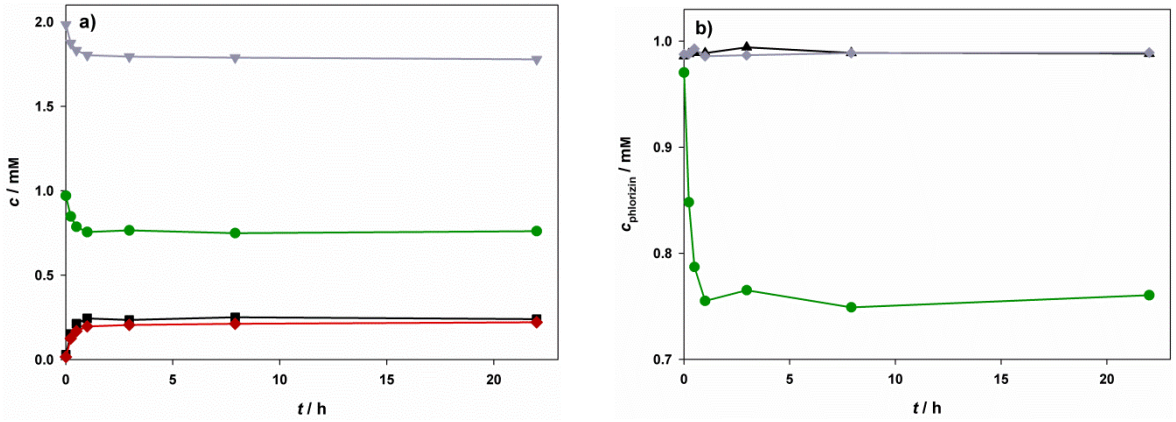


Fig. S2 Deglycosylation of phlorizin by *Pc*OGT at pH 7.0: a) 1 mM phlorizin (green) was partially deglycosylated in presence of 2 mM UDP (grey) and *Pc*OGT (25 mU mL^{-1}). Equal amounts of phloretin (black) and UDP-glucose (red) were formed. b) Phlorizin was only deglycosylated when UDP as well as *Pc*OGT were present (green). In presence of only UDP (grey) or solely *Pc*OGT (black) phlorizin was not converted. Deglycosylation of phlorizin is therefore solely achieved by enzyme-catalysed transfer of glucose to UDP and not by a potential hydrolytic side reaction.

2.4 Effect of organic cosolvents on *O*- to *C*-glucoside rearrangement

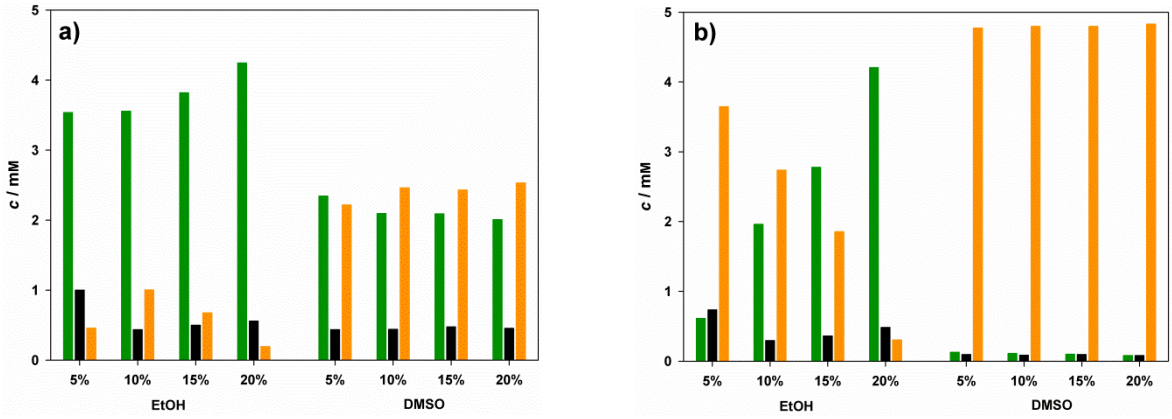


Fig. S3 Yields of *O*- to *C*-glucosidic bond rearrangement at various cosolvent conditions after (a) 2h and (b) 24h of conversion (5 mM phlorizin, 2 mM UDP, 100 mU mL^{-1} *Pc*OGT, 50 mU mL^{-1} *Os*CGT). Phlorizin (green), phloretin (black), nothofagin (orange); Any concentration of ethanol reduced conversions drastically. In contrast even the highest applied DMSO concentration of 20% showed no significant reduction of initial or final conversion.

2.5 *O*- to *C*-glucoside rearrangement in the absence of side reactions

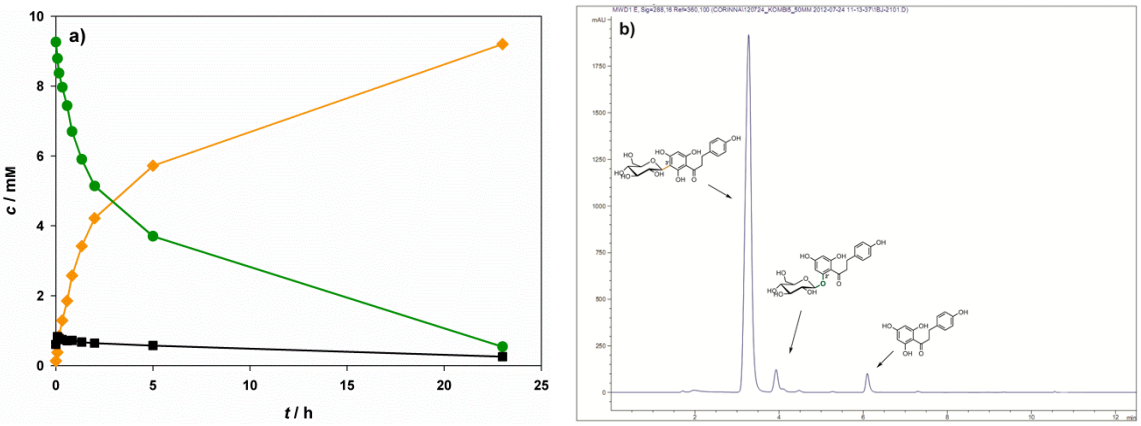


Fig. S4 10 mM phlorizin were almost completely converted to the *C*-glucoside nothofagin within 24 h (10 mM phlorizin, 2 mM UDP, 100 mU mL^{-1} *Pc*OGT, 50 mU mL^{-1} *Os*CGT). a) Time course of rearrangement: Quick initial consumption of phlorizin (green) gradually slows down after around one hour. The aglycon phloretin (black) remains relatively constant over the entire conversion. After 24 h of incubation 92% (9.2 mM) of applied phlorizin are converted to nothofagin (orange). b) HPLC analysis after 24h of conversion: The UV-trace (288 nm) shows a dominant nothofagin signal (3.3 min) with minor peaks for phlorizin (3.9 min) and phloretin (6.1 min). Absence of any further signals indicates that no side products were formed.

2.6 Potentiometric titration of phlorizin, phloretin, UDP and UDP-glucose

To investigate the cause for pH dependency of the equilibrium constant in *O*-glycosylation, potentiometric titrations of phloretin and its 4'-*O*-glucoside phlorizin (Fig. S5a) as well as UDP and UDP-glucose (Fig. S5b) were performed. Observed pK_a values were in general agreement with literature data, which allowed their assignment to functional groups (Fig. S5c).[5-6] Slight discrepancies between observed and reported pK_a values are probably caused by differences in experimental conditions (addition of DMSO, temperature, ionic strength).
The titration curve of the aglycon phloretin (Fig. S5a, black) was in good agreement with literature, the reported pK_a values being 7.0, 9.4, and 10.5.[5] Phlorizin shares with phloretin a pK_a value in the neutral pH range. This pK_a supposedly represents the first deprotonation event on one of the three phenolic hydroxyl groups on the aromatic ring involved in glycosylation. The second pK_a value of phloretin of around pH 9.3 is missing in phlorizin (green). This indicates that the pK_a represents a second deprotonation event on the same ring which is missing in phlorizin because of glucose bound to the oxygen at position 2'. A further pronounced plateau at around pH 10.8 is visible in both compounds. It most likely reflects unresolved pK_a values from a final deprotonation on the glycosylated aromatic ring and also deprotonation of the single hydroxyl groups on the second aromatic ring. The observed large increase of the equilibrium constant of *O*-glycosylation with increased pH in the range from 6.5 to 8.8 cannot be explained by deprotonation of phloretin upon glycosylation. Phloretin glycosylation causes no significant proton transfer below pH 8. At higher pH, proton release would rather occur in deglycosylation than in glycosylation direction due to increase of acidity in going from phlorizin to phloretin.
Protonation UDP and UDP-glucose can, however, explain a drastic preference of glycosylation at high pH. UDP (Fig. S5b, grey) displays a pK_a value of around 5.6, which is missing in UDP-glucose. It represents the second deprotonation of the terminal phosphate group, which is glycosylated in the case of UDP-glucose. Fig. S5b reveals that UDP is predominantly deprotonated throughout the relevant pH range (pH 6.5-8.8) where the reaction equilibrium constant exhibited pronounced pH-dependency (Fig. 1b). Sugar donor activation by a substituted phosphate-leaving group is a common feature of glycosyltransferases.[7] Therefore, directional preference for deglycosylation at low pH might be expected as a rather general feature of glycosyltransferase reactions, except in cases where the reactive group of the acceptor substrate matches the low pK_a of the sugar donor's (secondary) phosphate group. The vast majority of hydroxyl acceptor groups have comparably high pK_a and are therefore only deprotonated at high pH. Only a few studies have been performed in which pH effects on glycosyltransferase reaction equilibrium have been examined in detail. However, common trend is that reverse glycosylation is preferred at low pH (equilibrium constants and reaction rates).[8-11] Net proton release during forward glycosyl transfer from NDP-sugar was also exploited for glycosyltransferase assay development based on pH-shift.[12-13]

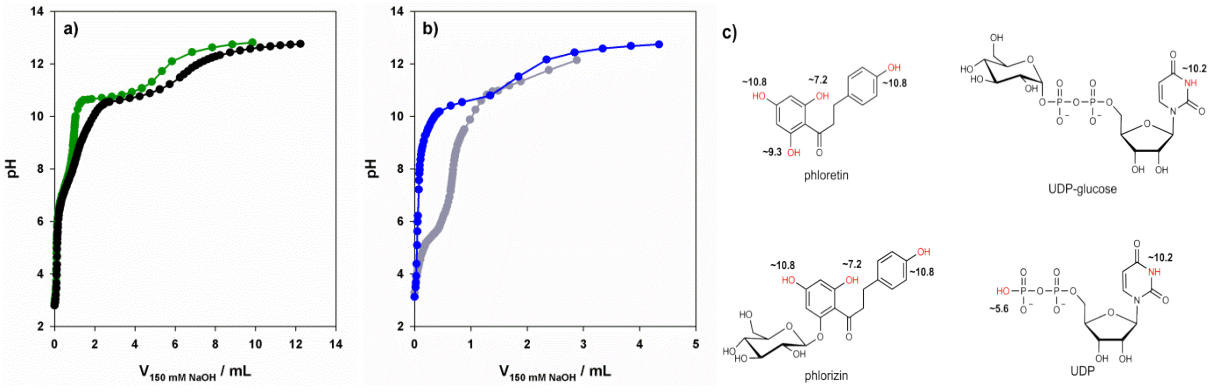


Fig. S5 Potentiometric titrations of substrates and products of *Pc*OGT-catalysed *O*-glycosylation, using 150 mM NaOH at 30°C (7.5 mM analyte in 20% DMSO, 13 mM $MgCl_2$, 50 mM KCl). a) Titration of 20 mL phlorizin (green) and phloretin (black); b) Titration of 10 mL UDP (blue) and UDP-glucose (grey); c) Structures of substrates (phloretin, UDP-glucose) and products (phlorizin, UDP) of the glycosylation reaction catalysed by *Pc*OGT. Approximate pK_a values are indicated at the respective positions.

2.7 Dependency of rearrangement rate on substrate concentrations

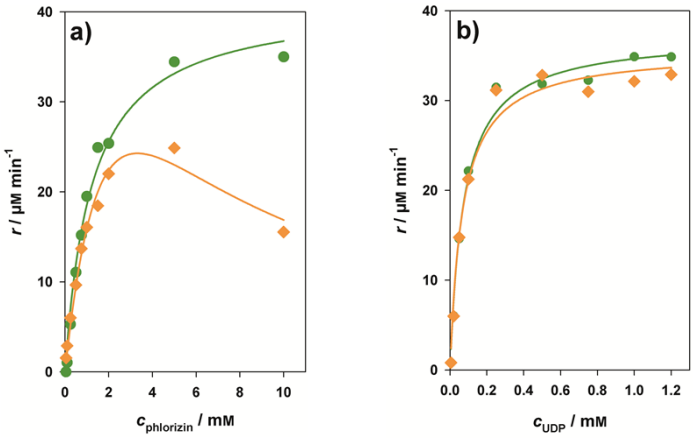


Fig. S6 Dependencies of phlorizin consumption (r_P; green) and nothofagin formation (r_N; orange) rates on the substrate concentration used. The constant substrate concentrations were 2 mM UDP (a) and 5 mM phlorizin (b). All reactions contained 100 mU mL^{-1} *Pc*OGT and 50 mU mL^{-1} *Os*CGT. Symbols show measurements and lines are fits to the data.

2.8 Optimisation of *O*- to *C*-glucoside rearrangement by *Os*CGT_I121D

Feasibility of *O*- to *C*-glucoside rearrangement by a single dual-specific *O*- and *C*-glycosyltransferase was evaluated with the promiscuous *Os*CGT mutant I121D. Initially different concentrations of UDP (0.5 and 5 mM) and phlorizin (0.05, 0.5 and 5 mM) were tested (1.25 mU mL^{-1} *Os*CGT_I121D at pH 7). Time courses for 24 h were recorded and a summary of conversions after 6h reflecting the overall trend is depicted in Fig. S7. An increase of UDP concentration from 0.5 to 5 mM as well as rising phlorizin to 5 mM enhanced the formation of nothofagin (Fig. S7a). While presenting a proof of principle, a maximum yield of around 30 µM was by far too low for synthetic use. In all cases more than 90% of the starting material was still present as phlorizin or phloretin (Fig. S7b-c) and only small amounts of 4'-*O*- and 3'-*C*-glycoside (nothofagin) were found. Highest conversion to nothofagin (4.6 %) was obtained when 0.5 mM phlorizin and 5 mM UDP were applied which served as a starting point for optimizations.

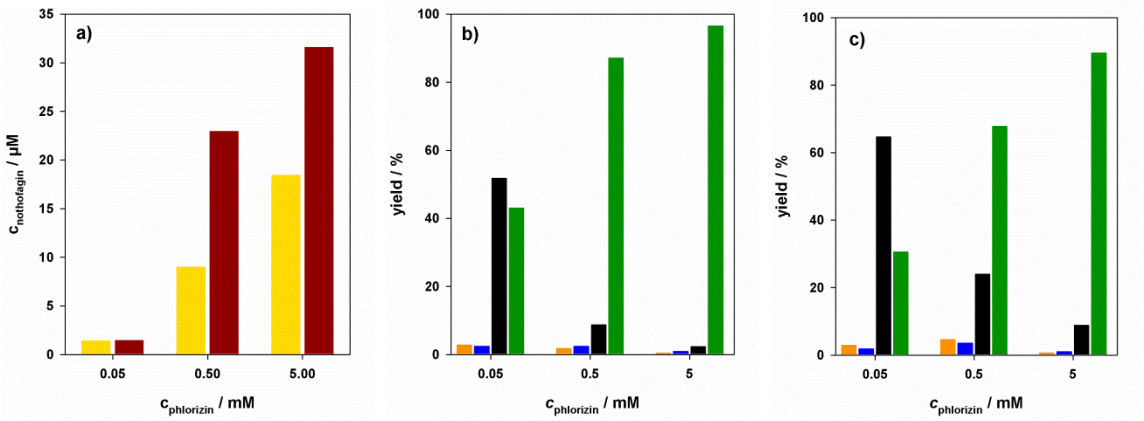


Fig. S7 Yields after 6 h of *O*- to *C*-glycoside rearrangement catalysed by *Os*CGT_I121D (1.25 mU mL^{-1}, pH 7.0) using various substrate concentrations. a) Nothofagin formed from different phlorizin concentrations using 0.5 (yellow) and 5 mM (red) UDP, respectively; Distribution of phloretin and glycosides in presence of 0.5 mM (b) and 5 mM (c) UDP: phlorizin (green), phloretin (black), nothofagin (orange), 4'-*O*-glucoside (blue).

Optimisation of the *O*- to *C*-rearrangement with wild type glycosyltransferases revealed that well balanced *O*- and *C*-glycosyltransferase activities and a suitable equilibrium constant of *O*-glycosylation are key for efficient conversions. Besides substrate concentrations, pH is the key parameter to influence the ratio of these enzymatic activities. Therefore, conversion of 0.5 mM phlorizin in presence of 5 mM UDP was evaluated at pH 6.0, 7.0 and 8.0. The effect of pH on yields after 6 h of incubation (Fig. S8) was in general comparable to results with wild type enzymes (Fig. 2a). At low pH formation of the aglycon was favoured but *C*-glycosylation was impaired. Best yield of *C*-glycoside as well as a favourable ratio of 3'-*C*- to 4'-*O*-glycoside were obtained at pH 7.0.

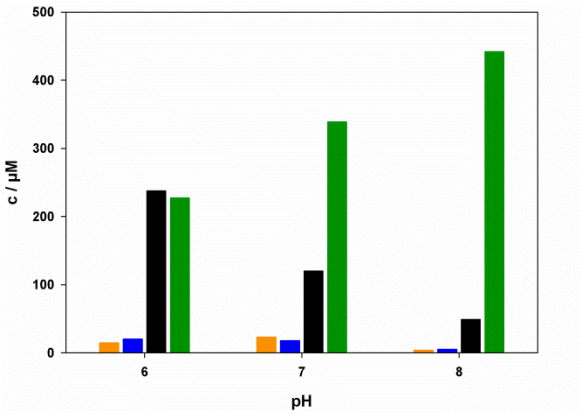


Fig. S8 Distribution of phloretin (glycosides) after 6 h of rearrangement by *Os*CGT_I121D (1.25 mU mL^{-1}) at various pH (50 mM BisTris, 0.5 mM phlorizin, 5 mM UDP): phlorizin (green), phloretin (black), nothofagin (orange), 4'-*O*-glucoside (blue).

Since accumulation of the aglycon phloretin was omnipresent we hypothesized that availability of UDP-glucose could limit *C*-glycosylation. Therefore addition of UDP-glucose (0, 0.5, 1 and 5 mM) was tested. Addition of UDP-glucose should boost final *C*-glycosylation but disfavour initial phloretin production. To analyse potential improvement on *C*-glycoside formation, we focussed on the final yields after 43 h (Fig. S9) rather than on effects on the initial rate. The expected negative effect of UDP-glucose on deglycosylation of phlorizin was detrimental using 5 mM UDP-glucose. At low concentrations (0.5, 1 mM), however, the positive effect of UDP-glucose on *C*-glycosylation prevailed. Although the highest nothofagin yield was obtained in presence of 0.5 mM UDP-glucose we decided to continue working with 1 mM UDP-glucose. We anticipated it to be beneficial for maximizing nothofagin formation since phloretin concentrations (~70 µM) did not seem to be limiting. By doubling the amount of enzyme and increasing the reaction time to 70 h it was finally possible to obtain 70% yield (0.35 mM) in *O*- to *C*-glycoside rearrangement catalysed by *Os*CGT_I121D.

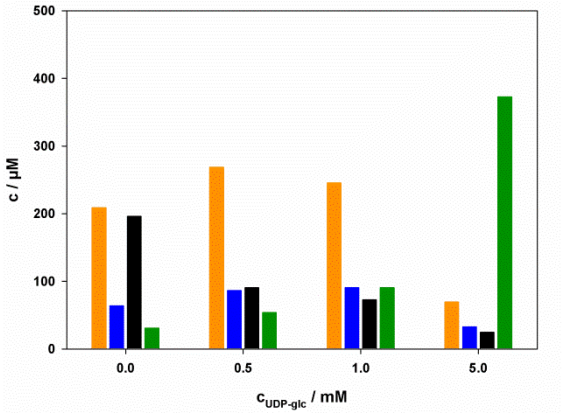


Fig. S9 Distribution of phloretin (glycosides) after 43 h of rearrangement by *Os*CGT_I121D (1.25 mU mL^{-1}) in presence of different UDP-glucose concentrations (0.5 mM phlorizin, 5 mM UDP): phlorizin (green), phloretin (black), nothofagin (orange), 4'-*O*-glucoside (blue).

3. References

1. M. Brazier-Hicks, K. M. Evans, M. C. Gershater, H. Puschmann, P. G. Steel, R. Edwards, *J. Biol. Chem.*, 2009, **284**, 17926-17934.
2. D. P. Dixon, A. Lapthorn, P. Madesis, E. A. Mudd, A. Day, R. Edwards, *J. Biol. Chem.*, 2008, **283**, 20268-20276.
3. A. Gutmann, B. Nidetzky, *Angew. Chem. Int. Ed.*, 2012, **51**, 12879-12883.
4. C. Gosch, H. Halbwirth, B. Schneider, D. Holscher, K. Stich, *Plant Sci.*, 2010, **178**, 299-306.
5. C. Valenta, J. Cladera, P. O'Shea, J.Hadgraft, *J. Pharm. Sci.*, 2001, **90**, 485-492
6. R. M. C. Dawson, D. C. Elliot, W. H. Elliot, K. M. Jones, *Data for Biochemical Research*, Clarendon Press, Oxford, 3rd edn., 1989, ch. 5, pp. 103-114 (and references therein).
7. L. L. Lairson, B. Henrissat, G. J. Davies and S. G. Withers, *Annu. Rev. Biochem*., 2008, **77**, 521-555.
8. L. Elling, M. R. Kula, *Enzyme Microb. Technol.*, 1995, **17**, 929-934.
9. C. Zhang, E. Bitto, R. D. Goff, S. Singh, C. A. Bingman, B. R. Griffith, C. Albermann, G. N. Phillips, J. S. Thorson, *Chem. Biol*., 2008, **15**, 842-853.
10. K. D. Miller, V. Guyon, J. N. S. Evans, W. A. Shuttleworth, L. P. Taylor, *J. Biol. Chem.*, 1999, **274**, 34011-34019.
11. E. V. Chandrasekaran, J. Xue, J. Xia, R. D. Locke, K. L. Matta, S. Neelamegham, *Biochemistry*, 2008, **47**, 320-330.
12. C. Deng, R. R. Chen, *Anal. Biochem.*, 2004, **330**, 219-226.
13. M. Persson, M. M. Palcic, *Anal. Biochem.*, 2008, **378**, 1-7.

4. Appendix

4.1 Maps of expression vectors

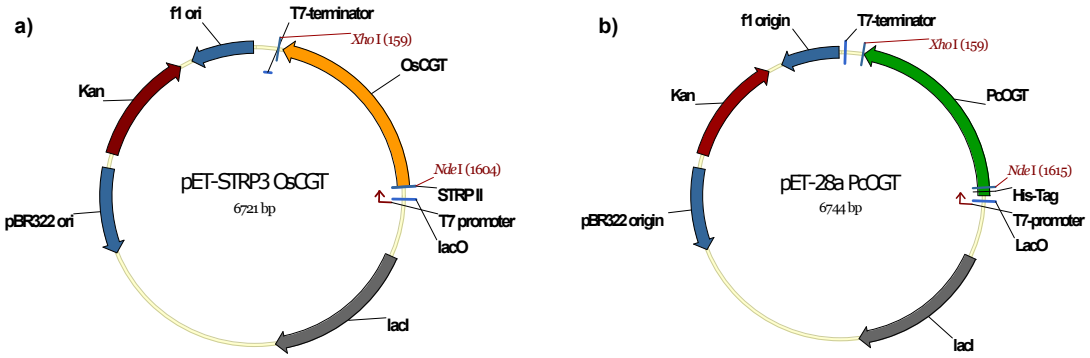


Fig. S10 Map of vectors for *Os*CGT (pET-STRP3) (a) and *Pc*OGT (pET-28a) (b) expressions. Genes are inserted into *Nde*I and *Xho*I sites for expression with N-terminally fused *Strep*-tag II and His-tag, respectively; f1 ori: f1 phage origin of replication, Kan: kanamycin resistance gene, pBR322 ori: pBR322 origin of replication, lacI: lactose repressor gene, lacO: lactose operator, STRP II: *Strep*-tag II, OsCGT: *Os*CGT gene (GenBank: FM179712), PcOGT: *Pc*OGT gene (GenBank: FJ854496)

4.2 DNA sequence of *Os*CGT in pET-Strep3 expression vector

Bold:	confirmed by sequencing
CAPITAL LETTERS:	Expressed open reading frame (*Os*CGT + *Strep*-tag II)
Grey:	*Strep*-tag II

>pET-STRP3_*Os*CGT

atccggatatagttcctcctttcagcaaaaaacccctcaagacccgtttagaggccccaaggggttatgctagttattgctcagcggtggcagc
agccaactcagcttcctttcgggctttgttagcagccggatctcagtggtggtggtggtggtgctcgagccaTATCGGTGGCCTGCAGGTCGAG
TCAATTAGTGCGACATGTTCCCCCTTGGCACAGCCGCGCGAACTCGGCCAGACAACGGTGGCTCGATCCACCGCCGGCGACGGCCTTCGCGGCG
GCCTCGGCGAGGCTCGCCGCCTTCATCCGCAACGCCTCGTCCGCCATCGCCGCCTTCACCTTCTCCGATATCTCCTCCGCGCCGATCACCCCGG
CCTCCCCCTCCCAGCTCCAGGTGTCCGCCCACACGCCGAGCCCGGCGCGCGCCACCACGCCGGAGTTCACCCGCTGGTCGCCGAACCTCGGCAG
CGCCAGCACCGGGACGCCGCTCGCCGCCGCCTCCGTCACCGAGTTCCAGCCGCAGTGGCTCACGAACAGCGCCACGGACTCGTGCTTCAGGACC
TCTTCCTGATCCACCCATGCCTTGGTGACGAGGCCTCGCTTCTCCACCCGCTCCAAGAACCCCTCGTCGAGCAGCTCGCCGAGCTCGGCGGCGT
CGTCCCTGTCCACGACGGTGCTCTTCACCACCCACAGGAACCGGTGGCCGCTGCCCTCCAGCCCGGCGGCGAGTTCCCTGAGCTGCTCCCTTGA
GATGGCCTTCCGGCTGCCGAAGCTCACGTACACCACCGACCGCGCCGGCTGCGCGTCGAGCCACTCCATGTAGTTCGCCTGCGGATCTTTCGCC
TGGTTGCTCGCCGGGAGAAGTGGCCCCACGGCGAACACCGGCGGGAAGCCGGAGGCGACCTTGCCCTGCTGCAGGGCCGCGACGGCCTCCGGCT
CCAAGGCGTCGAACGTGTTGACGAGGATGCCGGCGGCGCTCGTGAGGCTCCGGCCGTTGGCGACGAACTGGCGGGTGAAGAGGTGGTTGGGGTC
GTGCAGCGCCTGCGGGATGGAGGCCTTGGGGATGCGGTACACGCCGGGGATGTCGACGTCGCCGACGCCGCCGCCGCCGCCAGCGTTGGCGTCG
AGGTATGTGGGGAAGTAGGCGCAGAGGGAGAGCATCGCGGCGGAGGCGGTGAAGAGGATGTGGCACGGGAGGCCCTGCTCCTTCGCCACGGGTA
TGACGACGGATGTCAGCGCGATGTCCGTGGCGAGCGCCGACGCGCCGGCGCCGGTGAGGAGCGGGCCGAGGAGCGGCGCCGACCGCCGCATGGC
CTCGAACCGGAGGAAGAACGGGTCGGCGCCGGGGAACTCGGACGCGTCGAACGGCGCGAGCTCGAAGTCGAGCCGCCGCACCGCCGGGAACGCG
TCGAACAGCGCGTCGAGGTGCTTCGACTCCGCGGTGGACACGGTGGGGAGCACCGTGACGAGGGAGACGTCGCAGCCGTGGCCGGAGGACAGCG
CCACGGCGAGGCGGCCGAAGGGGACGAGGTGGCCCATGCCGGCGCTCGGGATGAGCACCACATGCGGCCGCCTGCCCGCCGCGTCGCCAGAGCT
CGGCATATGGTTAATTAAGCCTTTCTCGAACTGCGGGTGGCTCCAGCTAGCCAT**ggtatatctccttcttaaagttaaacaaaattatttctag**
aggggaattgttatccgctcacaattcccctatagtgagtcgtattaatttcgcgggatcgagatctcgatcctctacgccggacgcatcgtgg
ccggcatcaccggcgccacaggtgcggttgctggcgcctatatcgccgacatcaccgatggggaagatcgggctcgccacttcgggctcatgag
cgcttgtttcggcgtgggtatggtggcaggccccgtggccgggggactgttgggcgccatctccttgcatgcaccattccttgcggcggcggtg
ctcaacggcctcaacctactactgggctgcttcctaatgcaggagtcgcataagggagagcgtcgagatcccggacaccatcgaatggcgcaaa
acctttcgcggtatggcatgatagcgcccggaagagagtcaattcagggtggtgaatgtgaaaccagtaacgttatacgatgtcgcagagtatg
ccggtgtctcttatcagaccgtttcccgcgtggtgaaccaggccagccacgtttctgcgaaaacgcgggaaaaagtggaagcggcgatggcgga
gctgaattacattcccaaccgcgtggcacaacaactggcgggcaaacagtcgttgctgattggcgttgccacctccagtctggccctgcacgcg
ccgtcgcaaattgtcgcggcgattaaatctcgcgccgatcaactgggtgccagcgtggtggtgtcgatggtagaacgaagcggcgtcgaagcct

gtaaagcggcggtgcacaatcttctcgcgcaacgcgtcagtgggctgatcattaactatccgctggatgaccaggatgccattgctgtggaagc
tgcctgcactaatgttccggcgttatttcttgatgtctctgaccagacacccatcaacagtattattttctcccatgaagacggtacgcgactg
ggcgtggagcatctggtcgcattgggtcaccagcaaatcgcgctgttagcgggcccattaagttctgtctcggcgcgtctgcgtctggctggct
ggcataaatatctcactcgcaatcaaattcagccgatagcggaacgggaaggcgactggagtgccatgtccggttttcaacaaaccatgcaaat
gctgaatgagggcatcgttcccactgcgatgctggttgccaacgatcagatggcgctgggcgcaatgcgcgccattaccgagtccgggctgcgc
gttggtgcggatatctcggtagtgggatacgacgataccgaagacagctcatgttatatcccgccgttaaccaccatcaaacaggattttcgcc
tgctggggcaaaccagcgtggaccgcttgctgcaactctctcagggccaggcggtgaagggcaatcagctgttgcccgtctcactggtgaaaag
aaaaaccaccctggcgcccaatacgcaaaccgcctctccccgcgcgttggccgattcattaatgcagctggcacgacaggtttcccgactggaaa
agcgggcagtgagcgcaacgcaattaatgtaagttagctcactcattaggcaccgggatctcgaccgatgcccttgagagccttcaacccagtc
agctccttccggtgggcgcggggcatgactatcgtcgccgcacttatgactgtcttctttatcatgcaactcgtaggacaggtgccggcagcgc
tctgggtcattttcggcgaggaccgctttcgctggagcgcgacgatgatcggcctgtcgcttgcggtattcggaatcttgcacgccctcgctca
agccttcgtcactggtcccgccaccaaacgtttcggcgagaagcaggccattatcgccggcatggcggccccacgggtgcgcatgatcgtgctc
ctgtcgttgaggacccggctaggctggcggggttgccttactggttagcagaatgaatcaccgatacgcgagcgaacgtgaagcgactgctgct
gcaaaacgtctgcgacctgagcaacaacatgaatggtcttcggtttccgtgtttcgtaaagtctggaaacgcggaagtcagcgccctgcaccat
tatgttccggatctgcatcgcaggatgctgctggctaccctgtggaacacctacatctgtattaacgaagcgctggcattgaccctgagtgatt
tttctctggtcccgccgcatccataccgccagttgtttaccctcacaacgttccagtaaccgggcatgttcatcatcagtaacccgtatcgtga
gcatcctctctcgtttcatcggtatcattacccccatgaacagaaatccccttacacggaggcatcagtgaccaaacaggaaaaaaccgccct
taacatggcccgctttatcagaagccagacattaacgcttctggagaaactcaacgagctggacgcggatgaacaggcagacatctgtgaatcg
cttcacgaccacgctgatgagctttaccgcagctgcctcgcgcgtttcggtgatgacggtgaaaacctctgacacatgcagctcccggagacgg
tcacagcttgtctgtaagcggatgccgggagcagacaagcccgtcagggcgcgtcagcgggtgttggcgggtgtcggggcgcagccatgaccca
gtcacgtagcgatagcggagtgtatactggcttaactatgcggcatcagagcagattgtactgagagtgcaccatatatgcggtgtgaaataacc
gcacagatgcgtaaggagaaaataccgcatcaggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcg
gtatcagctcactcaaaggcggtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggcca
ggaaccgtaaaaaggccgcgttgctggcgtttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcg
aaacccgacaggactataaagataccaggcgtttccccctggaagctccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctg
tccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgttcgctccaagctgggct
gtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgccact
ggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctacactaga
aggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggta
gcggtggtttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctca
gtggaacgaaaactcacgttaagggattttggtcatgaacaataaaactgtctgcttacataaacagtaatacaaggggtgttatgagccatat
tcaacgggaaacgtcttgctctaggccgcgattaaattccaacatggatgctgatttatatgggtataaatgggctcgcgataatgtcgggcaa
tcaggtgcgacaatctatcgattgtatgggaagcccgatgcgccagagttgtttctgaaacatggcaaaggtagcgttgccaatgatgttacag
atgagatggtcagactaaactggctgacggaatttatgcctcttccgaccatcaagcattttatccgtactcctgatgatgcatggttactcac
cactgcgatccccgggaaaacagcattccaggtattagaagaatatcctgattcaggtgaaaatattgttgatgcgctggcagtgttcctgcgc
cggttgcattcgattcctgtttgtaattgtccttttaacagcgatcgcgtatttcgtctcgctcaggcgcaatcacgaatgaataacggtttgg
ttgatgcgagtgattttgatgacgagcgtaatggctggcctgttgaacaagtctggaaa**gaaatgcataaacttttgccattctcaccggattc**
agtcgtcactcatggtgatttctcacttgataaccttatttttgacgaggggaaattaataggttgtattgatgttggacgagtcggaatcgca
gaccgataccaggatcttgccatcctatggaactgcctcggtgagttttctccttcattacagaaacggctttttcaaaaatatggtattgata
atcctgatatgaataaattgcagtttcatttgatgctcgatgagtttttctaagaattaattcatgagcggatacatatttgaatgtatttaga
aaaataaacaaatagggttccgcgcacatttccccgaaaagtgccacctaaattgtaagcgttaatattttgttaaaattcgcgttaaatttt
tgttaaatcagctcatttttaaccaataggccgaaatcggcaaaatcccttataaatcaaaagaatagaccgagatagggttgagtgttgttc
cagtttggaacaagagtccactattaaagaacgtggactccaacgtcaaagggcgaaaaaccgtctatcagggcgatggcccactacgtgaacc
atcaccctaatcaagttttttggggtcgaggtgccgtaaagcactaaatcggaaccctaaagggagcccccgatttagagcttgacggggaaag
ccggcgaacgtggcgagaaaggaagggaagaaagcgaaaggagcgggcgctagggcgctggcaagtgtagcggtcacgctgcgcgtaaccacca
cacccgccgcgcttaatgcgccgctacagggcgcgtcccattcgcca

4.3 DNA sequence of *Pc*OGT

>PcOGT (internal NdeI and XhoI restriction sites removed from GenBank: FJ854496)
atgggagacgtcattgtactgtacgcatctccagggatggggcacatcgtcgccatggtggagctgggcaagttcattgtccaccgctacggcc
cccacaaattctccatcaccattctctacacctgcggcagcattgtcgacaccgctagcatccccgtctacatccgccgcatctcccactccca
ccctttcatttccttccgccaattccctcgcgtcaccaataatattacccgaaacataagcgtccccgcaatcacgttcgacttcatccgccag
aacgatcctcatgtccgcagtgccctccaagaaatctctaaatccgccaccgttcgcgccttcatcatcgacctcttctgcacctccgctcttc
ccatcgggaaggaattcaacatcccaacatactacttccacacttctggtgccgcagttcttgctgctttttttgtatttgcccaagatcgatga
gcaaaccaaaaccaccgagagtttcaaagacctccgcgacaccgttttcgaattccccggatggaagtcctcctctgaaggctacacacatggtc
caactggtgctcgaccggaacgaccctgcttattcggacatgatctatttctgctcacatcttcccaaatccaacggaatcatcgtcaacacgt
tcgaagagctggagccacctagcgtcctccaggccattgctggaggcctgtgtgttcctgatgggccaactccgcccgtgtactacgttggtcc
attgattgaggaagagaaagaattgagtaaggatgcagatgcggccgagaaggaggactgcttgtcatggctcgataagcagccaagtcgaagc
gtgctgtttctctgtttcggaagcatgggatcatttccggctgctcaactgaaggagatagcgaacgggttggaggcgagcgggcagaggttcc
tgtgggtggtgaagaagccgccggttgaagagaaatcaaagcaggtccatggagttgacgactttgatttgaagggtgtgttgccagaagggtt
tttggagaggacggcagacaggggatggtagtgaagtcatgggcgccgcaggtggtggtgttgaagaaggagtcggttggtgggttcgtgaca
cattgcggatggaactcggtactggaagcagtggttgcgggggtgccgatgattgcttggccgctttacgcggagcagcacatgaacaggaatg
ttctagtgacggacatggaaatcgcgatcggggtggagcagagagacgaggaaggtgggttcgtgagcggggaagaagtggagaggagagtgag
ggagttgatggagtcggaaggaggaagagcgcttagagagaggtgcaagaaacttggggagatggcttcggctgctttgggagagaccggttcg
tccaccagaaacttggtcaactttgttagtagcattaca