Title: Glycosynthases from *Thermotoga neapolitana* β-glucosidase 1A: A comparison of α-glucosyl fluoride and in situ-generated α-glycosyl formate donors

Author: Tania Pozzo Merichel Plaza Javier Romero-García Magda Faijes Eva Nordberg Karlsson Antoni Planas

PII: S1381-1177(14)00171-4
DOI: http://dx.doi.org/doi:10.1016/j.molcatb.2014.05.021
Reference: MOLCAB 2969

To appear in: *Journal of Molecular Catalysis B: Enzymatic*

Received date: 28-1-2014
Revised date: 26-5-2014
Accepted date: 27-5-2014

Please cite this article as: T. Pozzo, M. Plaza, J. Romero-García, M. Faijes, E.N. Karlsson, A. Planas, Glycosynthases from *Thermotoga neapolitana rmbeta*-glucosidase 1A: A comparison of *rmalpha*-glucosyl fluoride and in situ-generated *rmalpha*-glycosyl formate donors, *Journal of Molecular Catalysis B: Enzymatic* (2014), http://dx.doi.org/10.1016/j.molcatb.2014.05.021

Glycosynthases from *Thermotoga neapolitana* β-glucosidase 1A: A comparison of α-glucosyl fluoride and in situ-generated α-glycosyl formate donors

Tania Pozzo[a], Merichel Plaza[b], Javier Romero-García[c], Magda Faijes[c], Eva Nordberg Karlsson[a*], Antoni Planas[c]

[a]Biotechnology, Lund University, P.O. Box 124, SE-22100 Lund, Sweden

[b]Department of Chemistry, Centre for Analysis and Synthesis, P.O. Box 124, SE-22100, Lund, Sweden

[c]Laboratory of Biochemistry, Institut Quimíc de Sarrià, Universitat Ramon Llull, Via Augusta 390, E-08017 Barcelona, Spain

*Corresponding Author. Tel.: +46-46-2224626 *E-mail address:* eva.nordberg-karlsson@biotek.lu.se

Highlights

- Selective synthesis with α-glucosyl fluoride donors and thermophilic glucosynthase.
- Synthesis with *in situ* generated glycosyl-donor also confirmed, but less selective.
- Flavonoids are alternative acceptors for the synthase with both types of donors.

Abstract

*Tn*Bgl1A from the thermophile *Thermotoga neapolitana* is a dimeric β-glucosidase that belongs to glycoside hydrolase family 1 (GH1), with hydrolytic activity through the retaining mechanism, and a broad substrate specificity acting on β-1,4-, β-1,3- and β-1,6-linkages over a range of glyco-oligosaccharides. Three variants of the enzyme (*Tn*Bgl1A_E349G, *Tn*Bgl1A_E349A and *Tn*Bgl1A_E349S), mutated at the catalytic nucleophile, were constructed to evaluate their glycosynthase activity towards oligosaccharide synthesis. Two approaches were used for the synthesis reactions, both of which utilized 4-nitrophenyl β-D-glucopyranoside (4NPGlc) as an acceptor molecule: the first using an α-glucosyl fluoride donor at low temperature (35°C) in a classical glycosynthase reaction, and the second by *in situ* generation of the glycosyl donor with (4NPGlc), where formate served as the exogenous nucleophile under higher temperature (70° C). Using the first approach, *Tn*Bgl1A_E349G and *Tn*Bgl1A_E349A synthesised disaccharides with β-1,3-linkages in good yields (up to 61%) after long incubations (15h). However, the GH1 glycosynthase Bgl3_E383A from a mesophilic *Streptomyces sp*., used as reference enzyme, generated a higher yield at the same temperature with both a shorter reaction time and a lower enzyme concentration. The second approach yielded disaccharides for all three mutants with predominantly β-1,3-linkages (up to 45%) but also β-1,4-linkages (up to 12.5%), after 7h reaction time. The *Tn*Bgl1A glycosynthases were also used for glycosylation of

flavonoids, using the two described approaches. Quercetin-3-glycoside was tested as an acceptor molecule and the resultant product was quercetin-3,4′-diglycosides in significantly lower yields, indicating that *Tn*Bgl1A preferentially selects 4NPGlc as the acceptor.

Keywords: β-glucosidase; α-glucosyl fluoride; Formate; 4-nitrophenyl β-D-glucopyranoside; Quercetin

1. Introduction

Glycoside hydrolases (GH) are enzymes widespread in nature, with the main function of breaking down glycosidic bonds within a diverse array of biological processes including cellular metabolism, defence mechanisms and the degradation of biomass [1]. In order to better understand GHs, they have been classified into more than 132 families based on amino-acid sequence homology data that has been archived within the Carbohydrate Active Enzyme database [2]. The β-glucosidase *Tn*Bgl1A from *Thermotoga neapolitana* belongs to glycoside hydrolase family 1 (GH1 and displays both the typical TIM $(\beta/\alpha)_8$ barrel-fold characteristic of the family and hydrolytic activity over a broad range of glyco-oligosaccharides by the retaining mechanism [3].

The main hydrolytic function of GHs in nature can be changed *in vitro* towards glyco-oligosaccharide synthesis, and the formation of glycosidic linkages can occur via reverse hydrolysis (thermodynamically controlled) or transglycosylation (kinetically controlled) [4,5]. However, the major obstacle and disadvantage to using GHs for glyco-oligosaccharide synthesis lies in the fact that hydrolytic activity remains present as the main function of these enzymes. In general, the use of GHs for these two synthetic reactions results in low yields due to the frequent hydrolysis of glycosidic linkages formed during oligosaccharide synthesis [6,7,8].

In the investigation for efficient methodologies on enzymatic synthesis of glycosidic linkages *in vitro*, an alternative approach has emerged with the creation of hydrolytically-inactive-nucleophile, mutated variants called "glycosynthases", which efficiently catalyse the glycoside linkage formation as a primary function and are capable of achieving high yields for oligosaccharide synthesis [9]. Glycosynthases are hence engineered glycoside hydrolases in which the catalytic nucleophile (Glu or Asp residues) has been replaced by a non-nucleophilic residue with a shorter side chain, commonly Gly, Ala, or Ser [10]. They are inactive hydrolases but efficiently catalyze glycosyl transfer to an acceptor when using an activated glycosyl fluoride donor (GlcF) with the opposite anomeric configuration than the original substrate of the parental wild type hydrolase reaction. [11,12]. An alternative glycosynthase reaction for thermostable glycosynthases was developed using external nucleophiles, such as formate, which mimics the glycosyl-enzyme intermediate mechanism that utilizes activated substrates [13].

In this study, thermostable glycosynthases from *Thermotoga neapolitana* (*Tn*Bgl1A) were created, by mutating the catalytic nucleophile E349 to either G, A or S, and characterised for use in gluco-oligosaccharide synthesis while following two different approaches: a) using α-glycosyl fluoride as a donor and b) using sodium formate as an exogenous nucleophile for in situ generation of an α-glycosyl-formate donor. In the former reaction, Bgl3_E383A from *Streptomyces sp*., reported as an efficient glycosynthase with regioselectivity for β-1,3-linkage [9], was also used to provide a mesophilic enzyme for comparison. In a follow up study, flavonoids were studied as alternative acceptor molecules for enzymatic glycosylation by thermostable glycosynthases under high temperature.

2. Materials and Methods

2.1 Chemicals

Commercial substrates, 4-Nitrophenyl β-D-glucopyranoside (4NPGlc), quercetin-3-β-D-glucoside, quercetin, and sodium formate were supplied by Sigma-Aldrich (Steinheim, Germany). Quercetin-3,4'-di-O-β-D-glucopyranoside was purchased from Polyphenols Laboratories AB (Sandnes, Norway). Formic acid was provided from Merck (Darmstadt, Germany) and methanol by Scharlau (Barcelona, Spain).

The α-glycosyl fluoride, donor molecule was chemically synthesized as described previously [9]. Briefly, peracetylated glucose was fluorinated with hydrogen fluoride pyridine, followed by de-O-acetylation with sodium methoxide in methanol, which resulted in high yields of α-glycosyl fluoride (more than 80%).

2.2 Engineering TnBgl1A glycosynthases

Three inactive-nucleophile variants, *Tn*Bgl1A_E349G, *Tn*Bgl1A_E349A and *Tn*Bgl1A_E349S, were constructed via site directed mutagenesis using both the plasmid encoding the β-glucosidase *Tn*Bgl1A gene from *T. neapolitana* as the template, and the mutagenic primers (Eurofins MWG Operon, Ebersberg, Germany) described in Table 1. PCR reactions were run in a T-Gradient thermocycler (Biometra, Göttingen Germany) using iProof high fidelity DNA polymerase (iProofTM High-Fidelity master mix Bio-Rad laboratories, California, USA); the thermal cycling protocol was as follows: an initial denaturing step of 30 s at 95° C,

followed by 30 cycles of denaturing 30 s at 95° C, annealing 60 s at 55° C and extension of 5 min 30 s at 68° C with a final extension of 10 min at 68° C.

PCR reactions were digested with *Dpn*I (New England, Biolabs) for one hour to remove the template and transformed into electrocompetent *Escherichia coli* NovaBlue (Novagen) using the GenePulser II (Bio-Rad), immediately incubated for 1 hour at 37° C and spread on LB agar containing 100 µg/mL ampicillin. Successfully transformed *E.coli* colonies were selected, and an overnight inoculum was prepared for plasmid extraction in 5 ml LB media with ampicillin (100 µg/mL). Plasmids were purified using the Zyppy plasmid miniprep kit (Zymo Research, CA, USA), and sent for sequencing using vector specific primers T7 and pET-RP supplied by the sequencing facility at GATC Biotech (Konstanz, Germany) to obtain complete coverage of the entire β-glucosidase genes. The mesophilic glucosynthase, Bgl3_E383A, from *Streptomyces sp*. was constructed previously by Vallmitjana *et al.* [14], where the site directed mutagenesis was carried out by using a mutagenic / reverse universal primer pair,and an overlapping PCR method that was described previously by Juncosa *et al.* [15].

2.3 Protein production and purification

Inactive nucleophile variants and wild type *Tn*Bgl1A were transformed into competent *E.coli* BL21 (DE3) (Novagen) for production. Production was carried out in a bioreactor of 2.5 L at 37° C, pH 7, using a defined medium, described in Ramchuran [16] with 100 µg/mL ampicillin. Enzyme production was induced with 0.1mM of IPTG when the OD at 620nm reached 0.8, and cultivation continued up to 3 hours. Cultures were centrifuged (10000 × g, 10 min, 4° C), and cell pellets were dissolved in binding buffer (20 mM imidazole, 20 mM Tris-HCl, 0.75 M NaCl, pH 7.5) and lysed by sonication for 3×2 min using a 14 mm titanium probe sound intensity of 60% and a cycle of 0.5 (UP400S, Dr. Hielscher) followed by a final centrifugation of 30 min, 39000 × g, 4° C as described by Turner [17].

The supernatant was heat treated at 70° C for 30 min followed by centrifugation (10000 × g, 20 min) and the resulting supernatant containing the His-tagged enzyme was purified by immobilised metal ion affinity chromatography (IMAC). The IMAC was performed using a HiTrap affinity 5 ml column (GE health care, Germany) with a

copper ligand on an ÄKTA prime system (Amersham Biosciences, Uppsala, Sweden). Fractions containing the purified enzyme were pooled and dialyzed against 20 mM citrate-phosphate buffer, pH 5.6, overnight using a dialysis membrane with a cut-off of 3,500 Da (Spectrum laboratories, Rancho Dominguez, CA, USA). The dialysed proteins were then stored at 4° C for future use.

The mesophilic Bgl3_E383A was expressed constitutively without the addition of IPTG in 200 ml of LB medium with 100 µg/ml. Mesophilic Bgl3_E383A was then purified by IMAC on HiTrap Chelating Sepharose (Pharmacia) using nickel as described by Vallmitjana [14], followed by dialysis of pure Bgl3_E383A with 20 mM phosphate buffer at pH 7. Enzyme purity was confirmed by visualisation with SDS-PAGE and the protein concentration was estimated by measuring absorbance at 280 nm with the NanoDrop 1000 (Thermo Fisher Scientific).

2.4 Hydrolytic activity

Initial velocities were determined to estimate hydrolytic activity of *Tn*Bgl1A and residual activity for inactive nucleophilic variants at two different pHs (5.6 and 7) and temperatures (35° C and 80° C). The reaction began by adding 20 µl of the enzyme (from 0.01 to 0.5 µM) to a pre-incubated 2 mM 4NPGlc in 50 mM citrate-phosphate buffer at both pH 5.6 and 7 under both thermal conditions. Extinction coefficient values were: $\varepsilon_{M\ 35°C\ pH5.6}$ 1898 $M^{-1}cm^{-1}$; $\varepsilon_{M\ 80°C\ pH5.6}$ 3789 $M^{-1}cm^{-1}$; $\varepsilon_{M\ 35°C\ pH7}$ 14905 $M^{-1}cm^{-1}$; $\varepsilon_{M\ 80°C\ pH7}$ 15203 $M^{-1}cm^{-1}$.

2.5 Chemical reactivation using Sodium Azide

Chemical reactivation experiments of inactive nucleophilic variants using sodium azide (from 0.5 to 3 M) as external nucleophiles were first evaluated on the 96 well plate assay at 35° C for 30 min (Thermo Scientific Multiskan GO UV/Vis Microplate). Different 4NPGlc concentrations were used (0.2 to 4 mM) in a 50 mM citrate-phosphate buffer at pH 5.6 alongside a fixed enzyme concentration of 0.08 µM within a total volume of 250 µl. Blanks for each 4NPGlc concentration were run to assure that spontaneous substrate hydrolysis was negligable.

184

185

186 *2.6 Glycosynthase activity using α-glycosyl fluoride donor*

187 Glycosynthase activity reactions were run, using 1 mM α-glucosyl fluoride (GlcF) and
188 5 mM 4NPGlc in a 50 mM buffer at pH 7. The reaction was pre-incubated for 5 min
189 before adding the enzyme (between 0.5 to 2.5 μM) to achieve a final volume of
190 300 μL for the total reaction. The reactions were monitored upon addition of the
191 enzyme; aliquots of 10 μL were withdrawn every hour for 16 hours in the case of
192 thermophilic glycosynthases, and every 15 min over a 1-hour time period for the
193 mesophilic glycosynthase [9]. Samples were diluted with deionized water (1:10)
194 before injection in the High Performance Liquid Chromatograph (HPLC) for
195 monitoring oligosaccharide synthesis.

196

197 *2.7 Glycosynthase activity by rescue with exogenous nucleophile*

198 Glycosynthase activity reactions were run for the thermophilic glycosynthases using
199 6 mM 4NPGlc in 50 mM citrate-phosphate over a pH-range of 3 to 7 with different
200 concentrations of the exogenous nucleophile sodium formate (0.1 to 2 M). The
201 reaction was pre-incubated for 5 min at 70° C before adding the enzyme (0.5 to
202 2.7 μM). Afterwards, monitoring began by withdrawing aliquots of 20 μL every hour
203 for 7 hours. The blanks that were run are as follows: blank 1, reaction without enzyme
204 to check spontaneous hydrolysis of the substrate with an external nucleophile and
205 blank 2, reaction without external nucleophile to check for the possibility of
206 spontaneous transglycosylation reactions. All aliquots were diluted with deionized
207 water (1:10 and 1:100) before injection to HPLC and High-Performance Anion-
208 Exchange Chromatography with Pulsed Amperometric Detection (HPAEC-PAD).

209

210 *2.8 Monitoring enzymatic synthesis of oligosaccharides.*

211 HPLC was used for monitoring oligosaccharide formation using a C18 column
212 (Novapak, 4 um Waters); identification of the linkage type was achieved by co-
213 injection of standards followed by UV detection at 300 nm. The mobile phase used
214 was 14% methanol (MeOH) in water at 1 mL/min flow rate. Reactions incubated after

16 h were visualised using thin layer chromatography (TLC). Sample was loaded onto a silica gel TLC (Silica 60 F254, Merck, Darmstadt, Germany) and separated using ethyl acetate/MeOH/water (7:2.5:1) as mobile phase. Plates were developed with sulfuric acid/MeOH/water (10:45:45) followed by heating at 120° C [9].

HPAEC-PAD was used to measure glucose liberation during the glycosynthase reaction, which was quantified by integration of glucose peak areas based on a standard curve (0.5 to 10 µM glucose), using a Dionex ICS-5000, Ion Chromatograph equipment (Dionex Corporation, Sunnyvale, CA, USA). The column was CarboPac PA-200 anion-exchange column (3 x 150 mm and 3 x 30 mm CarboPac PA-200 guard column). Glucose was eluted using a flow rate of 0.5 ml min^{-1} by a gradient of eluents: Constant 100 mM NaOH 0 to 20 min, Gradient 0-15 min of 0-34mM NaAc from 15 to 20 min, then constant 100 mM NaAc. Chromeleon 7 software was used to control the system and analyse the data.

2.9 Enzymatic glycosylation of flavonoids

Enzymatic glycosylations of the flavonoid molecules (quercetin and quercetin-3-β-D-glucoside) were carried out at donor:acceptor-ratios of 1:3 using either 4NPGlc or 2NPGlc as a donor and 2 M of sodium formate as exogenous nucleophiles, at 1 µM enzyme concentration in 50 mM citrate phosphate buffer at pH 5.6. Reactions were pre-incubated for 5 min at 70° C before adding the enzyme, after which monitoring began by withdrawing aliquots of 10 µL every hour during 16 hrs. The sample was diluted with MeOH (1:10) before HPLC injection. The blanks measured were as follows: blank 1, reaction without enzyme to check for spontaneous hydrolysis of the substrate with an external nucleophile and blank 2, reaction without an external nucleophile to check for the possibility of spontaneous transglycosylation reactions. The classical glycosynthase reaction using both GlcF and Q3Glc as the acceptor molecules was also tested in a 1:3 ratio at 35°C.

2.10 Monitoring enzymatic glycosylation of flavonoids

Glycosylated forms were detected using an UltiMate-3000® HPLC system from Dionex (Thermo Fisher, Germering, Germany) that consisted of an online degasser, a quaternary solvent pump, an autosampler with cooler, column oven, and photo diode

array detector (DAD) with scanning capabilities, all controlled by Chromeleon 6.80 (Thermo Fisher) software. The detection wavelengths were set at 280, 300 and 350 nm. A porous-shell fused core Ascentis Express C18 (150 mm x 2.1 mm, 2.7 μm.) from Supelco (Bellefonte, PA, USA) was used as an analytical column for LC separation. The column temperature was 50 °C and injection volume was 2 μL, the vial tray was held at 4 °C. Mobile phases consisted of (A) water with 0.5% (v/v) formic acid (B) methanol with 0.5% (v/v) formic acid in a gradient elution analysis programmed as follows: 0 min, 5% (B); 0-5 min, 5% (B); 5-35 min, 40% (B); 35-40 min, 40% (B); at a flow rate of 300 μL/min.

The glycosylated flavonoids were further identified by MS/MS, and the equipment used was a liquid chromatograph Accela equipped with DAD coupled to a Thermo Exactive orbitrap mass spectrometer (MS) via an electrospray interface (ESI) (Thermo Fisher, Germering, Germany). The MS was operated under ESI negative mode. Tuning and optimization were carried out using a direct injection of quercetin 3-glucoside standard (0.004 mmol) at 0.750 mL/min. The carrier gas was ultrahigh-purity nitrogen.

2.11 Structural studies of β-glycosynthases

Docking experiments were prepared using the crystal structure of *Tn*Bgl1A_E349G (PDB-code to be added). This structure was protonated at pH 5.6 with the server H++ [20]. An α-glucosyl fluoride (donor) molecule drawn using the server PRODRG [18], was placed at the -1 subsite of the*Tn*Bgl1A_E349G structure by means of a superimposition with the ligand (2-deoxy-2-fluoro-alpha-D-glucopyranose) of β-glucosidase *Tm*BglA from *Thermotoga maritima* (1OIN) and was subsequently minimized using GROMACS [19]. The acceptor molecule 4NPGlc was obtained from the crystal structure 3AI0, the quercetin 3-glucoside was drawn with PRODRG, and both were used as ligands for a docking experiment where they were placed along both the +1 and +2 subsites of the *Tn*Bgl1A_E349G structure. The docking was realized by AutoDock VINA [20] and the chosen docked ligand was minimized with both the protein and donor using GROMACS. The results and figures were visualized using VMD [22].

3. Results and Discussion

3.1 Engineering, expression and purification

The gene encoding the β-glucosidase, *Tn*Bgl1A from *T. neapolitana*, was previously cloned into the pET-22b(+) vector under control of the T7/lac promoter for production of the protein product in *E.coli* [23]. This construct was amplified in the methylating *E.coli* strain Nova Blue, and was used as a template for engineering three inactive nucleophile variants, *Tn*Bgl1A_E349G, *Tn*Bgl1A_E349A and *Tn*Bgl1A_E349S by site directed mutagenesis. The mutations were inserted using a ligation independent amplification of the template plasmid with overlapping mutagenic primers (Table 1), followed by selective digestion of the methylated template. Later on, for production, the separate plasmids encoding either one of the nucleophile variants or *Tn*Bgl1A wild type (*wt*) were transformed into *E.coli* BL21 (DE3) and each line produced the target protein in 2.5 l batch cultivations at 37° C, pH 7. Although overall expression was high (around 50% of total protein), the major fraction of recombinant enzyme was insoluble, resulting in a yield for the desired soluble target enzyme that accounted for approximately 5% of the total soluble protein present in the extract after cell disruption. There was no visible difference in the production yields between the *wt* and the three nucleophile-mutated variants.

The high thermostability of the enzymes motivated a heat treatment step, which resulted in approximately 50% purity, where more than 90% purity (as judged by SDS-PAGE) was reached after IMAC purification using a copper ligand that bound selectively to the C-terminal His6-tag that was fused to all of the enzymes. After the two purification steps, approximately 50 mg of purified TnBgl1A was obtained per litre of batch-culture.

3.2 Hydrolytic activity

The classical glycosynthase reaction, using fluorinated sugars, was performed at temperatures no higher than 40º C and pH 7, due to lability of the fluorinated donor [24,25]. Optimal conditions for *Tn*Bgl1A's hydrolytic activity were previously determined to be 80° C and pH 5.6 in Khan's work [26]. The results displayed in Figure 1 show that hydrolytic activity was evaluated under different conditions of

temperature (80°C and 35°C) and pH (5.6 and 7), where the initial velocity of *Tn*Bgl1A at 80°C, pH5.6 was considered to be 100%. When the temperature was reduced to 35°C (keeping the optimal pH of 5.6) the activity decreased to 25%. The effect of the pH was also significant within the interval, due to the fact that at 80°C an increase in pH from 5.6 to 7 resulted in a 30% reduction in activity. This observed level of enzymatic activity fell within the same range as the activity drop observed upon the temperature reduction from 80°C to 35°C at a constant pH. Despite the only 10% enzymatic activity that remained under the conditions of 35ºC, pH 7, these appeared to be appropriate for the glycosynthase reactions with mutated nucleophile variants. This low activity level also explains the demand of higher enzyme concentrations and longer incubation times (in addition to the fact that most glycosynthases have lower specific activity compared to their primary *wt*-hydrolase activity).

Any residual hydrolytic activities in the mutated nucleophile variants was also assayed under the previously described experimental conditions, and resulted in either little or no detectable activity, the results of which are presented in Table 2.

3.3 Chemical reactivation of mutated nucleophile variants with Sodium Azide

Chemical reactivation with sodium azide (NaN_3) was performed next in order to evaluate the possibility to utilize the mutated nucleophile variants as glycosynthases. Different NaN_3 concentrations (0.5 to 4 M) were assayed in an effort to restore activity to the inactive nucleophile variants at 35° C by using 4NPGlc (from 0.2 to 3 mM) as a substrate in yielding the α-glucosyl azide product [9]. The results indicated that 1 M NaN_3 was the appropriate concentration for restoring maximum activity to both *Tn*Bgl1A_E349A and *Tn*Bgl1A_E349S. The observed level of reinstated activity by *Tn*Bgl1A_E349S was shown to be twice as high as for *Tn*Bgl1A_E349A, whereas *Tn*Bgl1A_E349G failed to demonstrate any activity as shown in Figure 2.

3.4 Glycosynthase activity using α-glucosyl fluoride donor

As all three mutated nucleophile variants *Tn*Bgl1A_E349A, *Tn*Bgl1A_E349S, and *Tn*Bgl1A_E349G showed low residual hydrolytic activity (Table 2) and two (*Tn*Bgl1A_E349A, *Tn*Bgl1A_E349S) of the three were reactivated with an exogenous azide (Fig. 2), it was considered amenable to test all three as glycosynthases. The

classical glycosynthase reaction mechanism, described in Fig.3, was carried out using α-glucosyl fluoride and 4NPGlc as the donor and acceptor molecules respectively in a ratio of 1:5.

All three glycosynthase variants (including *Tn*Bgl1A_E349G, which was not chemically rescued with azide) synthesized a single disaccharide product that was identified by HPLC and TLC as 4-nitrophenyl-β- laminaribioside (4NPLam), which is an indication of a glycosynthase activity with selectivity towards a disaccharide with a β-1,3 glycosidic linkage. In a parallel experiment, the well-known glycosynthase Bgl3_E383A, derived from *Streptomyces sp.* [9] was assayed under similar experimental conditions, producing substantial yields of 4NPLam (Table 3).

In a first trial (60 min duration at 35° C, pH 7), the nucleophile variants of *Tn*Bgl1A seemed to have low initial rates and overall efficiencies compared to Bgl3_E383A, which reached 100 % yield in 4NPLam production. However, this was not surprising due to the fact that these thermostable enzymes were operating at least 45° C below their activity optimum, where only 10% of the optimum activity was expected to be achieved (Fig.1, Table 2).

In an attempt to improve the oligosaccharide yields first obtained at 0.5 μM enzyme, the reaction was run at a higher enzyme concentration for a longer incubation time (Table 3). After 15 hours, *Tn*Bgl1A_E349G (2.5 μM) improved to show yields up to 62%, *Tn*Bgl1A_E349A (1.7 μM) yielded 21% oligosaccharide, and *Tn*Bgl1A_E349S (0.5 μM) showed a 45% yield (Table 3). This was in contrast to cases reported by Moracci [13] on glycosynthases derived from thermophilic enzymes where reactions performed at low temperature (35ºC) generated product yields lower than 10%, suggesting that syntheses performed under these temperatures following the classical approach are capable of producing good yields if experimental conditions are optimized.

3.5 Glycosynthase activity using exogenous nucleophile

Another strategy that has been developed to conduct the glycosynthase reaction is *in situ* generation of the activated glycosyl donor, which is particularly interesting for thermophilic enzymes [13,27,28] as the conventional glycosyl fluoride donor is unstable at high temperatures and requires the glycosynthase reaction to be carried out at 30-37° C; a temperature far below the optimal activity of a thermophilic enzyme.

The *in situ* generation of the α-glucosyl formate complex (by chemical rescue with sodium formate, leading to a reactive adduct that mimics the glycosyl-enzyme intermediate) creates a potential glycosyl donor for the glycosynthase reaction with 4NPGlc as the acceptor molecule (Fig.4).

Glycosynthase reactions were run using 6 mM 4NPGlc and 2 M sodium formate at 70° C, in 50 mM citrate-phosphate at the optimal pH 5.6. Further optimization of the glycosynthase reaction was carried out by varying the sodium formate concentration from 0.1 M to 3 M at different pH values (3 to 7), but the oligosaccharide synthetic yields were lower at pHs above 6 (data not shown), and at low pH (3 to 4) no glycosynthase reaction was detected at all. This indicated that the synthase reaction performed by the *Tn*Bgl1A-glycosynthases shared the pH optimum of 5.6 with hydrolytic reactions of the *wt* enzyme (Fig. 5 and Table 4).

When using the exogenous nucleophile approach, the enzyme was less regioselective than in the previous approach that used α-glucosyl fluoride as the donor. The product was mainly 4NP-β- laminaribioside, but 4NP-β- cellobioside was also detected in lower amounts. Low concentrations (not quantifiable) of tri- and tetra-saccharides were also detected by HPLC when the enzyme concentration was increased for the reaction.

An observation made, during the glycosynthetic reaction using formate as the exogenous donor molecule, was that glucose was formed by spontaneous hydrolysis of the labile α-glucosyl formate intermediate. For this reason, it may be asserted that this hydrolytic activity was in competition with the transglycosylation of the acceptor molecule. Both reactions have been plotted in Figure 5, demonstrating that both the hydrolytic activity and the transglycosylation reaction occur to approximately the same extent for *Tn*Bgl1A_E349A and *Tn*Bgl1A_E349G.

In addition, it has been reported by Perugino and Trincone [29,30] that enhanced activity of glycosynthases from a thermophilic origin can occur when using *in situ* generation of the donor at low concentrations of exogenous nucleophile and low pH, which was suggested to be caused by increased protonation of the acid/base that facilitates an increased efficiency for the first step of the reaction (Fig. 4). However, this was not observed for *Tn*Bgl1A glycosynthases and the synthetic activity disappeared at lower pH (data not shown).

3.6 Enzymatic glycosylation of flavonoids

A previous report showed that *Tn*Bgl1A has a high affinity for hydrolysing flavonoid glucosides, with K_M values of 0.06 and 0.13 mM for quercetin 4′-glucoside (Q4′Glc) and quercetin 3-glucoside (Q3Glc) respectively; K_M values lower than that of 4NPGlc at 0.2 mM [26] provide a strong indication that these flavonoids could be used as acceptors, thus expanding the use of *Tn*Bgl1A-glycosynthases towards antioxidant modification. The benefit of modifying flavonoid structures is that their physicochemical properties can be influenced, as well their bioavailability.

The results from the screening of flavonoid acceptors revealed that the synthesis of glycosylated flavonoids was possible with *Tn*Bgl1A-glycosynthase. However, these glycosynthases were shown to preferentially select a glycosylated flavonoid rather than a non-glycosylated one as the acceptor molecule. The use of quercetin aglycone as an acceptor did not result in any detectable glycosylation (data not shown), while the use of Q3Glc was found to be more successful. The major product of the reactions with Q3Glc was quercetin 3,4′diglucoside (Q3,4′diGlc), which was identified using standards of this molecule by HPLC and Mass Spectrometry (MS). The yields observed from the synthesis of di-glucosylated flavonoids were determined to be low (Table 5) when compared to the amount of product formed in oligosaccharide synthesis, holding true over both synthetic approaches (using either the GlcF or 4NPGlc with formate donor, respectively). The use of 2NPGlc as the donor molecule resulted in a significant yield improvement (Table 5), which indicated that accommodation at the -1 subsite (see also paragraph 3.7) provided better intermolecular interactions, facilitating the glycosynthase activity.

The flavonoid reactions were carried out in presence of 50% (v/v) methanol, a solvent that was added to dissolve the flavonoid. Optimization of the solvent concentration may thus lead to increased yields, as in this case a co-solvent concentration was used that has been observed to reduce activity of many enzymes. High co-solvent concentration can however be more advantageous in combination with a thermostable enzyme like *Tn*Bgl1A, as many thermostable enzymes are able to tolerate high solvent concentrations [17, 31, 32].

442 *3.7 Structural considerations on glycosynthase reactions*

443 Putative enzyme/substrate interactions in *Tn*Bgl1A were analysed by molecular
444 modelling to rationalize the observed regiospecificity in the glycosynthase reactions.
445 In addition, comparisons between the crystal structures of *Tn*Bgl1A_E349G (PDB-
446 code, to be added) and Bgl3_E383A (1GNX), an efficient glycosynthase previously
447 reported by Faijes [9], were carried out. Both enzymes were shown to be dimers in
448 their respective crystal structure, with a TIM $(\beta/\alpha)_8$ barrel fold and two conserved
449 motifs, TXNE and TENG, containing the conserved acid/base and nucleophile of the
450 GH1 family. In general, *Tn*Bgl1A_E349G was found to have an open active site with
451 more loops than Bgl3_E383A. The residues building the glycone (-1) subsite in GH1
452 are highly conserved interacting tightly with the donor molecule by hydrogen bonds
453 The residues building the glycone (-1) subsite in the GH1 enzymes are highly
454 conserved and interact tightly with the donor molecule by hydrogen bonding [3]. Both
455 enzymes were regioselective, as the linkage in the major glycosynthetic product of
456 both enzymes was found to be a β-1,3-glycosidic linkage. This would be influenced
457 by the accommodation of the donor and acceptor via tight interactions with the
458 residues at the -1 and +1 subsites, respectively. Important interacting residues of the -
459 1 subsite in TnBgl1A were determined to be W120, Y293, and the general base E164,
460 while W322 and W166 correspond to important residues at the +1 subsite (Fig 6A).
461 Similar interactions were observed in Bgl3; Figure 6b demonstrates the putative
462 important roles played by residues W148, Y311, and E292 in the -1 subsite, as well as
463 W370 in the +1 subsite. In both structures, the closer hydroxyl group of the acceptor
464 (Glc-pNP) to the anomeric carbon of the donor is the 3-hydroxyl (distances of 3.8-3.7
465 Å, respectively, Fig 6) in agreement with the observed β-1,3 specificity. At 70°C,
466 *Tn*Bgl1A-glycosynthases also produced a minor β-1,4-disaccharide product (using the
467 formate approach). Despite the longer distance between the anomeric carbon of the
468 donor and the 4-hydroxyl of the acceptor (about 6Å in the modelled structure in Fig
469 6a), formation of a β-1,4-glycosidic linkage may reflect the higher flexibility of the
470 enzyme at 70°C that results in less stringent substrate specificity.

471 *Tn*Bgl1A_E349G residues at the aglycone area (+2 and +3 subsites) interacting with
472 the Q3Glc were also identified, most of which were characterised as aromatic
473 hydrophobic residues: Trp322 displays stacking interaction with the flavonoid rings B
474 and C. Tyr175 is also well located to interact with the A-ring of the flavonoid, while

E406 and H178 interact with the glucosyl moiety of the Q3Glc acceptor (Figure 7). The 4′-hydroxyl of Q3Glc was determined to have a distance of 3.7 Å to the anomeric carbon of GlcF and in this manner, the flavonoid appeared to fit as the acceptor molecule at the *Tn*Bgl1A active site to form the Q3,4′-diGlc product.

4. Conclusion

The "classical" glycosynthase approach that utilized α-glycosyl fluoride donors was, despite the lability of the fluoride donors and the necessity to use a sub-optimal temperature of the thermophilic enzyme *Tn*Bgl1A, shown as a promising method for oligosaccharide synthesis. Synthesis resulted in at least 60% yield and selectivity for 4NPLam formation. The yield by an exogenous nucleophile was higher than the yield using *in situ* generation of the glycosyl donor. The approach using formate (as exogenous nucleophile) also resulted in lower selectivity as both 4NPLam and 4NPCel were produced, albeit the latter at low concentrations.

In an attempt to explore new areas for the application of thermostable glycosynthases, quercetin-3-glycoside was also used as an acceptor molecule for the glycosylation of flavonoids. Although product yields were lower, production of quercetin-3,4′-diglycoside was confirmed and this demonstrated the potential for the use of glycosynthases in synthetic reactions that utilize a much broader range of acceptor molecules.

Acknowledgements

We acknowledge the Swedish research council Formas (supporting the research program SuReTech) and EU FP7 (supporting the project AMYLOMICS) for financial support. Geneco Research School is highly appreciated for the travel grant that supported Tania Pozzo as an exchange researcher at the University Ramon Llull, Barcelona, Spain.

References

[1] C. S. Rye and S. G. Withers, Curr. Opin. Chem. Biol. 4, (2000) 573–580.

[2] B. L. Cantarel, P. M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, B. Henrissat, Nucl. Acids Res. 37 (2009) D233–D238.

[3] S. R. Marana, IUBMB Life 58 (2008) 63–73.

[4] T. Desmet and W. Soetaert, Biocat. Biotrans. 29 (2011) 1–18.

[5] J. Thimm and J. Thiem, In: Glycoscience (eds B.O. Fraser-Reid, K. Tatsuta, J. Thiem) Springer-Verlag Berlin Heidelberg (2008) 1387-1409.

[6] P. Monsan and F. Paul, FEMS Microbiol. Rev. 16 (2006) 187–192.

[7] D. H. Crout and G. Vic, Curr. Opin. Chem. Biol. (1998) 98–111.

[8] A. G. Watts and S. G Withers, Biochem. J. 380 (2004) e9-e10.

[9] M. Faijes, M. Saura-Valls, X. Perez, M. Conti, A. Planas, Carbohydr. Res. 341 (2006) 2055–2065.

[10] E. Andrés, H. Aragunde, A.Planas. Biochem. J. 458, (2014) 355-363.

[11] L. F. Mackenzie, Q. Wang, R. A. J. Warren, S. G. Withers, J. Am. Chem. Soc. 120 (1998) 5583–5584.

[12] C. Malet and A. Planas, FEBS Letters 440 (1998) 208–212.

[13] M. Moracci, A. Trincone, G. Perugino, M. Ciaramella, M. Rossi, Biochem. 37 (1998) 17262–17270.

[14] M. Vallmitjana, M. Ferrer-Navarro, R. Planell, M. Abel, C. Ausín, E. Querol, A. Planas, J.-A. Pérez-Pons, Biochem. 40 (2001) 5975–5982.

[15] M. Juncosa, J. Pons, T. Dot, E. Querol, A. Planas, J. Biol. Chem. 269 (1994) 14530-14535.

[16] S. O. Ramchuran, O. Holst, E. Nordberg Karlsson, J. Biosci. Bioeng. 99 (2005) 477-84.

[17] P. Turner, D. Svensson, P. Adlercreutz, E. Nordberg Karlsson, J. Biotechnol. 130 (2007) 67-74.

[18] A.W. Schüttelkopf and D.M. Van Aalten, Acta Crystallogr. D60 (2004) 1355–1363.

[19] B. Hess, C. Kutzner, D Van Der Spoel, E Lindahl, J. Chem. Theory Comput. (2008) 435-447.

534 [20] V. Hornak R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling,
535 Proteins, 15 (2006) 712-725.

536 [21] O. Trott, A. J. Olson, J. Comput. Chem. 31 (2010) 455-461.

537 [22] W. Humphrey, A. Dalke, K. Schulten, J. Mol. Graphics 14 (1996) 33-38.

538 [23] C. Turner, P. Turner, G. Jacobson, M Waldebäck, P Sjöberg, E. Nordberg
539 Karlsson, K. Markides. Green Chem 8 (2006) 949-959.

540 [24] S.J. Williams and S.G. Withers, Carbohydr. Res. 327 (2000) 27-46.

541 [25] M. Faijes, X. Pérez, O. Pérez, A. Planas, Biochem. 42 (2003) 13304–13318.

542 [26] S. Khan, T. Pozzo, M. Megyeri, S. Lindahl, A. Sundin, C. Turner, E. Nordberg
543 Karlsson, BMC Biochem. 12 (2011) 1-15.

544 [27] A. Trincone, G. Perugino, M. Rossi, M. Moracci, Bioorg. Med. Chem. Letters 10
545 (2000) 365–368.

546 [28] G. Perugino, A. Trincone, A. Giordano, J. van der Oost, T. Kaper, M. Rossi, M.
547 Moracci, Biochem. 42 (2003) 8484–8493.

548 [29] G. Perugino, P. Falcicchio, M. Michela Corsaro, I. Matsui, M. Parrilli, M. Rossi,
549 M. Moracci, Biocat. Biotrans. 24 (2006) 23–29.

550 [30] A. Trincone, A. Giordano, G. Perugino, M. Rossi, M. Moracci, ChemBioChem 6
551 (2005) 1431–1437.

552 [31] T. Hansson, and P. Adlercreutz, Biotechnol. Bioeng. 75 (2001) 656–665.

553 [32] T. Hansson, M. Andersson, E. Wehtje, P. Adlercreutz, Enz. Microb. Technol. 29
554 (2001) 527-534.

555

555 **Table 1**

556 Primers used for site directed mutagenesis of *Tn*Bgl1A.

Primer	Sequence (5' to 3')
BglAE349GF	GAGGTGTACATCACA**GGG**AACGGAGCTGCATTC
BglAE349GR	GAATGCAGCTCCGTT**CCC**TGTGATGTACACCTC
BglAE349AF	GAGGTGTACATCACA**GCG**AACGGAGCTGCATTC
BglAE349AR	GAATGCAGCTCCGTT**CGC**TGTGATGTACACCTC
BglAE349SF	GGAGGTGTACATCACA**TCG**AACGGAGCTGCATTCG
BglAE349SR	CGAATGCAGCTCCGTT**CGA**TGTGATGTACACCTCC

557 Codons changed are underlined in bold.

558

559

560

Table 2

Hydrolase activity determined at different temperatures (35°C and 80°C) and pH values (5.6 and 7) for *Tn*Bgl1A and its nucleophile variants.

Enzyme	Vo/[E)] (s^{-1})			
	pH 5.6 80°C	pH 5.6 35°C	pH 7 80°C	pH 7 35°C
*Tn*Bgl1A	87	17	23	6
*Tn*Bgl1A_E349G	0.04	0.04	0.24	0.01
*Tn*Bgl1A_E349A	N.D.	0.03	0.02	N.D.
*Tn*Bgl1A_E349S	0.20	0.13	1.2	0.03

N.D = not detected. Reaction of 2mM of 4NPGlc and different enzyme concentrations (0.01 to 0.5 μM)

568

569 **Table 3**

570 Glucosynthetic evaluation of 4NPLam formation at 1 mM GlcF and 5 mM 4NPGlc

571 condition run at 35° C, pH 7, monitored using HPLC.

Enzyme	[E] (μM)	Vo ($M/s \times 10^{-7}$)	Vo/[E] (s^{-1})	Time (h)	Yield β-1,3 (%)
*Tn*Bgl1A_E349G	0.5	0.05±0.0002	0.009	15	29
*Tn*Bgl1A_E349G	2.5	0.23±0.0010	0.009	15	62
*Tn*Bgl1A_E349A	0.5	N.D			
*Tn*Bgl1A_E349A	1.7	0.04±0.003	0.002	15	21
*Tn*Bgl1A_E349S	0.5	0.07±0.049	0.013	15	45
Bgl3_E383A	0.05	4.14±0.566	8.299	0.5	60

572 N.D =Not Detected

573

574

575

575

576 **Table 4**

577 Oligosaccharide synthesis using exogenous nucleophile.

Enzyme	pH	Conc [μM]	Vo [M/s × 10^{-7}]	Vo/(E) [s^{-1}]	Time [h]	Yield β1,3 [%]	Yield β1,4 [%]
*Tn*Bgl1A_ E349G	5.6	0.5	0.398±0.009	0.079	7	45	6
	4	0.5	N.D				
	7	0.5	0.151±0.008	0.003	7	28	3
	5.6	2.7	1.940±0.182	0.072	1	35	10
*Tn*Bgl1A_ E349A	5.6	0.5	0.272±0.047	0.054	7	42	13
	4	0.5	N.D				
	7	0.5	0.074±0.001	0.015	7	15	3
	5.6	2.5	0.371±0.006	0.014	1	16	4
*Tn*Bgl1A_ E349S	5.6	0.5	0.059±0.011	0.012	1	2	0.17

578 Reaction run with, 6 mM 4NPGlc and 2 M sodium formate, at 70° C product were
579 monitored by HPLC. N.D = Not Detected.

580

581

Table 5

Enzymatic glycosylation of flavonoids.

Enzyme	Conc. [μM]	Donor 1mM	Acceptor 3mM	Time [h]	Yield Q3,4-DiGlc [%]
*Tn*Bgl1A_ E349G	1	4NPG[a]	Q3Glc	16	12
	1	2NPG[a]	Q3Glc	16	37
*Tn*Bgl1A_ E349A	1	4NPG[a]	Q3Glc	16	12
	1	2NPG[a]	Q3Glc	16	19
*Tn*Bgl1A_ E349G	1	GlcF	Q3Glc	16	10
*Tn*Bgl1A_ E349A	1	GlcF	Q3Glc	16	6

Glycosynthase reaction using a ratio 1:3 donor acceptor, using both glycosynthetic approaches, product monitored using HPLC. [a]*In situ* donor formation using 2 M sodium formate N.D = Not Detected

*Graphical Abstract (for review)

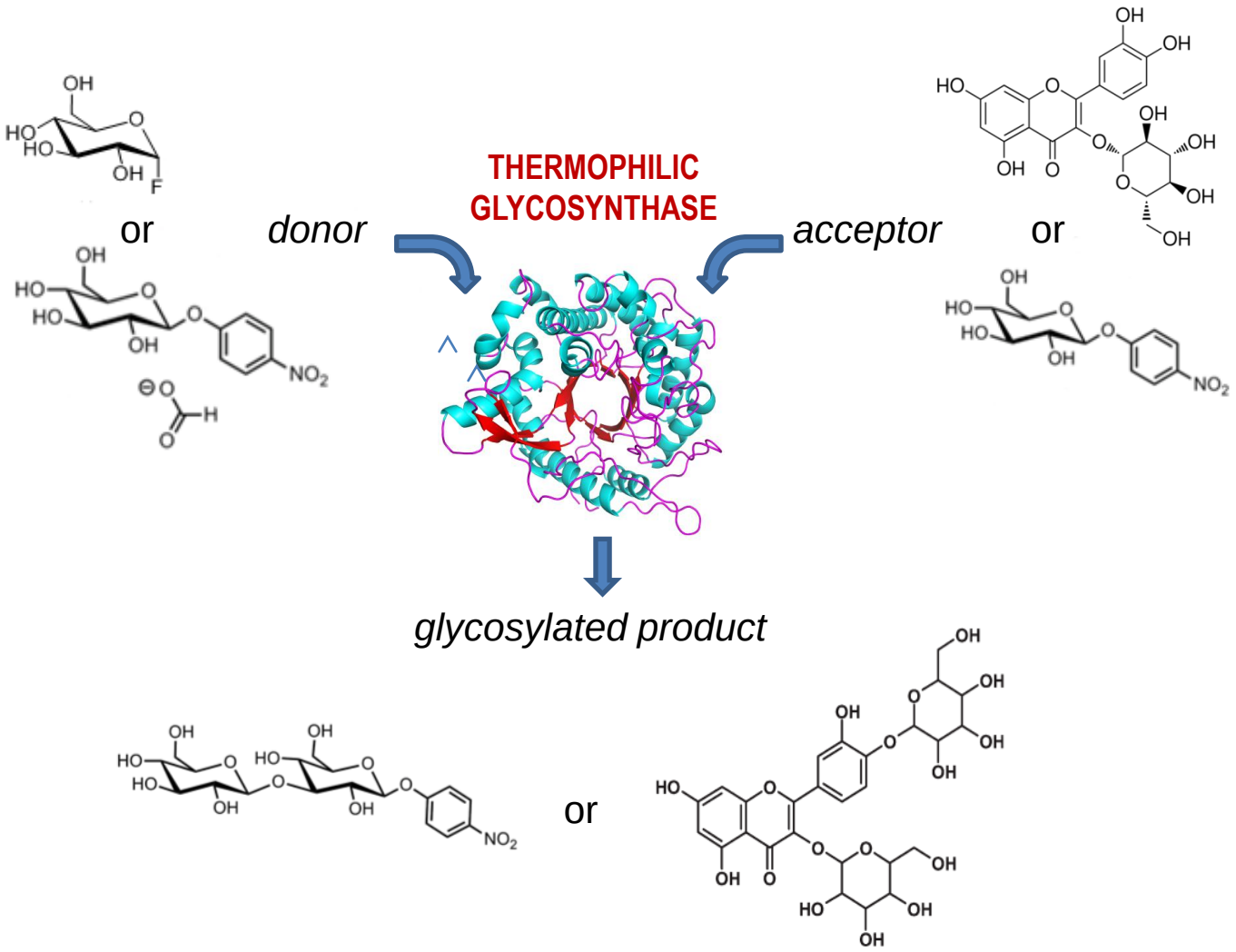


Figure 1.

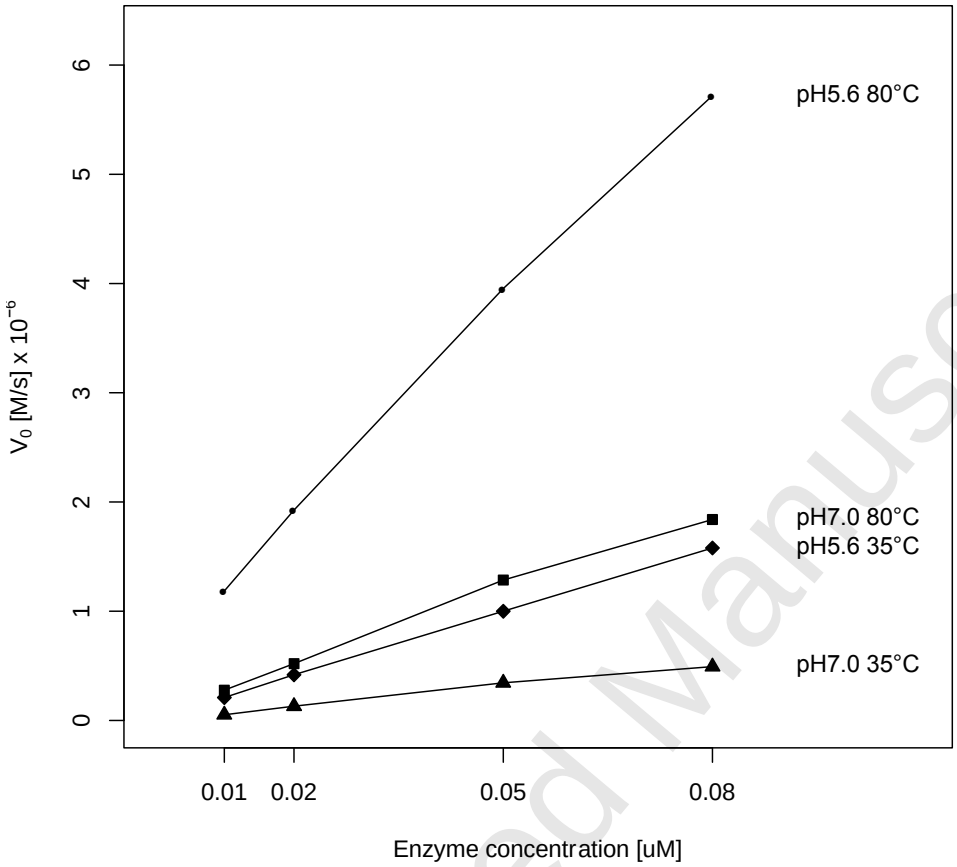


Fig.1. Linear dependence of initial velocity and enzyme concentration for *Tn*Bgl1A hydrolase activity, determined at different conditions of pH (5.6 and 7) and temperature (35°C and 80°C).

Figure 2

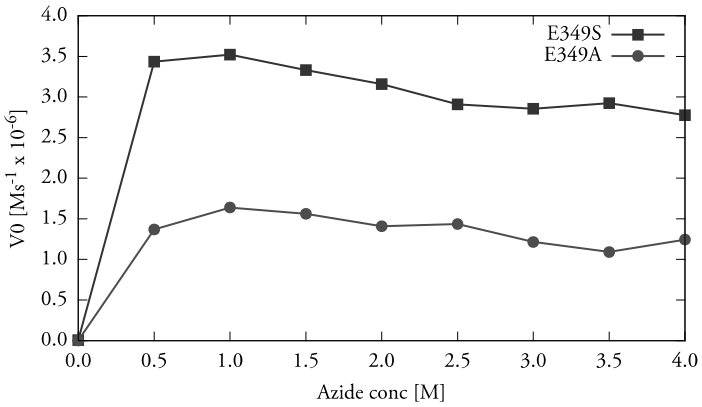


Fig. 2. Chemical reactivation for *Tn*Bgl1A inactive nucleophile variants using different concentrations of NaN_3 at 35° C, pH 5.6 and a fixed enzyme concentration of 0.08 μM.

Figure 3

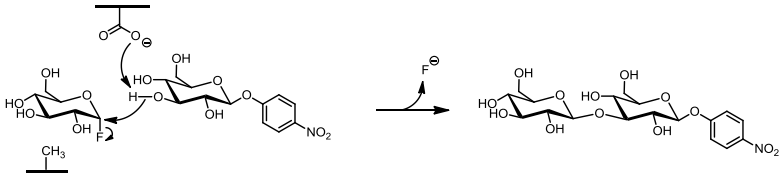


Fig.3. Classical glycosynthase mechanism.

Figure 4

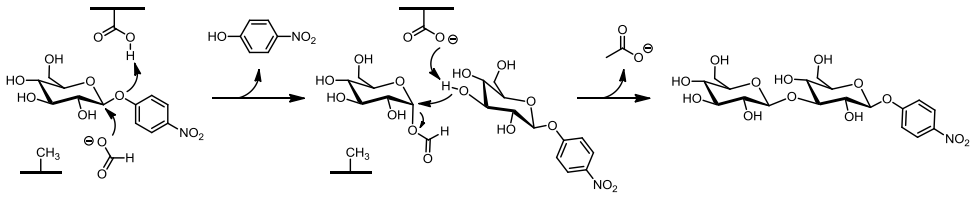
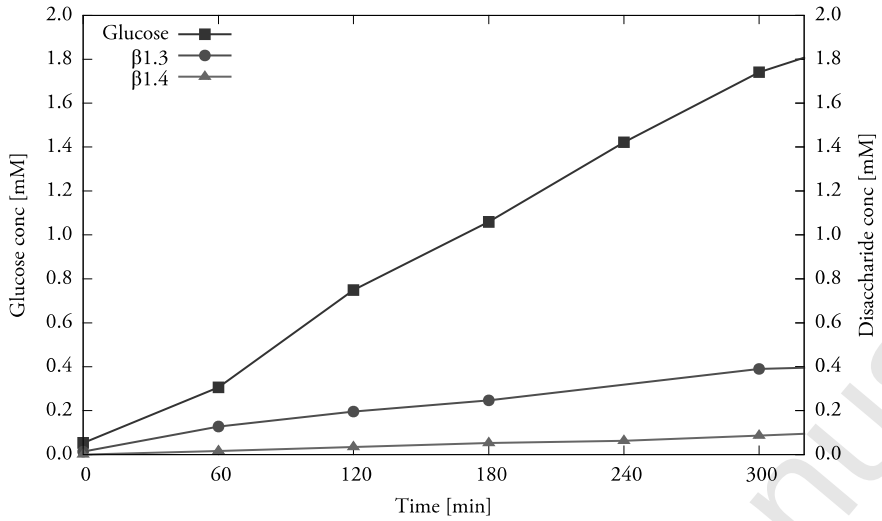


Fig.4. Glycosynthase activity using formate as exogenous nucleophile

Figure 5

A.



B.

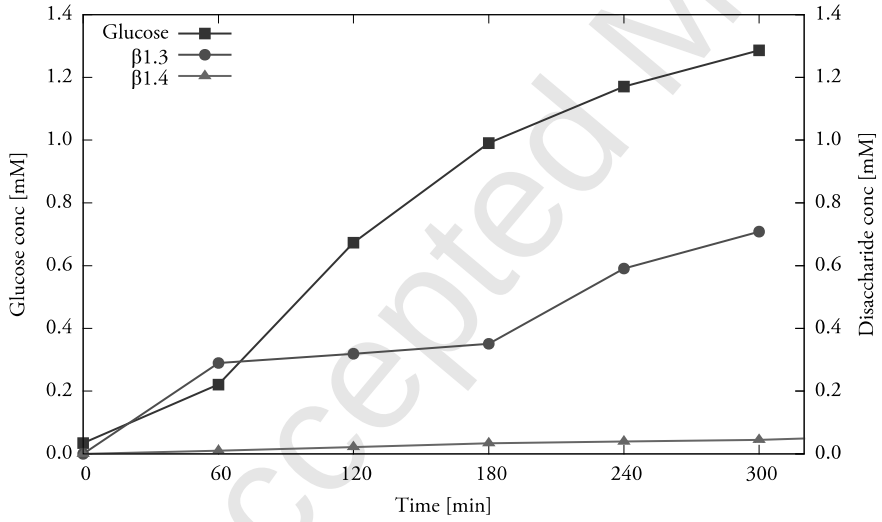


Fig. 5. Glucose and disaccharide formation during glycosynthase reaction in presence of formate for A. *Tn*Bgl1A_E349A and B. *Tn*Bgl1A_E349G. Glycosynthase reaction carried out at 6mM 4NPGlc, 0.5 μM of enzyme at pH 5.6 and 70° C.

Figure 6

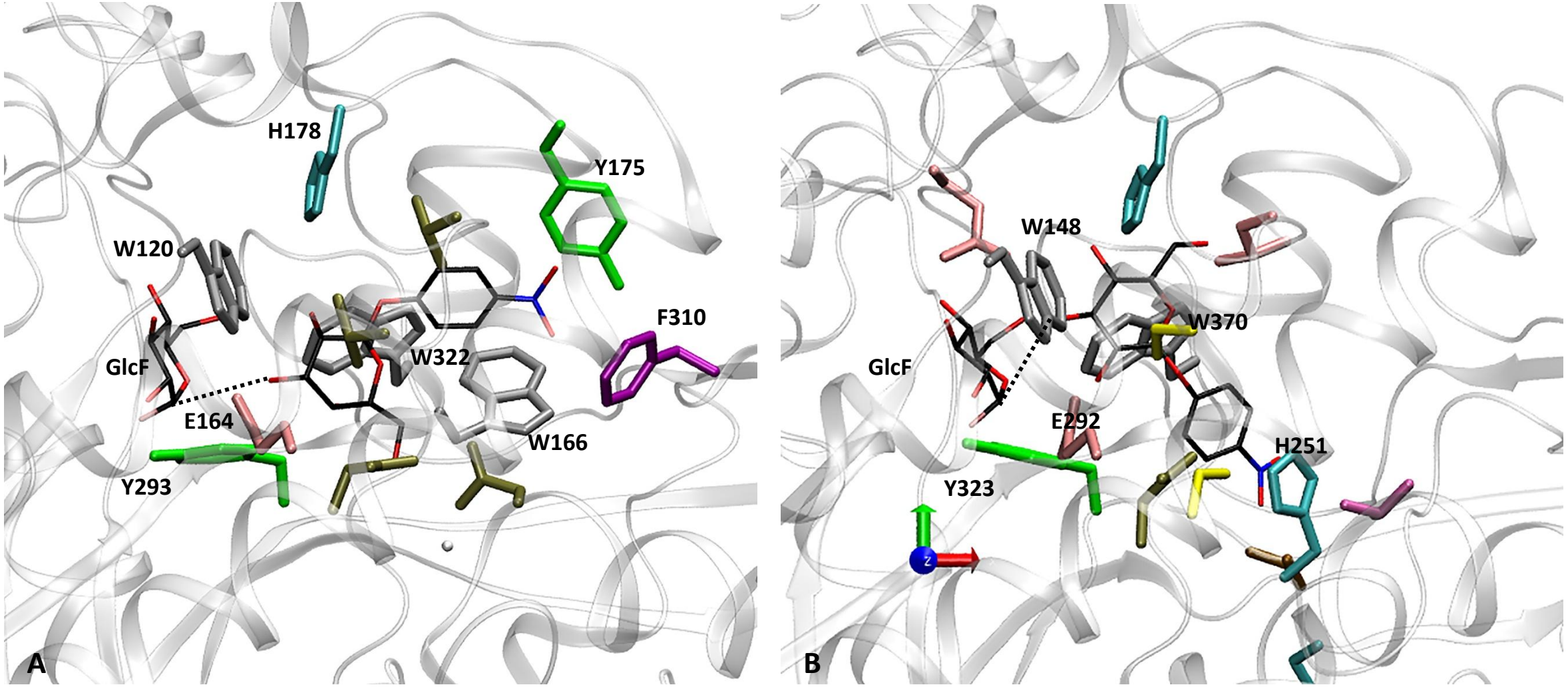
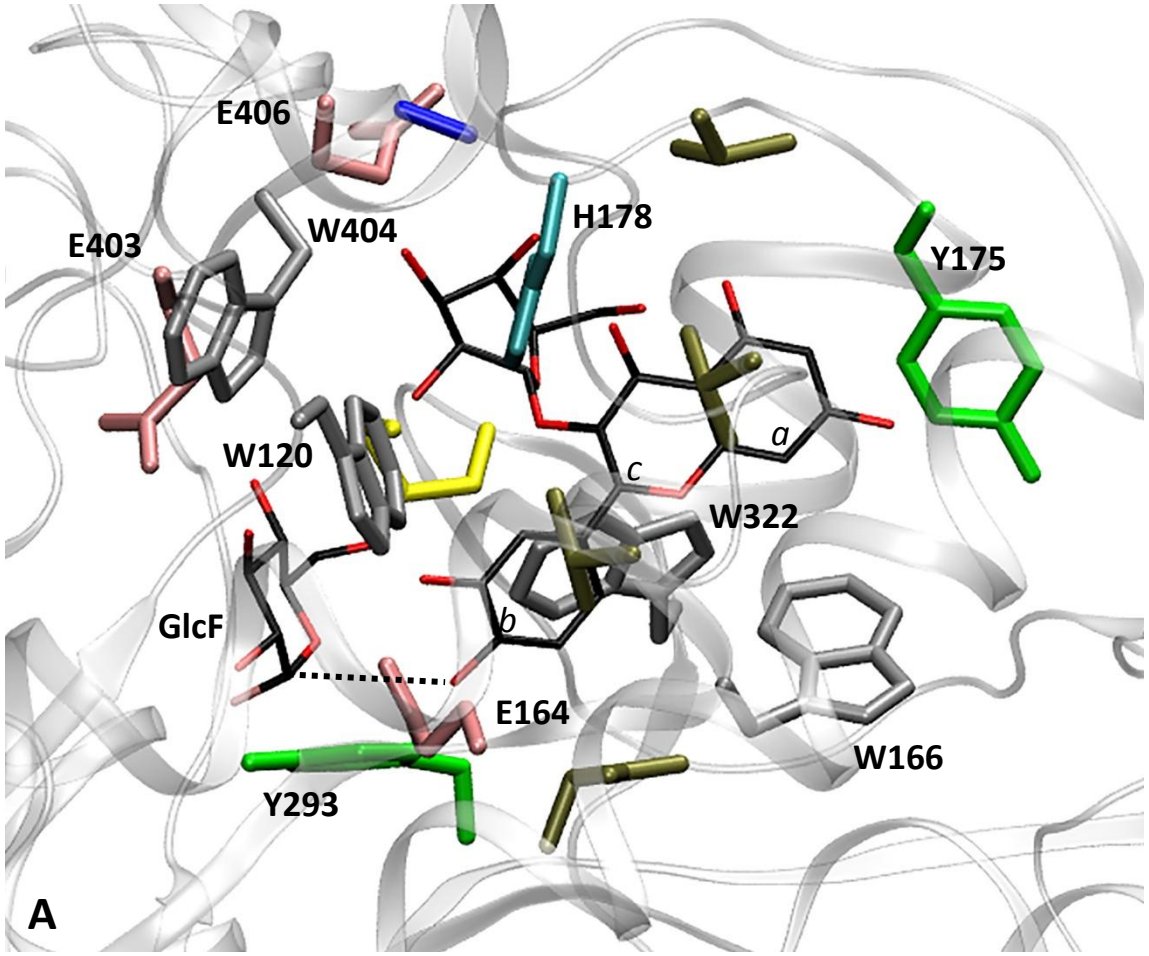


Figure 6. A. Docking GH1_TnBgl1A_E349G with PNPGlc. Dashed line, distance between PNPGlc(O3) and GlcF(C1), 3,8 Å. **B. Docking Bgl3_E383A with PNPGlc**. Dashed line, distance between PNPGlc(O1) and GlcF(C1), 3,71 Å.

Figure 7



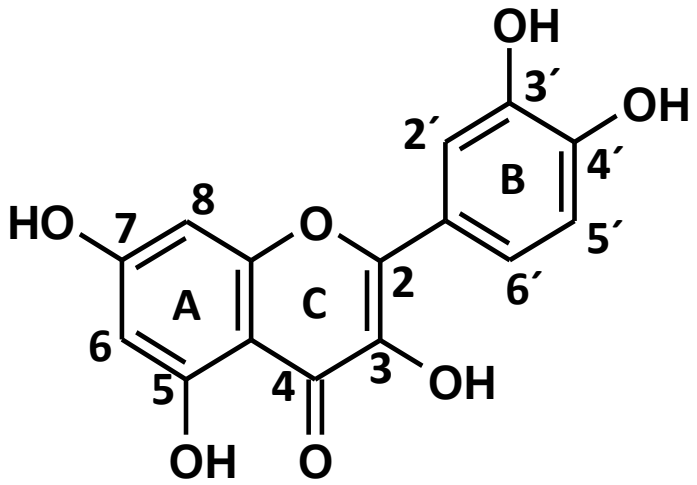


Figure 7. A. Docking GH1_TnBgl1A_E349G with Q3Glc. Dashed line, distance between Q3Glc(O4') and GlcF(C1), 3,73 Å. The numbering of the rings in the quercetin-molecules are indicated as: a, (A-ring); b. (B-ring) and c. (C-ring). **B.** Schematic structure of the quercetin backbone with the designation of the rings (A,B,C) and numbering of carbons indicated.