Semi-rational approach for converting a GH36 α-glycosidase into an α-transglycosidase.

David Teze, Franck Daligault, Vincent Ferrières, Yves-Henri Sanejouand, Charles Tellier

Semi-rational approach for converting a GH36 α-glycosidase into an α-transglycosidase

David Teze[2], Franck Daligault[2], Vincent Ferrières[3,4], Yves-Henri Sanejouand[2] and Charles Tellier[1,2]

[2]Université de Nantes, CNRS UMR 6286, UFIP, 2 rue de la Houssinière, F-44322 Nantes cedex 03, France.

[3]Ecole Nationale Supérieure de Chimie de Rennes, CNRS, UMR 6226, 11 allée de Beaulieu, CS 50837, 35708 Rennes cedex, France.

[4]Université européenne de Bretagne, France.

[1]To whom correspondence should be addressed: tel: +332 51 12 57 33; fax: +332 51 12 56 32; e-mail: charles.tellier@univ-nantes.fr

Running title: Semi-rational design of transglycosidases

KEYWORDS:

Evolution / Glycosidase / Mutagenesis / Oligosaccharide / Transglycosylation

SUPPORTING INFORMATION AVAILABLE: Primers used for site-directed mutagenesis are listed in Tables **TS1** and **TS2**. Fifteen electropherograms obtained for GalGal-*p*NP synthesis study of the best variant (N334T) are presented in Figures **S1** to **S15.** Kinetic behaviors of mutants W336F and C526A are presented in Figure **S16.** Kinetic

behaviors of mutants W411F and D367N are presented in Figure **S17.** (Supplementary Data).

ABSTRACT: A large number of retaining glycosidases catalyze both hydrolysis and transglycosylation reactions. In order to use them as catalysts for oligosaccharide synthesis, the balance between these two competing reactions has to be shifted towards transglycosylation. We previously designed a semi-rational approach to convert the *Thermus thermophilus* β-glycosidase into transglycosidases by mutating highly conserved residues located around the -1 subsite. In an attempt to verify that this strategy could be a generic approach to turn glycosidases into transglycosidases, *Geobacillus stearothermophilus* α-galactosidase (AgaB) was selected in order to obtain α-transgalactosidases. This is of particular interest as, to date, there are no efficient α-galactosynthases, despite the considerable importance of α-galactooligosaccharides. Thus, by site-directed mutagenesis on 14 AgaB residues, 26 single mutants and 22 double mutants were created and screened, of which 11 single mutants and 6 double mutants exhibited improved synthetic activity, producing 4-nitrophenyl α-D-galactopyranosyl-(1,6)-α-D-galactopyranoside (GalGal-*p*NP) in 26 to 57% yields against only 22% when native AgaB was used. It is interesting to note that the best variant was obtained by mutating a second-shell residue, with no direct interaction with the substrate or a catalytic amino acid. As this approach has proved to be efficient with both α- and β-glycosidases, it is a promising route to convert retaining glycosidases into transglycosidases.

INTRODUCTION

Carbohydrates play major roles in most biological processes such as inflammation reactions, cell death, cellular interactions, stabilization of proteins, adhesion of pathogenic microorganisms and binding of toxins (Hart and Copeland 2010; Lichtenstein and Rabinovich 2013; Ghazarian, Idoni, and Oppenheimer 2011). In particular, most naturally-occurring human antibodies recognize a carbohydrate, the α-gal epitope (Macher and Galili 2008), and α-galactosyl residues are present in glycan antigens targeted by carbohydrate vaccines against various pathogens such as *Neisseria meningitidis, Shigella dysenteriae* and *Leishmania* spp. (Astronomo and Burton 2010). Furthermore, α-1,6 linked galactooligosaccharides have been recognized for their prebiotic properties (Andersen et al. 2013; Cervera-Tison et al. 2012; Hachem et al. 2012; Nakai et al. 2010; Wang et al. 2014). However, their use is hindered by their poor availability: unlike for proteins and nucleic acids, there is no general synthetic route for oligosaccharides and, despite the considerable development of efficient methods in this field (Wang et al. 2007; Zhu and Schmidt 2009), the assembly of oligosaccharides remains a substantial challenge. Consequently, alternative synthetic strategies have been explored, particularly those based on enzymatic processes.

The enzymatic synthesis of glycosidic bonds can be performed by glycosyltransferases or transglycosidases. The former can catalyze the formation of glycosidic bonds with both high yield and selectivity, but their use is limited in the large-scale production of oligosaccharides by the difficulty in accessing activated sugar donors (sugar nucleotides for Leloir-type transferases or phosphate sugars for non-Leloir types) (Lairson et al. 2008), even though a recent approach seems promising (Gantt et al. 2013; Zhang et al. 2006). As an alternative, engineered retaining glycosidases have provided an efficient approach for the synthesis of

oligosaccharides. A first methodology was based on the substitution of the catalytic nucleophile by a neutral amino acid and the use of activated donors with an anomeric configuration opposite to that of the original substrate (such as α-glycosyl fluorides), which yielded new enzymes called glycosynthases (Wang et al. 1994; Mackenzie et al. 1998; Malet and Planas 1998; Cobucci-Ponzano et al. 2012). However, although these glycosynthases have been very efficient for the synthesis of β-glycosidic bonds, the results are disappointing when they are applied to α-galactosidases, despite recent advances (Cobucci-Ponzano et al. 2010). A second methodology based on the Hehre-resynthesis hydrolysis properties of inverting glycosidases (Hehre et al. 1979) has been proposed. Mutating residues involved in the hydrolysis reaction but not in the Hehre-resynthesis reaction yield variants that efficiently catalyze the synthesis β-linkages (Honda and Kitaoka 2006; Honda et al. 2008). This methodology was also extended to the synthesis of α-linkages (Wada et al. 2008), but with less efficiency.

As an alternative, natural α-transglucosidases, such as amylosucrase and dextransucrase, have been identified in CAZy (Cantarel et al. 2009) families GH13 and GH70 that can efficiently catalyze the formation of α-glucosidic bonds using sucrose as a natural activated donor (Irague et al. 2011; Champion et al. 2009; Pizzut-Serin et al. 2005; Plou, Segura, and Ballesteros 2007). These examples of natural transglycosidases suggested that directed evolution of glycosidases into transglycosidases could be possible and this strategy was validated in our laboratory on both α- and β-glycosidases (Feng et al. 2005; Osanjo et al. 2007; Koné et al. 2009). Although molecular modeling and structural analysis were tentatively applied to explain why these evolved mutants present improved transglycosidase activities (Teze et al. 2014; Osanjo et al. 2007; Feng et al. 2005; Teze et al. 2013b), the molecular events leading to an improved transglycosylation over hydrolysis ratio have not yet been determined precisely, even for natural transglycosidases such as cyclodextrin

glycosyltransferases (Leemhuis et al. 2002; Kelly et al. 2008). However, the data led us to propose the working hypothesis that mutating conserved residues within the donor subsite of *Thermus thermophilus* β-glycosidase would enable the creation of transglycosidases and indeed four new β-transglycosidases were obtained (Teze et al. 2014).

Therefore, the aim of this study was to investigate whether this approach could also enable the rapid creation of efficient α-transgalactosidases. AgaB was chosen as a model system as this enzyme, which belongs to the CAZy family 36 (Lombard et al. 2014), subgroup I (Fredslund et al. 2011; Hachem et al. 2012), has shown interesting transgalactosylation properties. It was shown to catalyze the synthesis of 4-nitrophenyl α-D-galactopyranosyl-(1,6)-α-D-galactopyranoside (GalGal-pNP) with yields up to 53% when high substrate concentration were used (100 mM) (Spangenberg et al. 2000). Higher yields of this disaccharide (74% with 40 mM of substrate) have been obtained with the α-galactosidase from *Aspergillus nidulans,* which belongs to the same subgroup and shares 39% identity with AgaB (Nakai et al. 2010). We also previously showed that the regioselectivity of the enzyme could be modulated by mutagenesis (Dion et al. 2001a, b). In addition, its crystallographic structure has recently been published (Merceron et al. 2012). Recently, the first reported α-galactosynthase was created from a GH36 enzyme, the *Thermotoga maritima* α-galactosidase (TmαGal, subgroup III) (Cobucci-Ponzano et al. 2010) and moderate improvement of the transglycosylation catalyzed by the same enzyme has been reported (Bobrov et al. 2013). Interestingly, the GH36 family also contains pure transgalactosidases, such as stachyose and verbascose synthases, which are important for carbon storage in seeds (Peterbaueret al. 2002a, b; Peterbauer et al. 2003). However, these plant enzymes, which belong to the GH36 subgroup II (Fredslund et al. 2011), are difficult to produce and their structure is not known. Their sequences present less than 15% identity with AgaB, preventing reliable sequence alignments, not to mention homology modeling. Nevertheless, their existence suggests that

glycoside hydrolases from this family could be evolved to pure transglycosidases, with little or no hydrolysis activity, while keeping their original double-displacement mechanism.

The present work describes a semi-rational approach to convert the *Geobacillus stearothermophilus* α-galactosidase (AgaB) into transglycosidases by mutating highly conserved residues located around the -1 donor subsite. An initial Ala scanning allowed identifying critical positions that affect the hydrolysis/tranglycosylation ratio, and the most effective positions were then mutated into a more conservative amino acid. We found that 50% of the mutated positions led to an improvement of the transglycosidase activity, suggesting that targeting the -1 site of glycosidases could significantly decrease the screening effort to evolve glycosidases into tranglycosidases.

RESULTS

Amino acid conservation within the GH36 family

To study the effect of mutations on conserved residues within the donor subsite of the enzyme, the conservation of all AgaB amino acids was investigated. Thus, using the AgaB sequence as the query, 205 sequences from the GH36 subgroup I were obtained that shared between 25 and 95% of sequence identity between all pairs of sequences and between 25 and 66% with the query sequence. The average conservation was 54.4%, with 5.2% of amino acids (38 AAs) at least 99.5% conserved (Fig. **1A**). Given the AgaB structure (PDB code 4FNQ) and the very homologous AgaA structure (97% identity) in complex with stachyose (PDB code 4FNU) (Merceron et al. 2012), an α-Gal residue was added to the AgaB structure by superimposition (RMSD=0.5 Å). The minimal distance between one of the non-hydrogen atoms of the aforementioned conserved residues and one of the glycosyl non-hydrogen atoms was measured (Fig. **1B**). It is particularly interesting to note that 10 residues have a minimal

distance less than 3.5 Å from the glycosyl, and that no other residues have a minimal distance less than 4.5 Å. These 10 residues constitute the first shell, directly involved in catalytic activity (D478 and D548) or in substrate binding (D366, D367, W411, K476, N480, C526, G529 and W545; Fig. **2A**). Six of these highly conserved residues are located from 4.5 to 6.1 Å away from the glycosyl, interacting with first-shell residues and thus constituting the conserved residues of the second interaction shell (N334, W336, R373, E413, R443 and S547; Fig. **2B**). These 16 residues located at less than 6.1 Å from the glycosyl are likely to be involved in the enzyme function and their impact on the transglycosylation is worth exploring. Then, the next conserved residue is the G368 at 7 Å, which might be involved in protein folding.

Ala-scanning

Among the 16 conserved AAs close to the donor subsite, two are the catalytic residues (D478 and D548) and the remaining 14 were individually mutated into alanine to assess if these residues have an influence on the transglycosylation yields. The resulting mutants were produced and their kinetic parameters k_{cat}/K_M as well as their ability to catalyze GalGal-*p*NP synthesis are summarized in Table **I**.

Most of the mutants, except for the E413A and W411A variants, display a non-Michaelian behavior as they do not present any saturation at high substrate concentration, thus catalytic constants, K_M and k_{cat} , could not be individually determined and only the apparent second-order rate constant k_{cat}/K_M ("catalytic efficiency") is presented (Table I). The WT enzyme and the G529A showed a slight substrate inhibition. Examples of these non-classical behaviors are presented in Figures **S16** and **S17** (Supplementary data).

Of the 14 mutants produced, only 6 displayed sufficient activity to measure their yield of GalGal-*p*NP synthesis when used with α-Gal-*p*NP 10 mM as a substrate (Table **I**). This Ala-scanning was successful for rapidly obtaining transglycosidases, with three variants

presenting transglycosylation yields in the range of 31 to 38% against 22% with the WT. However, the kinetic parameter k_{cat}/K_M of these six mutants dropped dramatically by two to three orders of magnitude. It should be noted that in three (G529A, C526A and S547A) of these six variants, Ala mutation resulted in low steric change and that two of these three presented the highest activities among all variants. While Ala-scanning was performed to analyze the effect of each conserved residue substitution in a similar fashion, it appeared that minimizing the steric difference between the variant and the native enzyme was needed in order to avoid drastic loss of activity. Thus, a second round of mutagenesis featuring "conservative mutations", i.e. mutations closest in structure to the wild-type enzyme, was performed.

Conservative mutations

Of the 14 chosen positions, 10 were selected for a second round of mutagenesis in which the novel amino acid replaced the original one in a more isosteric way. G529, C526 and S547 were excluded as Ala is already a good replacement choice and mutating R373 was excluded as it seemed to engender a strong deleterious effect on the transglycosylation yield, with a drastic drop from 22 to 2% (Table **I**). As the N334A mutant was the best variant selected by Ala-scanning, more substitutions on this specific position were added. The kinetic parameters k_{cat}/K_M as well as GalGal-*p*NP synthesis yields of the resulting mutants are summarized in Table **II**.

In the set of variants obtained by "conservative mutations", only three mutants were inactive: R443Q, D366N, E413Q, and a 10- to 75-fold increase in k_{cat}/K_M was observed on four positions with regard to Ala mutants as well as a 15- to 100-fold decrease in the amount of enzyme needed to reach reaction completion. Consequently, 8 out of 9 variants were active enough to assay GalGal-*p*NP synthesis and showed improved yields compared to the native enzyme (from 26 to 45% against 22% with the WT). Interestingly, none of the newly created

variants displayed yields lower than the corresponding Ala variant, which suggest that starting directly with a round of conservative mutations is most probably the best approach to create new transglycosidases.

As the resulting mutants after this second round of mutagenesis kept high levels of activity, combinations of the most successful mutations were performed.

Mutant recombination

After two rounds of mutagenesis, the variants obtained by mutating conserved residues of 7 out of 14 positions showed an increase in transglycosylation yields, therefore 21 recombination mutants, resulting from the combination of mutations S547A, N334T, W336F, D267N, W411F, K476Q and W545F, were created. The W545F-N334S double mutant was added to this set. Kinetic parameters k_{cat}/K_M as well as GalGal-*p*NP synthesis yield of the resulting mutants are summarized in Table **III**.

Among the 22 newly created double mutants, only 6 were active enough to allow quantification of GalGal-*p*NP synthesis (Table **III**), and these 6 mutants always contained W545F and/or N334T mutations, which are the least deleterious to the k_{cat}/K_M parameter among those used for recombination (Tables **I** and **II**). It can be concluded from these 15 inactive or low active mutants that there is a strong additivity of the decrease in catalytic efficiency, while the beneficial effect on transglycosylation shows a less additive effect: only the double mutant N334T-W545F displayed a higher yield than the single mutant N334T or W545F, with a 57% yield (against 45% for N334T, 35% for W545F and 22% for the WT).

For the best variants (N334T, W545F and N334T-W545F), the kinetics of GalGal-*p*NP synthesis were compared with that of the WT (Fig. **3**). One of the drawbacks of the use of glycosidases for oligosaccharide synthesis is often the secondary hydrolysis of the formed products. However, thanks to the use of an activated donor for which AgaB presents a much higher specificity than for natural substrates (the k_{cat}/K_M towards α-Gal-*p*NP is 500 times

greater than towards raffinose (Merceron et al. 2012)), small amounts of enzyme were used and the transglycosylation product was barely hydrolyzed and maximum yields of product formation remained for a long time (Fig. **3**). Only the N334T variant exhibited a slight secondary hydrolysis compared to the other mutants and the native enzyme.

After these three rounds of mutagenesis, AgaB variants with a strong improvement in their synthetic activity towards GalGal-*p*NP synthesis were obtained. However, as 4-nitrophenyl groups are hard to remove without breaking the glycosidic bond, this disaccharide is of little interest. Thus, a study of transglycosylation towards an acceptor bearing a thiophenyl group instead was undertaken, as this protecting group has proved successful in chemo-enzymatic synthesis (D'Almeida et al. 2009; Marton et al. 2008).

Effect of changes in donors and acceptors on synthetic yields of transglycosylation with AgaB variants

The closest analogue of α-Gal-*p*NP containing a thiophenyl group is phenyl 1-thio-α-D-galactopyranoside. However, given the low availability of this compound, the more affordable phenyl 1-thio-β-D-galactopyranoside (Gal-S-Ph) was used. Moreover, this compound made it possible to assess whether the N334T variant could use acceptors with a β-configuration as well as with an α-configuration. The N334T variant was chosen over the N334T-W545F double mutant for the sake of enzyme consumption, as it gives the best transglycosylation yield after the N334T-W545F variant, but needs 35-fold less enzyme to reach reaction completion (Table 4). Transglycosylation yields of phenyl α-D-galactopyranosyl-(1,6)-1-thio-β-D-galactopyranoside (GalGal-S-Ph) synthesis are given in Table **IV**. Both native AgaB and the N334T variant were able to catalyze the synthesis of GalGal-S-Ph, with a net increase from 14 to 33% with the latter, which demonstrates that newly created variants keep the promiscuity of the native enzyme towards acceptors. A comparative study with the α-galactosynthase approach (Cobucci-Ponzano et al. 2010) has been also performed: the D478G

AgaB variant (equivalent to the D327G TmαGal glycosynthase variant) was produced and tested with β-Gal-N_3 as donor. It showed an interesting increase in synthetic yields compared to the native enzyme (from 14 to 23%, Table **IV**). Nevertheless, the α-transgalactosidase N334T variant displayed higher yields (33 against 23%), while consuming 42-fold less enzyme. When using Gal-S-Ph instead of α-Gal-*p*NP as an acceptor, a drop of one-third in synthetic yields with both N334T and the native enzyme was observed. This decrease is mainly due to competition with the GalGal-*p*NP synthesis. This competition could be turned towards synthesis by adding α-Gal-*p*NP gradually, which resulted in an increase from 33 to 54% of GalGal-S-Ph synthesis yield when the N334T variant was used. In order to go further in avoiding competition with the GalGal-*p*NP synthesis, α-Gal-F was used as donor instead of α-Gal-*p*NP (Andre et al. 2001). This resulted in a drastic increase in transglycosylation yields from 14 to 37% with the native enzyme and from 33 to 62% with the N334T variant (Table **IV**).

DISCUSSION

With the best N334T AgaB variant, one-step glycosidic bond formation with high yields (from 45 to 62%) with unprotected sugars and moderate enzyme consumption was achieved, which makes this variant suitable for bioconversion uses. It also competes favorably with the glycosynthase D478G variant as it provides higher transglycosylation yields (62 against 23%), while using 151-fold less enzyme (Table **IV**). Thus, this approach, which targets highly conserved amino acids around the active site of AgaB, is very efficient for obtaining transglycosidase mutants rapidly and avoiding the tedious screening procedures imposed by random evolution approaches (Koné et al. 2009). Furthermore, the initial approach (Teze et al. 2014) has been improved, as we demonstrate in this work that one round of conservative

mutations allowed a fast access to new transglycosidases while avoiding drastic drops in activity. It is particularly interesting to note that the best variant, N334T, was obtained by mutating a second-shell residue, showing that very slight modifications may provoke strong effects on the balance of transglycosylation over hydrolysis. The addition of a second mutation induced a strong deleterious effect on the activity of all the 22 double mutants produced (Table **III**). On the other hand, the effect on transglycosylation yields seems less predictable: among the six double mutants that exhibited an activity high enough to study the transglycosylation yields, two displayed an interesting increase compared to the single mutants (N334T-W545F (+12%) and W411F-W545F (+9%)), two showed slight yield modifications (N334S-W545F (+1%) and N334T-W411F (-1%)) and two presented a drastic drop in yields (N334T-K476Q (-7%) and W336F-W545F (-9%)) (Table **III**).

Mutants were screened for the best transglycosidase activity by following the GalGal-*p*NP synthesis. As this product is of limited interest as an intermediate in oligosaccharide synthesis, we also demonstrated that by switching to other acceptors, mutants kept their improved transglycosylation properties since mutations did not affect the +1 and +2 acceptor sites. Moreover, by using alternative donors such as α-Gal-F, the transglycosylation yields improved due to the absence of the competing GalGal-*p*NP synthesis. As the α-Gal-F is a more activated donor than the α-Gal-*p*NP, the enzyme consumption was decreased by 3.6 to 5-fold, which also contribute to the transglycosylation yields increase due to the reduction of secondary hydrolysis.

The spatial arrangement of the mutated residues can be seen in the published structure of AgaB (Merceron et al. 2012) (Fig. 2). The N480 residue is located in the neighborhood of the acceptor subsite, which may explain the drop in transglycosylation yields obtained with the variant N480D as some interactions with an acceptor moiety could be missing. However, all the most beneficial mutations for transglycosidase activity are located farther from the

catalytic residues, at the bottom of the -1 site cavity (mutations on residues S547, N334, W545 and K476 (Fig. 2)). Thus, structural information does not give any clues for understanding the effect of mutation on the transglycosylation/hydrolysis ratio. To date, a possible explanation is that highly conserved amino acids in the strict vicinity of the glycosyl subsite are involved in stabilizing the transition states of the catalyzed reactions, i.e. glycosylation, hydrolysis and transglycosylation. Thus, mutating these residues might decrease the transition state stabilization of these three reactions (resulting in a drop in kinetic parameters), but their contributions to the stabilization of each of these three transition states is most probably not identical. Thus, the finely tuned equilibrium between these three reactions would be disturbed and it is likely that variants presenting a balance slightly shifted towards transglycosylation would be obtained (as well as those with a balance shifted towards hydrolysis). Indeed, on 14 mutated positions, the substitution of 4 of them led to variants with a decreased synthetic ability, 7 led to an increased ability and 3 gave no quantifiable yields.

As the rationalization of the results proved unsuccessful, it seems to be essentially impossible to identify by rational design the key positions to mutate. However, the methodology proposed in this work to create improved transglycosidases is much more rapid and efficient than directed evolution, which requires screening procedures on a large library of mutants. As it has been proved efficient for creating both α- (this work) and β-transglycosidases (Teze et al. 2014), it may be a general route to obtain transglycosidases rapidly.

MATERIALS AND METHODS

α-Gal-*p*NP and β-Gal-S-Ph were purchased from Carbosynth, oligonucleotides from Eurofins, and enzymes from Fermentas. Other laboratory reagents were purchased from Sigma-Aldrich unless otherwise indicated and used without further purification.

Site-directed mutagenesis

The QuikChange site-directed mutagenesis kit from Stratagene was used to make point mutations without any modifications: during PCR, Pfu polymerase extends the primers and generates a nicked plasmid with the desired mutation incorporated. Then the product is treated with Dpn I endonuclease, which digests the parental DNA. The nicked plasmid can then be transformed into XL-1 Blue competent cells. The 1.3-kb DNA fragments encoding the mutant AgaB enzymes were sequenced in both forward and reverse directions. Primers used for the site-directed mutagenesis are listed in Supplementary Data (Tables **TS1, TS2**).

Protein expression and purification

Recombinant strains expressing AgaB genes were grown in 1 L of LB medium at 37°C overnight, centrifuged and resuspended in 35 mL of lysis buffer (0.1 M phosphate, pH 8, 0.5 M NaCl, 10 mM imidazole, 10 $\mu g.mL^{-1}$ DNAse I). After sonication and centrifugation, 6xHis-tagged proteins were purified by immobilized ion metal-affinity chromatography (IMAC): 250 µL of Ni-NTA Superflow (Qiagen) was added to the supernatant and stirred for 1 h at 4°C, then loaded onto a 10 mL column. The column was washed using 25 mL of washing buffer (0.1 M phosphate, pH 8, 0.5 M NaCl, 25 mM imidazole), then 5 mL of elution buffer was added (0.1 M phosphate, pH 8, 0.5 M NaCl, 250 mM imidazole). Purity of the final product was checked by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Enzyme concentrations were determined by UV absorbance at 280 nm ($Abs_{0.1\%}$ = 2.18 and 2.07 for the W mutants) using NanoDrop 1000 (Thermo Scientific)

Bioinformatic analysis

GH36 sequences were retrieved from the Uniprot database (Apweiler et al. 2004) and multiple sequence alignment (MSA) was performed with CLUSTALW (Chenna et al. 2003). To assess the degree of amino acid residue conservation, 205 different sequences more than

30% identical to AgaB after the MSA were considered. Hereafter, amino acid residues found conserved in more than 99.5% of these sequences are termed “highly conserved”.

Kinetic studies

All kinetic studies were performed with a TECAN Infinite M1000 spectrophotometer at 30°C in phosphate buffer 0.1 M, pH 7. Initial reaction rates were calculated from the slope of the first-order plot of product concentration (*p*NP) against reaction time, then initial reaction rates were plotted against substrate concentration and the resulting curves were fitted to the hyperbolic equation $v=Vmax*[S]*[E]/(K_M+[S])$ using Origin 7.0 (OriginLab) to obtain k_{cat} and K_M parameters. For most mutants hyperbolic curves were not observed since the kinetic of pNP release was the result of both glycosylation, hydrolytic and transglycosylation reactions. The k_{cat}/K_m parameter was evaluated at low substrate concentrations, when kinetic can be approximate to the equation: $v=k_{cat}/K_M*[E]*[S]$. k_{cat}/K_M is then an apparent second order rate constant (referred as the “catalytic efficiency”), which is determined by the initial slope of the curve V/[E] = f([S]). All measurements were made in triplicate with at least 12 different substrate concentrations ranging from 9.8 µM to 20 mM.

Transglysosylation studies

Screening of the transglycosylation efficiency of AgaB variants was achieved by following GalGal-*p*NP synthesis (Spangenberg et al. 2000). Synthetic yields with 10 mM α-Gal-*p*NP were determined by capillary electrophoresis (Beckman P/ACE System 5000 with an uncoated fused silica capillary, 47 cm) as previously described (Teze et al. 2013a; Teze et al. 2014). 200 µL of media containing 10 mM imidazole (used as an internal standard), 10 mM α-Gal-*p*NP and the enzyme (enzyme concentrations are indicated in Tables **II**, **III**, and **IV**) was incubated at 30°C in phosphate buffer 0.1 M, pH 7, within the capillary electrophoresis apparatus and analyzed every 24 min for 2.8 h then each hour for the remaining 7 h. If specified, α-galactosyl-fluoride (α-Gal-F) 10 mM or β-galactosyl-azide (β-Gal-N_3) 10 mM

was used as a donor instead of α-Gal-*p*NP. If specified, 1-thio-β-D-galactopyranoside (Gal-S-Ph) 10 mM was used as an acceptor, in which case the given transglycosylation yields refer to the synthesis of phenyl-(α-D-galactopyranosyl)-(1,6)-1-thio-β-D-galactopyranoside (GalGal-S-Ph). Separations were performed at 17 kV with 50 mM Borax, pH 9.5, as running buffer. Donors (except α-Gal-F), acceptor and products were detected by UV absorbance at 214 nm and quantified by comparison with imidazole. Four experiments with the native enzyme and the N334A mutant gave maximum yields within a percent precision. As an example, all 15 electropherograms obtained for the best variant (N334T) study are presented in Figures **S1** to **S15** (Supplementary Data). Due to its high thermostability, native enzyme and also AgaB variants are fully stable in the timescale of experiments.

FUNDING

This work was supported by the Pays de la Loire and Bretagne regions (Glyconet Network) and the French Agence Nationale de la Recherche [ANR-10-BLAN-0718].

ACKNOWLEDGMENT

The authors are grateful to Genadi Chikov, Corentin Léger, Amina Fateh and Claude Solleux for their help and technical support.

SUPPORTING INFORMATION AVAILABLE

Primers used for site-directed mutagenesis are listed in Tables **TS1** and **TS2**. Fifteen electropherograms obtained for GalGal-*p*NP synthesis study of the best variant (N334T) are presented in Figures **S1** to **S15.** Kinetic behaviors of mutants W336F and C526A are presented in Figure **S16.** Kinetic behaviors of mutants W411F and D367N are presented in Figure **S17** (Supplementary Data).

ABBREVIATIONS

AgaB: *Geobacillus stearothermophilus* α-galactosidase B; TmαGal: *Thermotoga maritima* α-galactosidase; GH36: Glycosyl Hydrolase family 36; WT: wild type; GalGal-*p*NP: 4-nitrophenyl-(α-D-galactopyranosyl)-(1,6)-α-D-galactopyranoside; α-Gal-*p*NP: 4-nitrophenyl-α-D-galactopyranosyl; α-Gal-F: α-Galactosyl-fluoride; β-Gal-N_3: β-galactosyl-azide; Gal-S-Ph: 1-thio-β-D-galactopyranoside; GalGal-S-Ph: Phenyl-(α-D-galactopyranosyl)-(1,6)-1-thio-β-D-galactopyranoside.

REFERENCES

Andersen JM, Barrangou R, Abou Hachem M, Lahtinen SJ, Goh YJ, Svensson B, and Klaenhammer TR. 2013. Transcriptional analysis of oligosaccharide utilization by *Bifidobacterium lactis* Bl-04. *BMC Genomics* 14:312.

Andre C, Spangenberg P, and Rabiller C. 2001. In Situ 19 F NMR spectroscopy study of enzymatic transglycosylation reactions using α-D-aldohexopyranosyl fluorides as donors and acceptors. *Tetrahedron: Asymmetry* 12: 779–783.

Apweiler R, Bairoch A, Wu CH, Barker WC, Boeckmann B, Ferro F, Gasteiger E, et al. 2004. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* 32 (Database issue): D115–9.

Astronomo RD and Burton DR. 2010. Carbohydrate vaccines: developing sweet solutions to sticky situations? *Nature Reviews Drug Discovery* 9: 308-324.

Bobrov KS, Borisova AS, Eneyskaya EV, Ivanen DR, Shabalin K A, Kulminskaya AA, and Rychkov GN. 2013. Improvement of the efficiency of transglycosylation catalyzed by α galactosidase from *Thermotoga maritima* by protein engineering. *Biochemistry-Moscow* 78: 1112–1123.

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, and Henrissat B. 2009. The Carbohydrate-Active EnZymes Database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res.* 37 (Database issue): D233–8.

Cervera-Tison M, Tailford LE, Fuell C, Bruel L, Sulzenbacher G, Henrissat B, Berrin JG, Fons M, Giardina T, and Juge N. 2012. Functional analysis of family GH36 α-galactosidases from *Ruminococcus gnavus* E1: insights into the metabolism of a plant oligosaccharide by a human gut symbiont. *Appl. Environ. Microbiol.* 78: 7720–32

Champion E, Andre I, Moulis C, Boutet J, Descroix K, Morel S, Monsan P, Mulard LA, and Remaud-Simeon M. 2009. Design of α-transglucosidases of controlled specificity for programmed chemoenzymatic synthesis of antigenic oligosaccharides. *J. Am. Chem. Soc.* 131: 7379–7389.

Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, and Thompson JD. 2003. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.* 31 (13): 3497–3500.

Cobucci-Ponzano B, Zorzetti AC, Strazzulli SC, Bedini E, Corsaro MM, Comfort DA, Kelly RM, Rossi M, and Moracci M. 2010. A Novel α-D-galactosynthase from *Thermotoga Maritima* converts β-D-galactopyranosyl azide to α-galacto-oligosaccharides. *Glycobiology* 21 (4): 448–456.

Cobucci-Ponzano B, Perugino G, Strazzulli A, and Moracci M. 2012. Thermophilic glycosynthases for oligosaccharides synthesis. *Cellulases*. 1st ed. Vol. 510. Elsevier Inc.

D'Almeida A, Ionata M, Tran V, Tellier C, Dion M, and Rabiller C. 2009. An expeditious and efficient synthesis of β-D-galactopyranosyl-(1-3)-D-N-acetylglucosamine (lacto-N-biose) using a glycosynthase from *Thermus Thermophilus* as a catalyst. *Tetrahedron: Asymmetry* 20: 1243-1246.

Dion M, Nisole A, Spangenberg P, Andre C, Glottin-fleury A, Mattes R, Tellier C and Rabiller C. 2001a. Modulation of the regioselectivity of a *Bacillus* alpha-galactosidase by directed evolution. *Glycoconjugate J.* 18, 215–223.

Dion M, Osanjo G, Andre C, Spangenberg P, Rabiller C and Tellier C. 2001b. Identification by saturation mutagenesis of a single residue involved in the α-galactosidase AgaB regioselectivity. *Glycoconjugate J.* (18), 457–464.

Feng HY, Drone J, Hoffmann L, Tran V, Tellier C, Rabiller C, and Dion M. 2005. Converting a β-glycosidase into a β-transglycosidase by Directed Evolution. *J. Biol. Chem.* 280 (44): 37088–37097.

Fredslund F, Hachem MA, Larsen RJ, Sørensen PG, Coutinho PM, Leggio LL and Svensson B. 2011. Crystal structure of α-galactosidase from *Lactobacillus acidophilus* NCFM Insight into tetramer formation and substrate binding. *J. Mol. Biol.* 412: 466–480.

Gantt RW, Peltier-Pain P, Singh S, Zhou M, and Thorson JS. 2013. Broadening the scope of glycosyltransferase-catalyzed sugar nucleotide synthesis. *Proc Natl Acad Sci USA* 110 (19): 7648–7653.

Ghazarian H, Idoni B, and Oppenheimer SB. 2011. A glycobiology review: carbohydrates, lectins and implications in cancer therapeutics. *Acta Histochem* 113: 236–247.

Hachem MA, Fredslund F, Andersen JM, Larsen RJ, Majumder A, Ejby M, Van Zanten G, Lahtinen SJ, Barrangou R, Klaenhammer TR, Jacobsen S, Coutinho PM, Leggio LL and Svensson B. 2012. Raffinose family oligosaccharide utilisation by probiotic bacteria: insight into substrate recognition, molecular architecture and diversity of GH36 α-galactosidases. *Biocatal Biotransform* 30: 316-325.

Hart GW, and Copeland RJ. 2010. Essay glycomics hits the big time. *Cell* 143: 672–676.

Hehre EJ, Brewer CF and Genghof DS. 1979. Scope and mechanism of carbohydrase action. Hydrolytic and nonhydrolytic actions of beta-amylase on alpha- and beta-maltosyl fluoride. *J. Biol. Chem*. 254, 5942–5950.

Honda Y, Fushinobu S, Hidaka M, Wakagi T, Shoun H, Taniguchi H, and Kitaoka M. 2008. Alternative strategy for converting an inverting glycoside hydrolase into a glycosynthase. *Glycobiology*, 18(4), 325–30.

Honda Y and Kitaoka M 2006. The first glycosynthase derived from an inverting glycoside hydrolase. *J. Biol. Chem.*, 281(3), 1426–1431.

Irague R, Massou S, Moulis C, Saurel O, Milon A, Monsan P, Remaud-Simeon M, Portais JC, and Potocki-Veronese G. 2011. NMR-based structural glycomics for high-throughput screening of carbohydrate-active enzyme specificity. *Anal. Chem.* 83: 1202–1206.

Kelly RM, Leemhuis H, Rozeboom HJ, van Oosterwijk N, Dijkstra BW and Dijkhuizen L. 2008. Elimination of competing hydrolysis and coupling side reactions of a cyclodextrin glucanotransferase by directed evolution. *Biochem J.* 413(3), 517–25.

Koné FMT, Le Béchec M, Sine JP, Dion M, and Tellier C. 2009. Digital screening methodology for the directed evolution of transglycosidases. *Protein Eng. Des. Sel.* 22: 37–44.

Lairson LL, Henrissat B, Davies GJ, and Withers SG. 2008. Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* 77: 521–555.

Leemhuis H, Dijkstra BW and Dijkhuizen L. 2002. Mutations converting cyclodextrin glycosyltransferase from a transglycosylase into a starch hydrolase. *FEBS Lett.* 514(2-3), 189–92.

Lichtenstein RG, and Rabinovich GA. 2013. Glycobiology of cell death: when glycans and lectins govern cell fate. *Cell Death Differ* 2869.

Lombard V, Ramulu HG, Drula E, Coutinho PM, and Henrissat B. 2014. The Carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42:D490–D495.

Macher BA, and Galili U. 2008. The Galα1,3Galβ1,4GlcNAc-R (α-Gal) epitope: a carbohydrate of unique evolution and clinical relevance. *Biochim Biophys Acta*. 1780: 75-88.

Mackenzie LF, Wang Q, Warren RAJ, and Withers SG. 1998. Glycosynthases: mutant glycosidases for oligosaccharide synthesis. *J. Am. Chem. Soc.* 120: 5583-5584.

Malet C, and Planas A. 1998. From B-glucanase to β-glucansynthase: Glycosyl transfer to α-glycosyl fluorides catalyzed by a mutant endoglucanase lacking its catalytic nucleophile. *FEBS Lett.* 440: 208–212.

Marton Z, Tran V, Tellier C, Dion M, Drone J, and Rabiller C. 2008. Engineering of glucoside acceptors for the regioselective synthesis of β-(1-->3)-disaccharides with glycosynthases. *Carbohydr. Res.* 343: 2939-2946.

Merceron R, Foucault M, Haser R, Mattes R, Watzlawick H, and Gouet P. 2012. The Molecular mechanism of thermostable α-galactosidases AgaA and AgaB explained by X-ray crystallography and mutational studies. *J. Biol. Chem.* 287: 39642–39652.

Nakai H, Baumann M, Petersen B, Westpha Y, Hachem M, Dilokpimol A, Duus, JØ, Schols HA and Svensson B. 2010. *Aspergillus nidulans* α-galactosidase of glycoside hydrolase family 36 catalyses the formation of α-galacto-oligosaccharides by transglycosylation. *FEBS J.* 277: 3538–3551.

Osanjo G, Dion M, Drone J, Solleux C, Tran V, Rabiller C, and Tellier C. 2007. Directed evolution of the alpha-L-fucosidase from *Thermotoga maritima* into an α-L-transfucosidase. *Biochemistry* 46: 1022–1033.

Peterbauer T, Karner U, Mucha J, Mach L, Jones DA, Hedley CL, and Richter A. 2003. Enzymatic control of the accumulation of verbascose in pea seeds. *Plant Cell Enviro*n. 26: 1385–1391.

Peterbauer T, Mach L, Mucha J, and Richter A. 2002a. Functional expression of a cDNA encoding pea (*Pisum Sativum L* .) Raffinose synthase , partial purification of the enzyme from maturing seeds , and steady-state kinetic analysis of raffinose synthesis. *Planta* 215: 839–846.

Peterbauer T, Mucha J, Mach L, and Richter A. 2002b. Chain elongation of raffinose in pea seeds. Isolation, characterization, and molecular cloning of mutifunctional enzyme catalyzing the synthesis of stachyose and verbascose. *J. Biol. Chem*. 277: 194–200.

Pizzut-Serin S, Potocki-Veronese G, Van Der Veen BA, Albenne C, Monsan P, and Remaud-Simeon M. 2005. Characterisation of a novel amylosucrase from *Deinococcus radiodurans*. *FEBS Lett.* 579: 1405–1410.

Plou FJ, Segura AG, and BallesterosA. 2007. Application of glycosidases and transglycosidases in the synthesis of oligosaccharides. In *Industrial Enzyme,* 141–157.

Spangenberg P, Andre C, Dion M, Rabiller C, and Mattes R. 2000. Comparative study of new α-galactosidases in transglycosylation reactions. *Carbohydr. Res.* 329: 65–73.

Teze D, Dion M, Daligault F, Tran V, André-Miral C, and Tellier C. 2013a. Alkoxyamino glycoside acceptors for the regioselective synthesis of oligosaccharides using glycosynthases and transglycosidases. *Bioorg. Med. Chem. Lett.* 23: 448-451.

Teze D, Hendrickx J, Czjzek M, Ropartz D, Sanejouand YH, Tran V, Tellier C, and Dion M. 2014. Semi-rational approach for converting a GH1 β-glycosidase into a β-transglycosidase. *Protein Eng. Des. Sel.* 27: 13–19.

Teze D, Hendrickx J, Dion M, Tellier C, Woods VL, Tran V, and Sanejouand YH. 2013b. Conserved water molecules in family 1 glycosidases: a DXMS and molecular dynamics study. *Biochemistry* 52: 5900–5910.

Wada J, Honda Y, Nagae M, Kato R, Wakatsuki S, Katayama T, Taniguchi H, Kumagai H, Kitaoka M, and Yamamoto K. 2008. 1,2-alpha-L-Fucosynthase: a glycosynthase derived from an inverting alpha-glycosidase with an unusual reaction mechanism. *FEBS Letters*, *582*(27), 3739–43.

Wang CC, Lee JC, Luo SY, Kulkarni SS, Huang YW, and Lee CC. 2007. Regioselective one-pot protection of carbohydrates. *Nature* 446: 896–899.

Wang Q, Graham RW, Trimbur D, Warren RAJ, and Withers SG. 1994. Changing enzymatic reaction mechanisms by mutagenesis: converting a retaining glucosidase to an inverting enzyme. *J. Am. Chem. Soc.* 116: 11594–11595.

Wang H, Ma R, Shi P, Xue X, Luo H, Huang H, Bai Y, Yang P, and Yao, B. 2014. A New α-galactosidase from *Thermoacidophilic alicyclobacillus sp.* A4 with wide acceptor specificity for transglycosylation. *Appl. Biochem. Biotechnol.* 174: 328–38

Zhang C, Griffith BR, Fu Q, Albermann C, Fu X, Lee IK, Li L, and Thorson JS. 2006. Exploiting the reversibility of natural product glycosyltransferase-catalysed reactions. *Science*. 313: 1291-1294.

Zhu X, and Schmidt RR. 2009. New principles for glycoside-bond formation. *Angew. Chem. Int. Ed.* 48: 1900-1934.

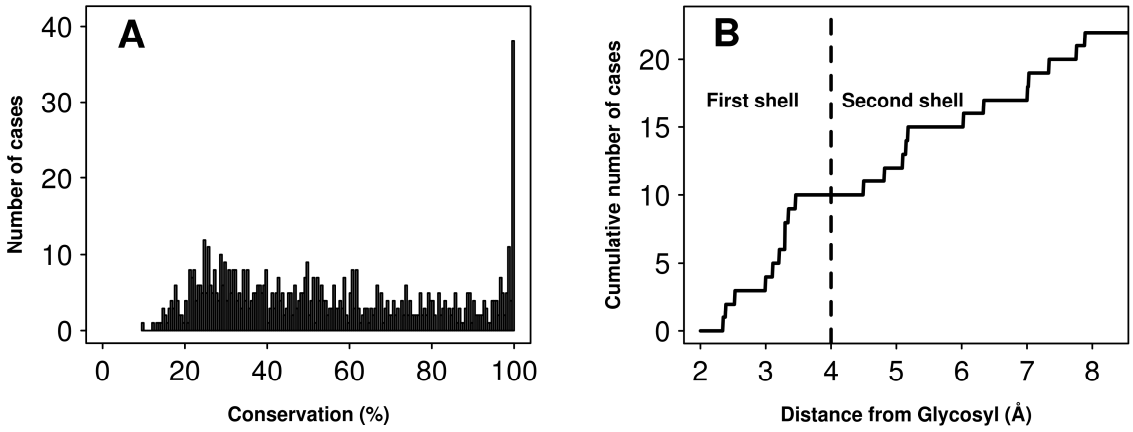

Fig. 1. Amino acid conservation within the GH36 family. (**A**) Histogram of the population of amino acids as a function of their percentage of conservation. Each population corresponds to an interval of 0.5% of conservation and the statistics are based on 205 sequences. (**B**) Highly conserved amino acid population as a function of the distance from the glycosyl, based on the AgaB structure (PDB code 4FNQ).

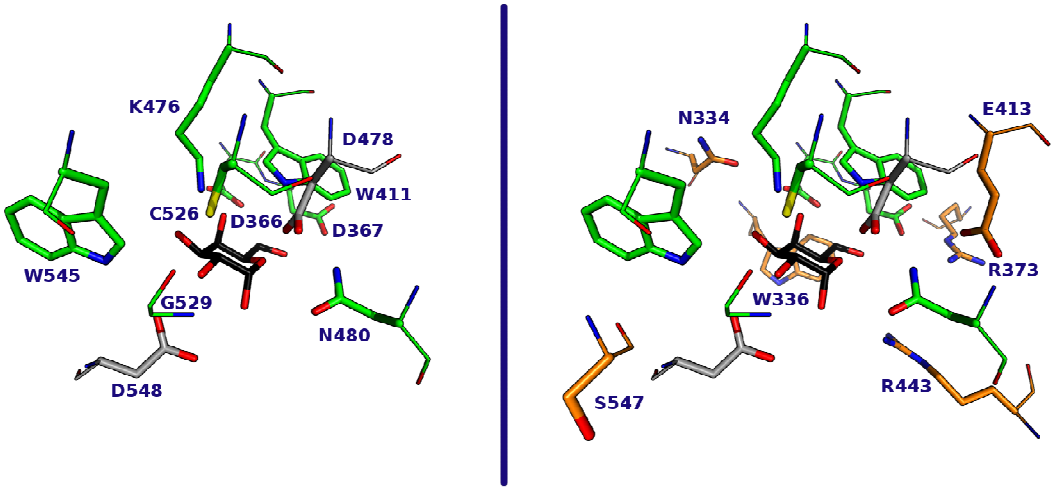

Fig. 2. Localization of highly conserved amino acids within the AgaB active site (PDB code 4FNQ). Catalytic and galactose carbons are in grey and black, respectively. Left: First-shell residues (in green). Right: Second-shell residues. Only Second-shell residues (in orange) are labeled. This figure was created with PyMol.

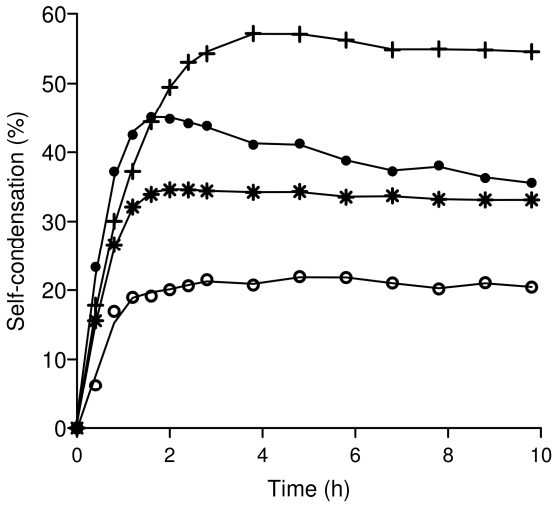

Fig. 3. Time course of GalGal-*p*NP synthesis in the presence of α-Gal-*p*NP 10 mM as substrate. Percentages are calculated with reference to initial substrate concentration using the following AgaB variants: WT (empty circles); W545F (stars); N334T (filled circles); N334T-W545F (plus symbols).

Table I. Kinetics and yields of GalGal-*p*NP synthesis of the AgaB variants obtained by Ala-scanning.

Enzyme	k_{cat}/K_M ($s^{-1}.mM^{-1}$)	GalGal-*p*NP (%)	Enzyme concentration ($\mu g.mL^{-1}$)[a]
WT	200 ± 4	22	1.0
G529A	3.06 ± 0.03	21	5.5
W545A	(6 ± 0.3) x 10^{-2}	34	790
C526A	0.33 ± 0.02	19	34
S547A	0.70 ± 0.05	31	29
N334A	0.57 ± 0.02	38	170
R373A	0.67 ± 0.04	2	11
W411A[b]	(1.3 ± 0.1) x 10^{-2}	ND[c]	ND[c]
E413A[b]	(1.2 ± 0.1) x 10^{-2}	ND[c]	ND[c]
K476A	(3.1 ± 0.1) x 10^{-3}	ND[c]	ND[c]
W336A; R443A	ND[d]	ND[c]	ND[c]
N480A; D366A; D367A	NDA[e]	ND[c]	ND[c]

[a]Enzyme concentration necessary to reach the maximum yield of GalGal-*p*NP formation within 4 hours.
[b] W411A (k_{cat}= (1.9 ± 0.02) x 10^{-2} s^{-1}; K_M =1.5 ± 0.07 mM) and E413A (k_{cat}= (3.0 ± 0.1) x 10^{-2} s^{-1}; K_M =2.4 ± 0.2 mM) display a classic Michaelis-Menten behavior.
[c]Not Determined. Activities of these enzymes were too low to measure GalGal-*p*NP formation.
[d]W336A and R443A presented a detectable activity, but too low to be measured.
[e]No Detectable Activity.

Table II. Kinetics and yields of GalGal-*p*NP synthesis of AgaB variants presenting "conservative mutations".

Enzyme	k_{cat}/K_M ($s^{-1}.mM^{-1}$)	GalGal-*p*NP (%)	Enzyme concentration ($\mu g.mL^{-1}$)[a]
WT[b]	200 ± 4	22	1.0
N334D	6.6 ± 0.2	44	3.0
N334S	2.4 ± 0.02	39	11.5
N334T	5.4 ± 0.03	45	10.5
W336F	0.38 ± 0.005	27	205
D367N[b]	0.23 ± 0.005	26	55
W411F[b]	0.75 ± 0.02	27	23
K476Q	$(6.7 \pm 0.4) \times 10^{-2}$	36	115
N480D	$(7.3 \pm 0.3) \times 10^{-2}$	14	140
W545F[b]	2.7 ± 0.02	35	7.5
R443Q; D366N; E413Q	NDA[c]	NDA	NDA

[a]Enzyme concentration necessary to reach the maximum amount of GalGal-*p*NP within 4 hours.
[b]These AgaB variants present a slight substrate inhibition.
[c]No Detectable Activity.

Table III. Kinetic parameters and yields of GalGal-*p*NP synthesis of the double mutants obtained by recombination.

Enzyme	k_{cat}/K_M ($s^{-1}.mM^{-1}$)	GalGal-*p*NP (%)	Enzyme concentration ($\mu g.mL^{-1}$)[a]
WT[b]	200 ± 4	22	1.0
N334T[b]	5.4 ± 0.03	45	10.5
W545F[b]	2.7 ± 0.02	35	7.5
N334T-W545F	0.21 ± 0.02	57	360
W545F-N334S	0.12 ± 0.004	40	425
W545F-W411F	$(1.1 \pm 0.1) \times 10^{-2}$	44	1.7×10^{3}
N334T-W411F	$(3.6 \pm 0.3) \times 10^{-2}$	44	280
W545F-W336F	$(3.8 \pm 0.1) \times 10^{-2}$	26	1.0×10^{3}
N334T-K476Q	$(1.9 \pm 0.2) \times 10^{-2}$	38	1.4×10^{3}

[a] Enzyme concentration necessary to reach the maximum amount of GalGal-*p*NP within 4 hours.

[b]WT, N334T and W545F were included as a reminder of the effect of single mutations.

Table IV. Yields of GalGal-S-Ph synthesis using AgaB variants with Gal-S-Ph (10 mM) as acceptor and various donors (each at 10 mM).

Enzyme	Donor	GalGal-*p*NP (%)[a]	Transglycosylation yield (%)	Enzyme concentration ($\mu g.mL^{-1}$)[b]
WT	α-Gal-*p*NP	9	14	1.5
WT	α-Gal-F	-	37	0.3
N334T	α-Gal-*p*NP	19	33	18
N334T	α-Gal-*p*NP[c]	9	54	27
N334T	α-Gal-F	-	62	5.0
D478G	β-Gal-N_3	0	23	755

[a]Yields of GalGal-*p*NP synthesis, measured at the time of the maximum yield of the transglycosylation product.

[b]Enzyme concentration necessary to reach the maximum transglycosylation yield within 4 hours.

[c]Instead of having 10 mM of donor at the beginning of the reaction, the same amount of donor was reached by 5 additions during the time course of the reaction.