# Semi-rational approach for converting a GH1 $\beta$-glycosidase into a $\beta$-transglycosidase.

David Teze, Johann Hendrickx, Mirjam Czjzek, David Ropartz, Yves-Henri Sanejouand, Vinh Tran, Charles Tellier, Michel Dion

▶ **To cite this version:**

David Teze, Johann Hendrickx, Mirjam Czjzek, David Ropartz, Yves-Henri Sanejouand, et al.. Semi-rational approach for converting a GH1 $\beta$-glycosidase into a $\beta$-transglycosidase.. Protein Engineering, Design and Selection, 2014, 27 (1), pp.13-9. 10.1093/protein/gzt057 . hal-01020940

**HAL Id: hal-01020940**

**https://hal.sorbonne-universite.fr/hal-01020940v1**

Submitted on 8 Jul 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

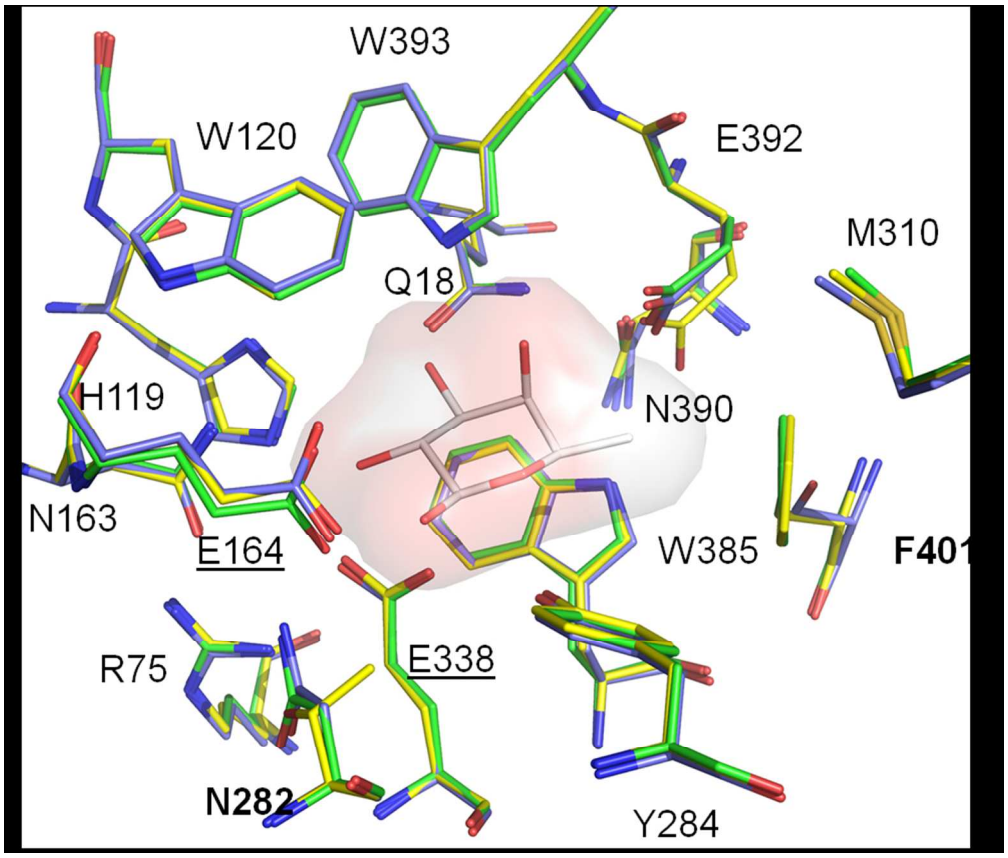
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.


# Semi-rational approach for converting a GH1 beta-glycosidase into a beta-transglycosidase

| | |
|---|---|
| Journal: | *Protein Engineering, Design, and Selection* |
| Manuscript ID: | PEDS-13-0075.R1 |
| Manuscript Type: | Original Article |
| Keywords: | Glycosidase, Mutagenesis, Oligosaccharide, Structure, Transglycosylation |
| | |

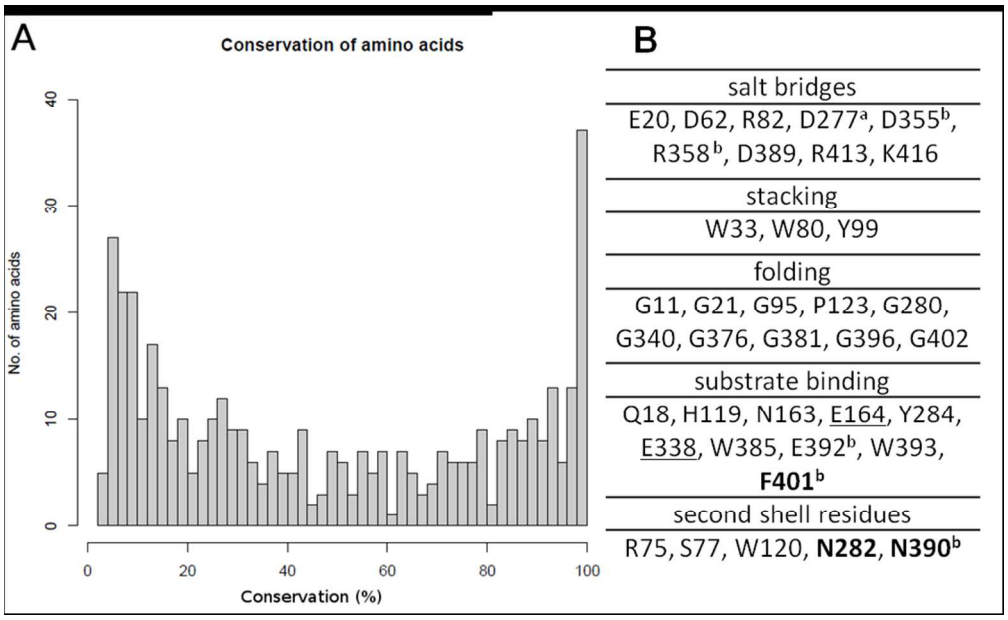
SCHOLARONE™
Manuscripts

For Review




168x142mm (150 x 150 DPI)

view




176x107mm (150 x 150 DPI)

er Review

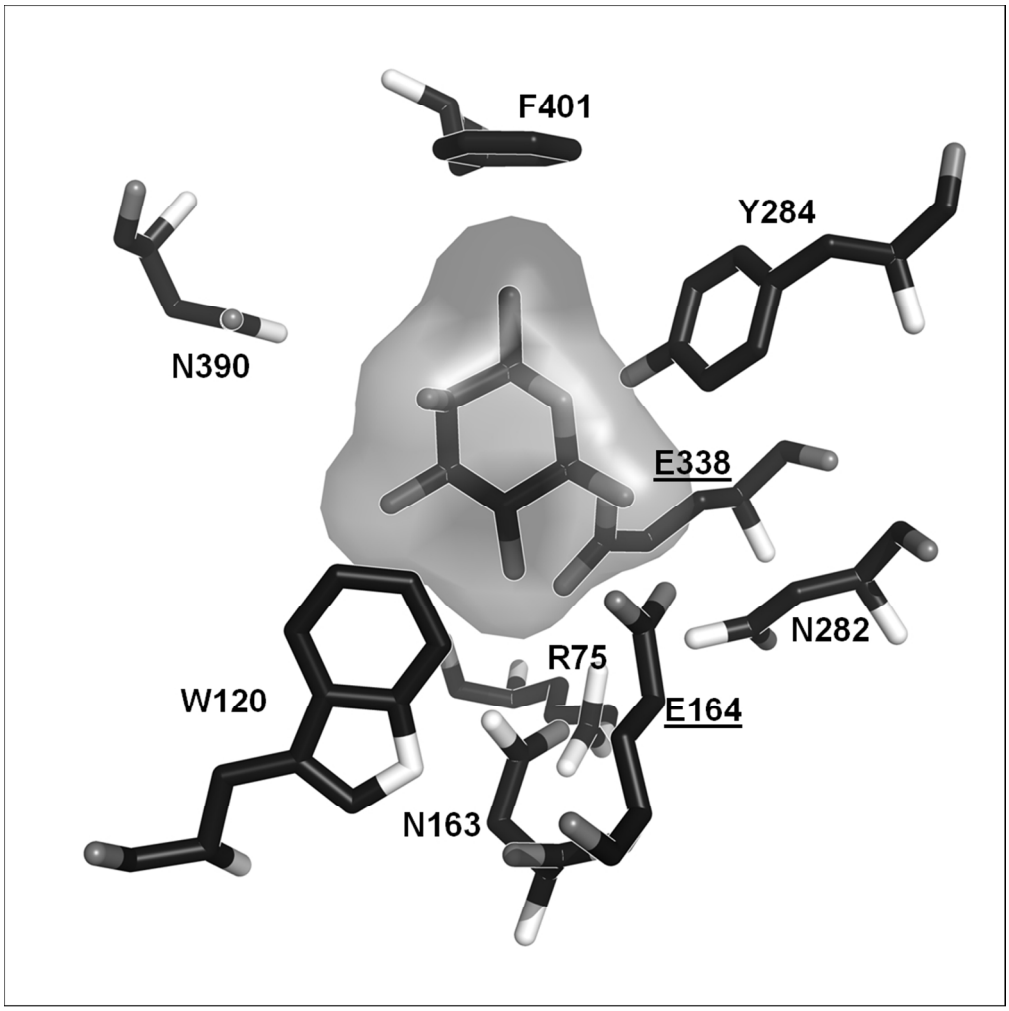
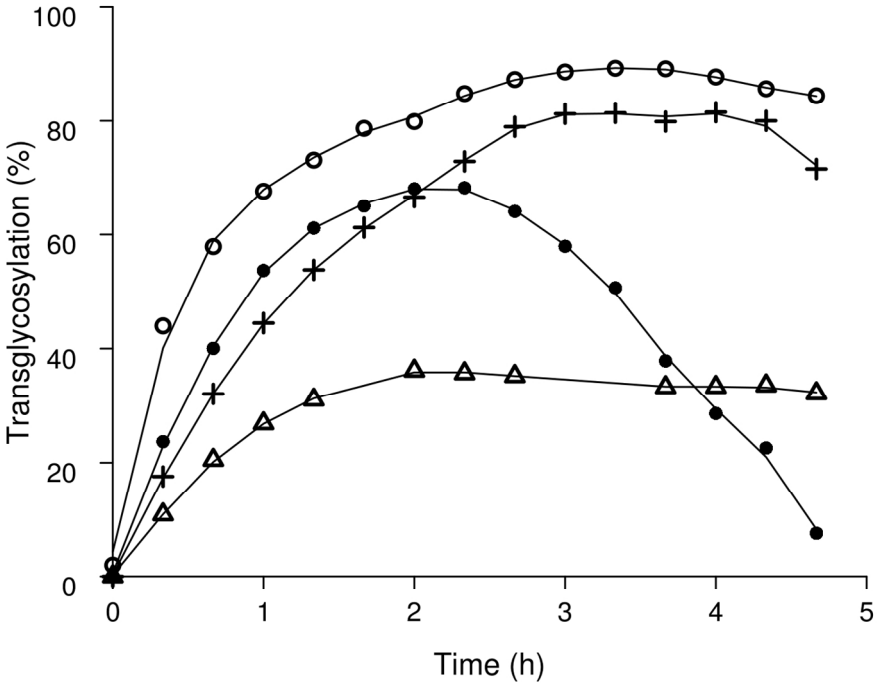



Figure 3. Localization of the targeted residues in native Ttβ-gly (PDB code: 1UG6). Carbons are shown in black, oxygens in grey and nitrogens in white. A β-D-fucose molecule is docked in the active site by superimposition with the 2-deoxy-2-fluoro-D-glucosyl using the glycosyl-enzyme structure of B. polymyxa β-glycosidase (PDB code: 1E4I). Catalytic residue identifiers are underlined. This figure was created with PyMol.
190x190mm (150 x 150 DPI)




eview

# Semi-rational approach for converting a GH1 β-glycosidase into a β-transglycosidase

*David Teze, Johann Hendrickx, Mirjam Czjzek[#], David Ropartz[§], Yves-Henri Sanejouand, Vinh Tran, Charles Tellier and Michel Dion**

Université de Nantes, CNRS FRE n°3478, UFIP, 2 rue de la Houssinière, F-44322 Nantes cedex 03, France.

[#]CNRS UMR 7139, Université Pierre et Marie Curie-Paris 6, Station Biologique, 29682 Roscoff, France

[§]INRA, UR1268 Biopolymers Interactions Assemblies F-44316 Nantes, France

**Corresponding Author**

*Email: michel.dion@univ-nantes.fr; Tel.: 33-2-51-12-56-11

**Running title:** semi-rational design of transglycosidases

For Peer Review

**ABSTRACT**: A large number of retaining glycosidases catalyze both hydrolysis and transglycosylation reactions, but little is known about what determines the balance between these two activities (T/H ratio). We previously obtained by directed evolution the mutants F401S and N282T of *Thermus thermophilus* β-glycosidase (Ttβ-gly, GH1), which display a higher T/H ratio than the WT enzyme. In order to find the cause of these activity modifications, and thereby set up a generic method for easily obtaining transglycosidases from glycosidases, we determined their X-ray structure. No major structural changes could be observed which could help to rationalize the mutagenesis of glycosidases into transglycosidases. However, as these mutations are highly conserved in GH1 β-glycosidases and are located around the -1 site, we pursued the isolation of new transglycosidases by targeting highly conserved amino acids located around the active site. Thus, by single-point mutagenesis on Ttβ-gly, we created four new mutants that exhibit improved synthetic activity, producing disaccharides in yields of 68 to 90% against only 36% when native Ttβ-gly was used. As all of the chosen positions were well-conserved among GH1 enzymes, this approach is most probably a general route to convert GH1 glycosidases into transglycosidases.

**KEYWORDS:**

Glycosidase / Mutagenesis / Oligosaccharide / Structure / Transglycosylation

## INTRODUCTION

Carbohydrates play major roles in biology such as in the inflammation reaction, cell death, cellular interactions, stabilization of proteins, adhesion of pathogenic microorganisms and binding of toxins (Hart & Copeland 2010; Lichtenstein & Rabinovich 2013; Ghazarian et al. 2011). However, their use as therapeutic agents is hindered by their poor availability. In fact, 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.





For Peer Review

2007), the assembly of oligosaccharides remains a substantial challenge. Because of the similar reactivities of hydroxyl groups, glycosidic bond formation requires meticulous control of both regio- and stereoselectivity, which involves extensive protecting group chemistry and results in low synthetic yields. Consequently, alternative synthetic strategies have been explored, in particular 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 nucleotide for Leloir-type transferases or phosphate sugars for non-Leloir types) (Lairson et al. 2008) even if a recent approach seems promising (Gantt et al. 2013). 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 (aspartate or glutamate) by a neutral amino acid (glycine, serine, cysteine or alanine) 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 & Planas 1998; Cobucci-ponzano et al. 2012). A second approach, which was developed recently in our laboratory, was based on the directed evolution of glycosidases into transglycosidases (Feng et al. 2005; Osanjo et al. 2007). These two approaches were validated on *Thermus thermophilus* β-glycosidase (Tβ-gly), an enzyme which belongs to the glycoside hydrolase family 1 (GH1) (Henrissat 1991; Cantarel et al. 2009), for which both efficient glycosynthases and transglycosidases were obtained (Feng et al. 2005; Drone et al. 2005; Marton et al. 2008). Directed evolution of glycosidases was carried out to obtain mutant enzymes having almost lost their hydrolytic activity while keeping significant transglycosylation activity (Feng et al. 2005; Osanjo et al. 2007). The



http://www.peds.oupjournals.org

For Peer Review

resulting mutants have an improved transglycosylation/hydrolysis ratio (T/H), and have similar efficiencies to glycosynthase enzymes (Teze et al. 2013a; Feng et al. 2005; Osanjo et al. 2007). The most challenging part of directed evolution is the screening of large mutant libraries. This has recently been overcome, however, by the development of a digital screening procedure which can discriminate, directly on colonies, the mutant enzymes presenting a better T/H balance (Koné et al. 2009).

Although molecular modeling was tentatively applied to explain why these evolved mutants have improved transglycosidase activities (Feng et al. 2005; Osanjo et al. 2007; Teze et al. 2013b), we could not determine precisely the molecular events leading to the improved T/H ratio. This knowledge would be interesting from a fundamental point of view to understand how glycoside hydrolases regulate the partitioning of the glycosyl-enzyme between hydrolysis and transfer products (Scheme 1). Moreover, it would be useful to set up a rational approach for the design of glycosidases into transglycosidases. In this way, by targeting only a few amino acids, library sizes could be reduced and this approach could be rapidly extended to other glycosidase families.

Therefore, the aim of this study was to investigate the molecular basis of the improved transglycosylation activity of two single mutants of Ttβ-gly: F401S and N282T (Feng et al. 2005) by determining their X-ray structures. Since no major structural changes were observed, bioinformatic studies were also undertaken within the CAZY family 1 (GH1), which revealed that these mutated positions belong to highly conserved positions around the -1 site. Several other positions that were identified as highly conserved were mutated and tested for their transglycosylation activity.

## MATERIALS AND METHODS

For Peer Review

Laboratory reagents were purchased from Sigma-Aldrich unless otherwise indicated and used without further purification.

**Site-directed mutagenesis**

Primers D (5'-CAATTAATCATCGGCTCG) and F (5'-AATCTTCTCTCATCCGCC) flank the gene upstream of the Ptac promoter and downstream of the PstI restriction site. 20 pmoles of D or F primer was mixed with 20 pmoles of mutagenic primers, 10 ng of plasmid pBBGly (pBTac2 vector containing the *ttβ-gly* gene (Feng et al. 2005; Dion et al. 1999)) in a 50-µl PCR tube. The reaction conditions were: 1x Dynazyme buffer, 0.14 mM of dCTP and dGTP, 0.06 mM of dATP and dTTP, 1.5 mM $MgCl_2$ and 2.5 U of polymerase Dynazyme Ext (Finnzymes). The reaction was thermocycled as follows: 96°C 5 min, then 30 cycles at 96°C 30 s; 55°C 30 s and 72°C 5 min. An overlap PCR was then performed using the obtained PCR products as a template and the primers D and F (Higuchi et al. 1988). These PCR products were digested by the *Eco*RI and *Pst*I restriction enzymes and cloned back into the pBTac2 vector digested by the same enzymes and the resulting plasmids were transformed into competent XL1 blue cells. The 1.3-kb DNA fragments encoding the mutant Ttβ-gly enzymes were sequenced in both forward and reverse directions.

**Protein expression and purification**

Recombinant strains expressing the Ttβ-gly mutant genes were grown in 1 L of LB medium at 37°C overnight, centrifuged and resuspended in 20 mL of lysis buffer (0.1 M phosphate, pH 8, 0.5 M NaCl, 10 mM imidazole). After sonication and centrifugation, supernatants were heated for 20 min at 70°C to precipitate most of the host proteins and centrifuged again. Then the 6xHis-tagged proteins were purified by immobilized ion metal-affinity chromatography (IMAC): 250 µL of Ni-NTA Superflow (Qiagen) are added to the supernatant and stirred for 1 h at 4°C, then loaded on a 10 mL column. The column is washed using 25 mL of washing

For Peer Review

buffer (0.1 M phosphate, pH 8, 0.5 M NaCl, 25 mM imidazole), then 5 mL of elution buffer are added (0.1 M phosphate, pH 8, 0.5 M NaCl, 250 mM imidazole). The purification of the different enzymes was achieved by size exclusion chromatography in TBE pH 8 on a Superdex S75 column (GE healthcare) and dialysis against HEPES buffer (pH 7, 50 mM) containing 150 mM NaCl. The purity of the final product was checked by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and protein chip (Agilent). Enzyme concentrations were determined by UV absorbance at 280 nm using a NanoDrop 1000 (Thermo Scientific).

**Crystallization and data collection**

Crystals of Ttβ-gly mutants were grown using the hanging drop vapor diffusion technique in which a drop consisting of 2 µL of the protein solution and 1 µL of the reservoir solution (0.15 M sodium cacodylate, pH 6.5; Tris-citrate sodium 0.6 M) was equilibrated against the reservoir solution at room temperature. Rod-shaped crystals grew in two months and were flash cooled under a nitrogen stream after soaking them briefly in reservoir solution containing 25% glycerol.

The data were collected on the beam lines BM30 (N282T and N390I) and ID14-1 (F401S) at ESRF (Grenoble, France). Data collection parameters are reported in Table S1.

**Structure determination refinement and analysis**

The three-dimensional structures of the mutants F401S, N282T and N390I of Ttβ-gly were determined by molecular replacement with the MOLREP program (Vagin & Teplyakov 1997). The structure of native Ttβ-gly (PDB code 1UG6) was used as a search model. The structures were refined by restrained refinement using the program REFMAC5 (Murshudov et al. 1997) from CCP4 (Collaborative Computational Project 1994) and COOT (Emsley & Cowtan 2004). Refinement statistics are reported in Table S1.

**Bioinformatic analysis**

For Peer Review

GH1 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, 240 different sequences more than 40% identical to the *Thermus thermophilus* beta-glycosidase (Ttβ-gly) sequence were considered. Hereafter, amino-acid residues found conserved in more than 98% of these sequences are termed "highly conserved".

**Kinetic studies**

All kinetic studies were performed with a TECAN spectrophotometer at 40°C in PBS 1x (NaCl 137 mM; KCl 2.7 mM; $NaH_2PO_4$ 10 mM; $KH_2PO_4$ 1.76 mM; pH 7.4). Initial reaction rates were calculated from the slope of the zero-order plot of product concentration (*p*NP) against reaction time, then the initial rates of reaction were plotted against substrate concentration and the resulting curves were fitted to the hyperbolic equation v=Vmax*[S]/(Km+[S]) using Origin 7.0 (OriginLab) to obtain the $k_{cat}$ and $K_m$ parameters. When initial velocities did not exhibit saturation behavior at high substrate concentration, only $k_{cat}/K_m$ were evaluated by measuring the initial slope of the curve v=f([S]) using R 2.15.0 (*www.**r**-project.org)*. All the measurements were made in duplicate with at least 12 different substrate concentrations from 4.9 µM to 10 mM.

**Capillary electrophoresis measurements**

Transglycosylation activities with *p*NP-Fuc as donor and N-methyl-O-benzyl-N-(β-D-glucopyranosyl)-hydroxylamine (Teze et al. 2013a) as acceptor 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). 200 µL of medium containing 10 mM imidazole (used as an internal standard), 15 mM *p*NP-Fuc, 15 mM BnON(Me)-Glc and the enzyme (enzyme amounts are indicated in Fig. 4) were incubated at 40°C in PBS in the capillary electrophoresis apparatus and analyzed every 20 min for 5 h. Separations were

For Peer Review

performed at 29 kV with 20 mM Borax pH 9.5 containing 50 mM SDS as running buffer. Donors, acceptor, products and *p*NP were detected by UV absorbance at 214 nm and quantified by comparison with imidazole (we previously checked that imidazole does not inhibit the enzyme at this concentration). Five experiments with the native enzyme and the F401S mutant gave maximum yields within a percent precision.

### Mass spectrometry

Enzymatic product was permethylated according to Anumula et al. (Anumula & Taylor 1992) and mass measurements were performed on an Autoflex III MALDI TOF/TOF mass spectrometer (Bruker Daltonics, Germany) equipped with a Smartbeam laser (200 Hz, 355 nm) in positive ion mode. For tandem mass spectrometry acquisition, argon was used as a collision gas with heCID fragmentation conditions (Lewandrowski et al. 2005). A DMA/DHB ionic matrix was prepared as described previously (Ropartz et al. 2011) and laser power was optimized for parent isolation and mass measurement of the fragments. 500 laser shots were added in parent mode and 4500 laser shots in fragment mode. Mass spectra were automatically processed with Flex Analysis 3.0 software (Bruker Daltonics, Germany) and manually annotated (see Figure S1).

## RESULTS

### Structural studies of the Ttβ-gly transglycosidase mutants

In order to assess a possible structural basis for the modification of the transglycosylation/hydrolysis ratio in the Ttβ-gly mutants obtained by directed evolution (Feng et al. 2005), the 3D structure of mutants F401S and N282T was determined by X-ray

For Peer Review

crystallography. Diffraction data for mutant N390I could not be collected with a good enough resolution. These mutants showed the highest transglycosylation activity compared to the wild type enzyme. They were crystallized as stable rod-shaped crystals, which belong to different space groups, $P2_1$ (F401S) and C2 (N282T), which are also different to that reported for WT Tt$\beta$-gly ($C222_1$) (1UG6.pdb). The structures of the F401S and N282T mutants were obtained at a resolution of 2.2 Å and 2.0 Å, respectively (Table S1), which was sufficient to make reliable comparisons with the wild type structure. The estimated coordinate error based on the R-factor (18) for these structures is 0.14 Å and 0.19 Å for N282T and F401S, respectively.

For each mutant, three molecules were observed per asymmetric unit, with a chain C more disorganized than the others (Table S1). Using the data for chain A, mutant structures were superimposed on that of the wild type enzyme to detect possible structural differences induced by the mutations but no overall change was observed (all-atom.rms < 0.13 Å). As shown in Figure 1, even within the active site, no significant reorientations of the amino acid side chains were detected, except for those corresponding to the mutations. A closer look at the structures of the active site revealed minor changes, such as, in the N282T mutant, a slight shift of 0.7 Å of the oxygen of the acid/base residue (E164) away from the 282 residue, but still at a classic distance (3 Å) to make a hydrogen bond. In the F401S structure, despite the marked local modification induced by the mutation, only slight shifts could be detected on M310 (0.8 Å) and W385 (0.4 Å) residues. This last movement might be functionally important, since W385 is known to interact strongly with the substrate (Nerinckx et al. 2003).

In order to detect possible structural differences in the glycosyl-enzyme intermediates, for which water and sugar compete in the hydrolysis and transglycosylation reactions (Scheme 1), protein crystals were soaked with DNP-$\beta$-D-2deoxy-2-fluoro-glucopyranoside (DNP-G2F). This substrate is known as a mechanism-based inhibitor which enables the glycosyl-

For Peer Review

intermediate to be trapped (Withers et al. 1987; Withers et al. 1990; Namchuk & Withers 1995; White et al. 1996). Unfortunately, no electronic density for G2F was detected in the active site, regardless of the mutant tested. Similar results were obtained when mutants were co-crystallized with DNP-G2F, suggesting that spontaneous hydrolysis of this substrate was faster than reaction with the nucleophilic residue (E338) to give the glycosyl-enzyme. Indeed, analysis of the mutant enzyme inactivation in PBS leads only to partial inactivation. Instead, a glycerol molecule, a compound used as a cryo-protectant, was found in the -1 subsite, as in the WT structure. This molecule superimposed well with the C2, C3, C4, F2, OH3 and OH4 of G2F in the structure of the glycosyl-enzyme of *Bacillus polymyxa* (1E4I.pdb).

Overall, these results show that the improvement in transglycosidase activity of the mutants cannot be attributed to significant active site structural modifications. However this does not exclude that significant structural or water dynamics (Teze et al. 2013b) changes could occur during the catalytic cycle which are not captured by these particular structures. Thus, our structural comparison of mutants with the wild type enzyme did not provide any general rule that could rationalize the mutagenesis of glycosidases into transglycosidases. We thus looked at other possible common structural characteristics that could link these three point mutations, apart from the fact that they are all located in close proximity to the -1 site.

**Amino acid conservation within the GH1 family**

Since the mutated residues in the transglycosidases previously obtained by directed evolution (N282, N390, F401) (Feng et al. 2005) are conserved among many GH1, we investigated the conservation of all Ttβ-gly amino acids. Thus, using the Ttβ-gly sequence as query, we obtained 240 sequences which share at least 29.47% of sequence identity between all pairs of sequences. The average conservation was 48.2% with a singular distribution, showing that 8.6% of amino acids (37 AAs) are at least 98% conserved (Figure 2A).

For Peer Review

Knowing the Ttβ-gly structure, their location could be assigned and some assumptions about their putative role in the structure and function within the GH1 family could be made (Figure 2B). Firstly, 9 residues seem to be involved in buried salt bridges while 3 aromatic residues are involved in stacking interactions far from the active site. These residues are probably responsible for the structural stability of the $(\beta/\alpha)_8$ fold. In addition, 9 glycines and one proline are highly conserved and localized in loops or within the β barrel, suggesting that they might be important for protein folding. As expected, 15 conserved amino acids are located close to the active site, particularly around the -1 site. 10 of them are directly involved in substrate binding or in catalytic activity (Q18, H119, N163, E164, Y284, E338, W385, E392, W393, F401). Five highly conserved residues (R75, S77, W120, N282 and N390) are also located in the second shell around the -1 site. However, no conserved residue was observed around the +1 and +2 site. This finding is probably related to the lack of common aglycone specificity within the GH1 family. More interestingly, among these well-conserved residues close to the active site, three of them (N282, N390 and F401) have been identified by directed evolution of Ttβ-gly as modulating the T/H balance (Feng et al. 2005). This observation prompted us to postulate that mutations of other residues presenting the same conservation characteristics could also modulate transglycosidase activity.

**Kinetic characterization of new transglycosidase mutants**

When choosing the highly conserved residues to be mutated, catalytic residues (E164 and E338) and those strongly involved in substrate binding by hydrogen bonding or stacking (Q18, H119, W385, E392 and W393) were eliminated to avoid a dramatic decrease in catalytic activity. Among the residues in the first shell of the -1 site, residue N163 was mutated to A and Y284 was mutated to F, since it was previously shown that a corresponding mutation (Y298F) in the β-glycosidase of *Agrobacterium sp.* (Abg) provoked a strong



http://www.peds.oupjournals.org

For Peer Review

reduction in the kinetics of the deglycosylation step (Gebler et al. 1995). Within the conserved second shell residues, there are two of the previously identified positions (N282 and N390), R75, S77 and W120. As S77 is far from the active site (more than 8 Å from the closest glycosyl hydroxyl 3), R75 and W120 were mutated (Figure 3). For position W120, three different mutations were tested: W120F, W120L and W120C, which conserved, respectively, the aromatic property of the residue (F), its hydrophobic nature (L) or switched to a hydrophilic residue (C). R75 was mutated to A.

The resulting mutants were analyzed for their hydrolytic properties by determining the $k_{cat}/K_m$ parameters with *p*NP-β-D-fucose (*p*NP-Fuc) as substrate (Table 1), since this is the preferred substrate of Ttβ-gly (Dion et al. 1999). Transglycosylation activities were determined with N-methyl-O-benzyl-N-(β-D-glucopyranosyl)-hydroxylamine (BnON(Me)-Glc) as acceptor (Scheme 1), since it provides a high transglycosylation yield with the native enzyme, and does not produce regio-isomers of the product N-methyl-O-benzyl-N-(β-D-fucopyranosyl(1→4)β-D-glucopyranosyl)-hydroxylamine (Teze et al. 2013a). Initial velocities and the maximum yield of the transglycosylation reaction for all mutants were assessed by capillary electrophoresis in conditions where the donor/acceptor molar ratio was set at 1 (Table 1). As revealed by TLC, W120C displayed the highest transglycosylation yield among the W120 mutants together with the lowest $k_{cat}/K_m$ (data not shown) so only this substitution was analyzed.

Kinetic parameters of the mutants (R75A; W120C; N163A; N282T; Y284F; N390I; F401S) were first determined with *p*NP-Fuc as the only substrate (Table 1). Saturation was not reached for the mutants, only for the WT enzyme. This behavior has already been observed for the F401S mutant with other substrates (Feng et al. 2005; Koné et al. 2009), and interpreted as the result of the self-condensation reaction (Scheme 1), which significantly contributes to the overall activity at high substrate concentration, especially if the hydrolytic





For Peer Review

activity is low. Thus, for the mutants, only the $k_{cat}/K_m$ parameters could be determined. In comparison with the WT enzyme, $k_{cat}/K_m$ values were lowered by a factor of 30 to 3500 depending on the mutant. W120C and Y284F correspond to mutations that most strongly decrease the overall activity of Ttβ-gly.

In contrast, the transglycosylation activity of the mutants exhibited a smaller decrease (10- to 200-fold reduction) compared to the wild type (Table 1). The maximum yield of transglycosylation products for the WT enzyme and the mutants was then determined (Figure 4). All mutants had an overall transglycosylation yield much higher than WT Ttβ-gly, reaching an exceptionally high value in some cases (90% for the Y284F mutant). Transglycosylation yields were always higher than self-condensation yields, which suggests that the BnON(Me)-Glc acceptor is a better substrate than *p*NP-Fuc for the acceptor site. However, the relative yield of the transglycosylation product and self-condensation depends on the mutation. Furthermore, except for the R75A mutant, this acceptor gave a transglycosylation product which was barely hydrolyzed by the enzymes, leading to its accumulation (Figure 4). For the R75A mutant, significant hydrolysis of the transglycosylation products was observed preventing high accumulation of the products and thus reducing the maximum yield (68%).

## DISCUSSION

Taken together, these results show that highly conserved amino acids around the active site of Ttβ-gly are relevant targets for obtaining very efficient mutants for transglycosylation reactions. Such an approach enabled us to identify new mutations, namely R75A, W120C, N163A and Y284F, which increase the transglycosylation/hydrolysis ratio, in addition to those that were previously identified by directed evolution (F401S, N282T and N390I) (Feng

For Peer Review

et al. 2005). With all these mutants, very high yields of transglycosylation products can be obtained (between 68% and 90%) with *p*NP-Fuc as donor and BnON(Me)-Glc as acceptor.

All these mutations led to lower activity than the WT enzyme. Since $k_{cat}/K_m$ is a measure of $k_2/K_s$ (Scheme 1) (Namchuk & Withers 1995), no general conclusions can be deduced from the reduction of $k_{cat}/K_m$ in Tt$\beta$-gly mutants. However, for all the mutants, the kinetics at high substrate concentration did not reach a plateau and there was an increased rate of *p*NP release in the presence of a high concentration of *p*NP-Fuc. The addition of competing acceptors, such as maltose or cellobiose, has been shown to produce similar effects on the kinetics of F401S and N390I mutants (Feng et al. 2005; Koné et al. 2009). Activation of *p*NP release was also consistently observed for all the new transglycosidase mutants in the presence of an increasing concentration of maltose as acceptor, even at the low substrate concentration of 1 mM (data not shown). So, whatever the effect of the mutations on the glycosylation step, these acceptor-dependent activation processes can be interpreted as an increase in the deglycosylation rate upon addition of an acceptor.

An intriguing feature of our results is that, wherever the position of the mutation around the -1 site, very similar effects were observed on the transglycosylation activity. Correspondingly, we previously showed by saturation mutagenesis that residues F401 and N390 could be substituted by several amino acids (at least by P, K, I, G, Q, N for F401 and I, S, T, D for N390) to give transglycosylation activities similar to that of F401S or N390I, respectively (Feng et al. 2005). As these latter mutations, obtained by directed evolution, are barely or not in direct contact with the reaction center, we can assume that most of the functional modifications of catalysis occur via indirect effects, which could induce subtle changes in the structure of the active site. Structural data obtained with the two different transglycosidase mutants confirm that these effects are indeed tiny. For instance, it was suggested that F401 belongs, with W385 and Y284, to a mechanistically relevant hydrophobic platform well



http://www.peds.oupjournals.org

For Peer Review

conserved in GH1 glycosidases (Nerinckx et al. 2003). The F401S structure indicates that W385 is slightly displaced by 0.4 Å, a movement that could modify the efficiency of this platform in the mechanism (Nerinckx et al. 2003). The role of F401 itself has been described as mainly a hydrophobic interaction with the methyl group of *p*NP-fucose. Additionally, it has been shown that the $k_{cat}/K_m$ of Abg towards *p*NP-Xyl (which is identical to *p*NP-Fuc except that it lacks its methyl) is 140 times lower than that of *p*NP-Fuc, and that the use of this substrate leads to an increase in the T/H ratio (Kempton & Withers 1992). Therefore, the loss of this interaction, either by mutating F401 or by modifying the substrate, leads to a drastic drop in the kinetic parameters and an increase in the T/H ratio. This strongly suggests that this interaction is important for the stabilization of transition states (TS*), but more for the hydrolysis one than for the transglycosylation one. Besides, the N282 residue is involved in a strong hydrogen bond with the acid-base catalytic residue (E164). In the N282T structure, a slight but significant shift of 0.7 Å was measured for the E164 carboxylate oxygens, which could also affect TS* stabilization. Similarly, the phenolic oxygen of Y284 was proposed to stabilize the TS* by hydrogen bonding to the intracyclic oxygen of the glycosyl and then the excess positive charge in the TS* (Nerinckx et al. 2005). This residue was also identified as a general acid catalyst during deglycosylation, stabilizing the leaving nucleophilic catalyst by partially donating a proton (Gebler et al. 1995). Finally, R75, W120 and N163 belong to a network of residues in close contact with the catalytic residues (Figure 3). Consequently, mutations of these residues are expected to disturb this network and affect TS* stabilization. So, we postulate that single mutations of well conserved residues around the -1 active site, wherever they are, induce lower stabilization of the transition states of the reaction, which affects hydrolysis more than transglycosylation. This results in mutant enzymes that have an improved T/H balance.

For Peer Review

This is consistent with previous results obtained by substrate modification: Withers and coauthors have shown with Abg and DNP-G2F as substrate that this compound allows the glycosyl-enzyme to be trapped, not because it is intrinsically more stable but because the TS* of its hydrolysis is strongly destabilized. However, this glycosyl-enzyme is efficiently deglycosylated by transglycosylation (Withers et al. 1987; Withers et al. 1990). This effect on the T/H ratio has been interpreted as being due to stabilizing interactions within the acceptor subsite: the interactions that stabilize the TS* are in both -1 and +1 subsites for glycosylation and transglycosylation but are located only in the -1 subsite for hydrolysis, therefore a decrease in stabilization within the -1 subsite may affect hydrolysis more than transglycosylation (Street et al. 1992). A similar phenomenon is probably responsible for the effect of our mutations.

In conclusion, 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.

For Peer Review

## ACKNOWLEDGMENT

We are indebted to the staff of the European Synchrotron Radiation Facilities (ESRF, Grenoble, France), beamline ID14 and BM30, for technical support during data collection and processing. This work was supported the Glyconet Network. The authors are grateful to Mathieu Fanuel from the BiBS facility for Mass Spectrometry measurements.

## SUPPORTING INFORMATION AVAILABLE

Annotated mass spectrum of N-methyl-O-benzyl-N-(β-D-fucopyranosyl(1→4)β-D-glucopyranosyl)-hydroxylamine, X-ray data collection and processing statistics and structures refinement and final model statistics.

## ABBREVIATIONS

Tt$\beta$-gly: *Thermus thermophilus* $\beta$-glycosidase; T/H ratio: transglycosylation over hydrolysis ratio; GH1: Glycosyl Hydrolase family 1; WT: wild type; BnON(Me)-Glc: N-methyl-O-benzyl-N-($\beta$-D-glucopyranosyl)-hydroxylamine; TLC: Thin Layer Chromatography; *p*NP-Fuc: 4-Nitrophenyl $\beta$-D-fucopyranoside; Abg: *Agrobacterium sp.* $\beta$-glycosidase; DNP-G2F: 2,4-Dinitrophenyl-$\beta$-D-2deoxy-2-fluoro-glucopyranoside.

## REFERENCES

Anumula, K. R., & Taylor, P. B. (1992). *Anal. Biochem.*, (203), 101–108.

Apweiler, R. *et al.* (2004). *Nucleic Acids Res.*, *32*(Database issue), D115–9.

Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., & Henrissat, B. (2009). *Nucleic Acids Res.*, *37*(Database issue), D233–8.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G., & Thompson, J. D. (2003). *Nucleic Acids Res.*, *31*(13), 3497–3500.

Cobucci-ponzano, B., Perugino, G., Strazzulli, A., & Moracci, M. (2012). *Thermophilic Glycosynthases for Oligosaccharides Synthesis*. *Cellulases* (1st ed., Vol. 510, pp. 273–300).

Collaborative Computational Project. (1994) *Acta Crystallogr. Sect. D*, *50*(Pt 5), 760–763.

Dion, M., Fourage, L., Hallet, J.-N., & Colas, B. (1999).*Glycoconj. J.*, *37*(16), 27–37.

Drone, J., Feng, H., Tellier, C., Hoffmann, L., Tran, V., Rabiller, C., & Dion, M. (2005). *Eur. J. Org. Chem.*, *2005*(10), 1977–1983.

Emsley, P., & Cowtan, K. (2004). *Acta Crystallogr. Sect. D*, *60*(Pt 12 Pt 1), 2126–2132.

For Peer Review

Feng, H.-Y., Drone, J., Hoffmann, L., Tran, V., Tellier, C., Rabiller, C., & Dion, M. (2005). *J. Biol. Chem.*, *280*(44), 37088–97.

Gantt, R.W., Peltier-pain, P., Singh, S., Zhou, M. & Thorson, J. S (2013). *Proc Natl Acad Sci USA*, 110(19), 7648–7653.

Gebler, J. C., Trimbur, D. E., Warren, a J., Aebersold, R., Namchuk, M., & Withers, S. G. (1995). *Biochemistry*, *34*(44), 14547–53.

Ghazarian, H., Idoni, B., & Oppenheimer, S. B. (2011). *Acta Histochem*, *113*(3), 236–47.

Hart, G. W., & Copeland, R. J. (2010). *Cell*, *143*(5), 672–676.

Henrissat, B. (1991). *biochem. J.*, *280*, 309–316.

Higuchi, R., Krummel, B., & Saiki, R. K. (1988). *Nucleic Acids Res.*, *16*(15), 7351–7367.

Kempton, J. B., & Withers, S. G. (1992). *Biochemistry*, *31*(41), 9961–9.

Koné, F. M. T., Le Béchec, M., Sine, J.-P., Dion, M., & Tellier, C. (2009). *Protein Eng. Des. Sel.*, *22*(1), 37–44.

Lairson, L. L., Henrissat, B., Davies, G. J., & Withers, S. G. (2008). *Annu. Rev. Biochem.*, *77*, 521–555.

Lewandrowski, U., Resemann, A., & Sickmann, A. (2005). *Anal. Chem.*, *77*(10), 3274–83.

Lichtenstein, R. G., & Rabinovich, G. a. (2013). *Cell Death Differ*, *2869*, 1–11.

Mackenzie, L. F., Wang, Q., Warren, R. A. J., & Withers, S. G. (1998). *J. Am. Chem. Soc.*, *120*(22), 5583–5584.

Malet, C., & Planas, A. (1998). *FEBS Lett.*, (440), 208–212.

For Peer Review

Marton, Z., Tran, V., Tellier, C., Dion, M., Drone, J., & Rabiller, C. (2008). *Carbohydr. Res.*, *343*(17), 2939–46.

Murshudov, G. N., Vagin, A. A., & Dodson, E. J. (1997). *Acta Crystallogr. Sect. D*, *53*, 240–255.

Namchuk, M. N., & Withers, S. G. (1995). *Biochemistry*, *34*(49), 16194–202.

Nerinckx, W., Desmet, T., & Claeyssens, M. (2003). *FEBS Lett.*, *538*(1-3), 1–7.

Nerinckx, W., Desmet, T., Piens, K., & Claeyssens, M. (2005). *FEBS Lett.*, *579*(2), 302–12.

Osanjo, G., Dion, M., Drone, J., Solleux, C., Tran, V., Rabiller, C., & Tellier, C. (2007). *Biochemistry*, *46*, 1022–1033.

Ropartz, D., Bodet, P.-E., Przybylski, C., Gonnet, F., Daniel, R., Fer, M., Helbert, W.,Bertrand, D. & Rogniaux, H. (2011). *Rapid Commun. Mass Spectrom.*, *25*(14), 2059–70.

Street, I. P., Kempton, julie B., & Withers, S. G. (1992). *Biochemistry*, *31*(41), 9970–9978.

Teze, D., Dion, M., Daligault, F., Tran, V., André-miral, C., & Tellier, C. (2013). *Bioorg. Med. Chem. Lett.*, *23*(2), 448–451.

Teze, D., Hendrickx, J., Dion, M., Tellier, C., Woods, V. L., Tran, V. & Sanejouand, Y.-H. (2013). *Biochemistry*, 52, 5900–5910.

Vagin, a., & Teplyakov, A. (1997). *J. Appl. Crystallogr.*, *30*(6), 1022–1025.

Wang, C., Lee, J., Luo, S., Kulkarni, S. S., Huang, Y., & Lee, C. (2007). *Nature*, *446*, 896–899.

Wang, Q., Graham, R. W., Trimbur, D., Warren, R. A. J., & Withers, S. G. (1994). *J. Am. Chem. Soc.*, *116*, 11594–11595.

For Peer Review

White, A., Tull, D., Johns, K., Withers, S. G., & Rose, D. R. (1996). *Nat. Struct. Mol. Biol.*, *3*(2), 149–154.

Withers, S. G., Street, I. P., Bird, P., & Dolphin, D. H. (1987). *J. Am. Chem. Soc.*, *109*(19), 7530–7531.

Withers, S. G., Warren, R. A. J., Street, I. P., Rupitz, K., Kempton, J. B., & Aebersoldg, R. (1990).*J. Am. Chem. Soc.*, *112*(15), 5887–5889.

For Peer Review

## Figure legends

**Scheme 1.** Hydrolysis and transglycosylation reactions catalyzed by the retaining glycosidase Ttβ-gly.

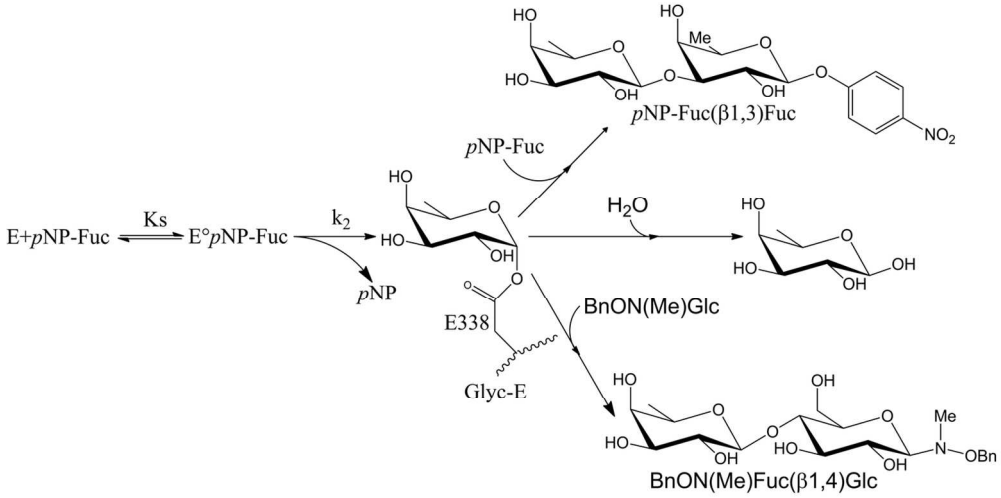
**Figure 1.** Superimposition of the active site structures of the WT (carbons in green, 1UG6), the N282T (carbons in yellow) and the F401S (carbons in blue) mutants of Ttβ-gly. Nitrogens are in blue, oxygens in red and sulfurs in yellow. Catalytic residue identifiers are underlined and the mutated residues are in bold. A β-D-fucose molecule (carbons in gray) is docked in the active site by superimposition with the 2-deoxy-2-fluoro-D-glucosyl using the glycosyl-enzyme structure of *B. polymyxa* β-glycosidase (PDB code: 1E4I). This figure was created with PyMol.

**Figure 2.** Study of the amino acid conservation within the GH1 family. A: Histogram of the population of amino acids as a function of their percentage of conservation. Each population corresponds to an interval of 2% of conservation and the statistics were based on 240 sequences. B: Putative function of the 37 most conserved amino acids, numbered according to the Ttβ−gly sequence. Previously known transglycosidase mutations are in bold and catalytic residues are underlined.[a]D277 is involved in a salt bridge with R207, which is not as highly conserved (85.4% only, K in 7.1% of cases).[b]In more remote GH1 sequences, mutations are observed for residues D355, R358, N390, E392, and F401.

**Figure 3.** Localization of the targeted residues in native Ttβ-gly (PDB code: 1UG6). Carbons are shown in black, oxygens in grey and nitrogens in white. A β-D-fucose molecule is docked in the active site by superimposition with the 2-deoxy-2-fluoro-D-glucosyl using the glycosyl-enzyme structure of *B. polymyxa* β-glycosidase (PDB code: 1E4I). Catalytic residue identifiers are underlined. This figure was created with PyMol.

**Figure 4.** Time course of formation of transglycosylation products with *p*NP-Fuc as donor (15 mM) and BnON(Me)-Glc (15 mM) as acceptor. The percentages are calculated with reference to the initial substrate concentration using the following Ttβ-gly variants: WT 2 $\mu g.mL^{-1}$ (△); R75A 125 $\mu g.mL^{-1}$ (●); N163A 40 $\mu g.mL^{-1}$ (+); Y284F 900 $\mu g.mL^{-1}$ (○).

For Peer Review




65x32mm (600 x 600 DPI)

Peer Review

**Supplementary data**

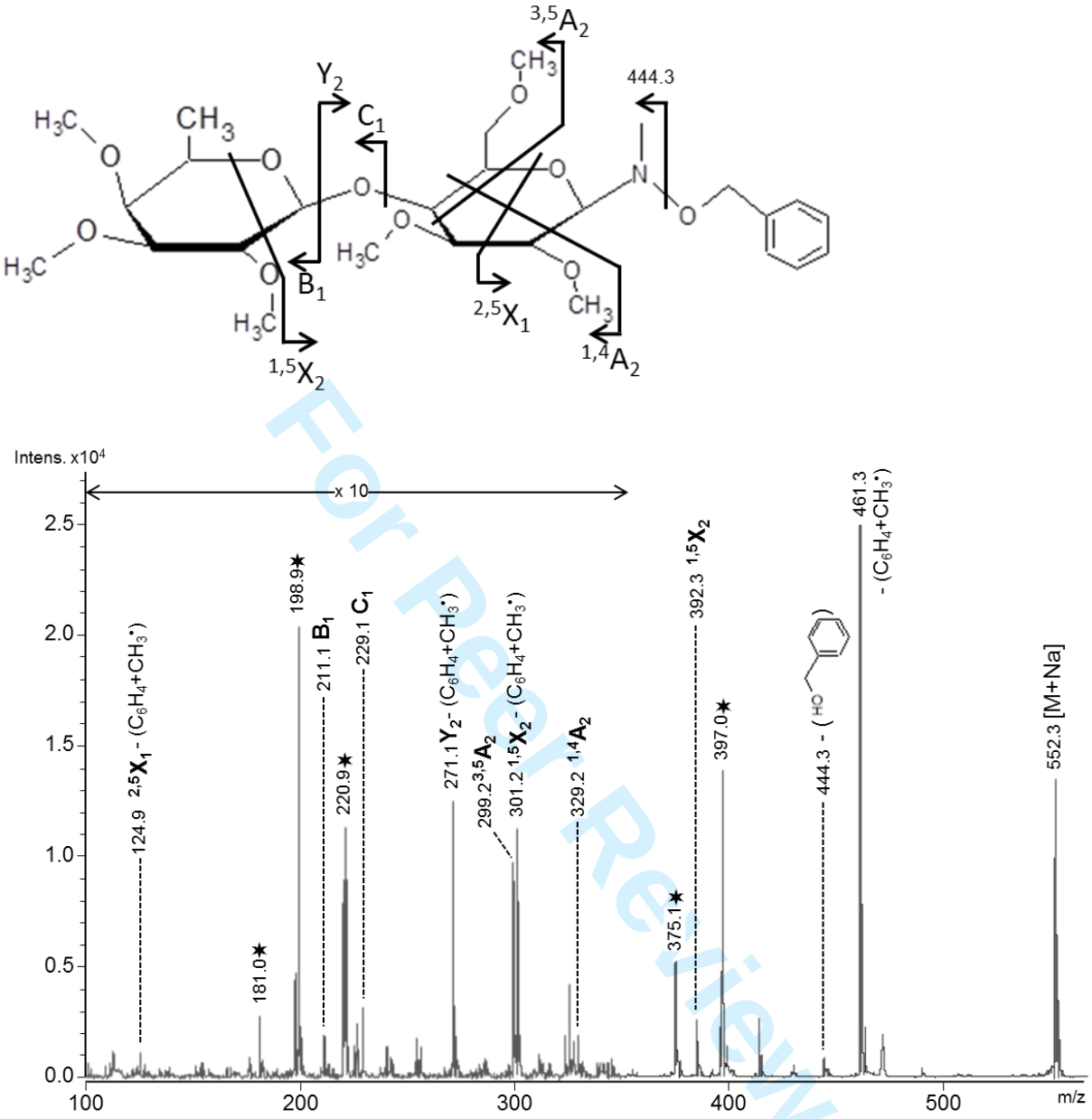



Figure S1. MALDI- TOF/TOF-heCID-MS/MS spectra of permethylated enzymatic product in sodiated form. Fragment ion nomenclature follows Domon and Costello *[Domon, B.; Costello, C. E. Glycoconjugate Journal 1988, 5, 397-409.]*. Annotated peaks $^{3,5}A_2$ and $^{1,4}A_2$ prove that the (1,4) regioisomer is present. Annotated peak with black star correspond to the dissociation of a co-isolated matrix (2,5-dihydroxybenzoïc acid (DHB)) trimer ([3DHB+3(Na-H)+Na]$^+$= m/z 551.015). In details m/z 397.0 = [2DHB+3(Na-H)+Na]$^+$, m/z 375.1 = [2DHB+2(Na-H)+Na]$^+$, m/z 220.9 = DHB+[2(Na-H)+Na]$^+$, m/z 198.9 = [DHB+(Na-H)+Na]$^+$ , m/z 181.0 = [DHB-$H_2O$+(Na-H)+Na]$^+$

For Peer Review

**Table S1**. X-ray data and structure quality statistics for Tt$\beta$-gly mutants

| | F401S | N282T |
|---|---|---|
| PDB code | 3ZJK | 4BCE |
| *Data collection and processing statistics* | | |
| Beamline (ESRF) | ID 14-1 | BM 30 |
| Space group | $P2_1$ | C2 |
| Unit-cell parameters | | |
| A | 98.557 | 143.167 |
| B | 77.322 | 77.665 |
| C | 101.767 | 122.015 |
| Resolution range high (Å) | 2.20 | 2.00 |
| Resolution range low (Å) | 60.18 | 120.13 |
| Completeness (%) | 99.84 | 99.31 |
| Number of reflections | 71852 | 84084 |
| Rmerge | 0.117 (0.549) | 0.074 (0.455) |
| Mean((I)/sd(I)) | 12.0 (3.1) | 11.7 (2.6) |
| Completeness | 99.8 (99.3) | 97.5 (99) |
| Multiplicity | 4.1 (4.0) | 3.1 (3.0) |
| *Refinement and final model statistics* | | |
| R Value (working + test set) | 0.17388 | 0.16784 |
| R Value (working set) | 0.17149 | 0.16761 |
| FREE R VALUE | 0.21869 | 0.17223 |
| Number of non-H atoms used in refinement | 10601 | 10827 |
| Rms Bond deviation | 0.0201 | 0.0264 |
| Rms Angle deviation | 1.7355 | 1.9404 |
| Overall B-factor | 33.453 | 33.520 |
| MolA | 25.254 | 28.872 |
| MolB | 28.290 | 30.834 |
| MolC | 48.165 | 41.832 |
| Water+ligands | 22.272 | 27.304 |

Values given in parentheses refer to reflections in the outer resolution shell (N282T: 1.95-2.06Å; F401S: 2.2-2.32Å)

For Peer Review

**Table 1**. Kinetic parameters of the hydrolytic reaction with *p*NP-Fuc and of the transglycosylation reaction in the presence of *p*NP-Fuc (15 mM) as donor and BnON(Me)-Glc (15 mM) as acceptor for WT and mutant Ttβ-gly at 40°C and pH 7.4.

| Enzyme | $k_{cat}/K_m$ [a] ($mM^{-1}.s^{-1}$) | Transglycosylation activity[b] ($\mu mol.min^{-1}.mg^{-1}$) | Transglycosylation yield [c] (%) | Self-condensation[d] (%) | Overall yield (%) |
|---|---|---|---|---|---|
| Y284F | $(3.83 \pm 0.66)x10^{-2}$ | 0.35 | 76 | 14 | 90 |
| F401S | $0.27 \pm 0.01$ | 1.1 | 82 | 5 | 87 |
| N390I | $0.48 \pm 0.05$ | 0.95 | 73 | 13 | 86 |
| N282T | $0.97 \pm 0.06$ | 1.4 | 75 | 11 | 86 |
| W120C | $(7.8 \pm 2.1)x10^{-3}$ | 1.2 | 57 | 28 | 85 |
| N163A | $0.176 \pm 0.005$ | 3.3 | 77 | 4 | 81 |
| R75A | $0.160 \pm 0.004$ | 1.4 | 65 | 3 | 68 |
| WT | $27.9 \pm 2.2$ | 37 | 36 | 0 | 36 |

[a]Kinetic constants were determined from the initial substrate-dependent variation of the *p*NP release.

[b]Transglycosylation activity corresponds to the initial condensation velocity and was determined by separation and quantification of products by capillary electrophoresis.

[c]Maximum yield corresponds to the transient maximum amount of transglycosylation product after long term incubation.

[d]Self-condensation yield is measured at the same time as the condensation product.

For Peer Review