Structural dissection of the active site of *Thermotoga marítima* β-galactosidase identifies key residues for transglycosylating activity

David Talens-Perales, Julio Polaina and Julia Marín-Navarro*

Instituto de Agroquímica y Tecnología de Alimentos, CSIC, Paterna, Valencia, Spain.

*Corresponding author.

E-mail address: jumana@iata.csic.es (J. Marín-Navarro)

ABSTRACT

Glycoside hydrolases, specifically β-galactosidases, can be used to synthesize galacto-oligosaccharides (GOS) due to the transglycosylating (secondary) activity of these enzymes. Site-directed mutagenesis of a thermoresistant β-galactosidase from *Thermotoga maritima* has been carried out to study the structural basis of transgalactosylation and to obtain enzymatic variants with better performance for GOS biosynthesis. Rational design of mutations was based on homologous sequence analysis and structural modelling. Analysis of mutant enzymes indicated that residue W959, or an alternative aromatic residue at this position, is critical for the synthesis of β-3′-galactosyl-lactose, the major GOS obtained with the wild-type enzyme. Mutants W959A and W959C, but not W959F, showed an 80% reduced synthesis of this GOS. Other substitutions: N574S, N574A and F571L increased the synthesis of β-3′-galactosyl-lactose about 40%. Double mutants F571L/N574S and F571L/N574A showed an increase of about twofold.

KEYWORDS: enzyme engineering, galacto-oligosaccharides, GH2 glycoside hydrolase, prebiotics

INTRODUCTION

Consumer demand of prebiotic compounds has risen sharply in the past years. Consequently, the development of methods for the synthesis of these products has become an industrially relevant target. Galacto-oligosaccharides (GOS) constitute one of the major types of prebiotic molecules. They consist of short chains (2 to 10) of galactose residues linked in some cases to a terminal glucose at the reducing end. GOS are resistant to gastric acid and digestive enzymes. When they reach the colon, they are preferentially metabolised by beneficial intestinal bacteria (prebiotic effect)[1]. GOS have been associated to multiple health-promoting effects, such as hindering enteropathogens adhesion to intestinal epithelium[2], improving calcium absorption and preventing cardiovascular pathologies[3]. Chemically, GOS differ in their polymerization degree and in the type of linkage that connects the galactosyl units (β-(1-2), β-(1-3), β-(1-4) or β-(1-6))[4]. These differences have been related to their stability and prebiotic potential[5,6].

GOS are produced by transgalactosylation of lactose by β-galactosidases. According to the carbohydrate active enzyme database (Cazy), β-galactosidases are found in four related families (GH1, GH2, GH35 and GH42) grouped in the GH-A clan of glycoside hydrolases[7]. A common structural feature of the proteins of this clan is the presence of a TIM barrel, defined by eight parallel beta strands surrounded by eight α-helices, harbouring the catalytic residues[7,8]. These enzymes proceed through a two-step retaining mechanism that enables them to catalyse both hydrolysis and transglycosylation[9,10] (Supplementary Figure S1). In the first

step, the catalytic residue acting as nucleophile forms a covalent intermediate with the galactosyl moiety from lactose, and the acid/base catalyst assists the departure of the glucosyl unit. In the second step, the galactosyl group is transferred to an acceptor deprotonated by the acid/base catalyst. This acceptor can be either a water molecule or another lactose molecule, resulting in final hydrolysis or transgalactosylation, respectively[5]. Different retaining glycosyl hydrolases have been effectively used to synthesize functional oligosaccharides. For instance, modified versions of yeast invertase that produce high yields of 6-kestotriose have been obtained[11,12].

High temperature is advantageous for large-scale GOS synthesis since it reduces the risk of microbial contamination and increases lactose solubility, which in turn enhances GOS yield[13]. To work under these conditions, thermostable enzymes are required. Some examples of thermoresistant β-galactosidases with transgalactosylating capacity are found in families GH1 (from *Sulfolobus solfactaricus*, *Pirococcus furiosus* and *Thermus sp.*) and GH2 (from *Streptococcus thermophilus* and *Thermotoga maritima*), showing different product specificity and yields[14-17].

In this work we have analysed structural and functional properties of a heat-stable GH2 β-galactosidase from the thermophilic bacterium *Thermotoga maritima*. This enzyme (henceforth TmLac) produces β-3'-galactosyl-lactose and β-6'-galactosyl-lactose[18,19]. We have investigated the role of different residues located at the catalytic centre in the transgalactosylation reaction. Our study provides information about the structural basis of transglycosylation efficiency and product specificity of TmLac. This information will be valuable to obtain new

versions of GH2 enzymes capable to synthesize GOS with higher yields or modified chemical structure.

MATERIALS AND METHODS

Strains and culture conditions.

Escherichia coli XL1-Blue (Stratagene) was the host strain for standard DNA manipulations. Overproduction of TmLac mutant enzymes was carried out in *E. coli* XRA. This strain was obtained by transformation of *E. coli* XL1-Blue with the pRARE2 plasmid, which encodes a set of tRNAs which are deficient in *E. coli* and confers resistance to chloramphenicol. pRARE2 was isolated from the *E. coli* strain Rosetta2 (Novagen). Transformants were grown at 37 ºC in LB media (0.5% yeast extract, 1% peptone, 0.5% NaCl) supplemented with 100 mg/L ampicilin and, in the case of XRA transformants, 68 mg/L chloramphenicol, with 2% agar for solid media.

Cloning and mutagenesis.

Site-directed mutagenesis was carried out by PCR, using as template the plasmid TmLac-pQE carrying the gene encoding the wild-type version of TmLac[19], the oligonucleotides showed in Table S1 (supplementary material), and a polymerase with proofreading activity (Phusion DNA polymerase, Thermo Scientific), following the procedure from Hemsley et al[20]. The resulting mutant

genes were sequenced by using ABI Prism BigDye Terminator cycle sequencing kit (Applied Biosystems), and the labeled products were analysed at the Sequencing Service of the University of Valencia (Spain) using a 3730 DNA analyzer (Applied Biosystems). Each plasmid was introduced in *E. coli* XRA strain to optimize gene expression.

Expresion of TmLac.

E. coli transformants were grown in 2XTY (1% yeast extract, 1.5% triptone, 0.5% NaCl) supplemented with 100 mg/L ampicillin and 68 mg/L chloramphenicol at 37ºC up to an optical density at 600 nm of 0.6 before the induction with 1 mM IPTG (isopropyl β-D-galactopyranoside) at 16 ºC during 14 hours. Cultures were centrifuged at 2400 x g for 10 min at 4 ºC, and the cells were concentrated 100-fold by resuspending in buffer A (20 mM phosphate, 10 mM imidazole, 500 mM NaCl; pH 7.4) supplemented with Complete EDTA-free protease inhibitor cocktail (Roche). Cells were broken by sonication (5 cycles of 15s at 185 W and 0.5 Hz, followed by 30 s at rest) with a Labsonic from Braun-Biotech keeping the samples on ice. The extract was clarified by centrifugation (15000 g for 20 minutes at 4ºC) and kept at 4ºC.

Enzyme assays.

Total β-galactosidase activity was assayed by incubating the clarified extract with 25% (w/v) lactose in reaction buffer (50 mM phosphate pH6.5, 10 mM NaCl, 1 mM $MgCl_2$) at 75ºC for 30 min. As previously mentioned, GOS yields

increase with higher lactose concentrations, and therefore a lactose concentration close to its limit of solubility at the assay temperature was chosen. The reaction was stopped by heating at 95 ºC for 10 min and the amount of released glucose was analysed with a glucose assay kit (Sigma). One unit of activity was defined as the amount of extract that produces 1 μmol of glucose per minute.

Analysis of transglycosylating activity was carried out by incubation of the enzyme extracts (0.7-0.8 U/mL) with 25% (w/v) lactose in reaction buffer at 75ºC. The reaction was stopped at different times to determine released glucose or GOS production. The pattern of oligosaccharides was analysed by high performance anion exchange chromatography, coupled to a pulsed amperometric detector (Dionex, Thermo Fisher Scientific) equipped with a CarboPac PA-100 column as previously described[19]. β-3'-galactosyl-lactose and β-6'-galactosyl-lactose standards were kindly provided by Francisco Javier Cañada.

Analysis of enzyme thermal stability.

Clarified cell extracts of *E. coli* were subjected to a heat-shock treatment which was carried out by incubating 200 μL of the extract at 85 ºC for 5 minutes. Soluble proteins after heat-shock were recovered after centrifugation at 19000 g for 20 minutes and analysed by SDS-PAGE, as previously described[11] in parallel to untreated samples.

Bioinformatic tools.

Multiple sequence analysis was carried out with ClustalW[21]. Molecular modelling was performed with the I-Tasser server[22]. The Pymol software (Delano Scientific LLC 2006) was used for structural analysis and visualization of modelled structures. Docking analysis between β-3´-galactosyl-lactose and the modelled TmLac was carried out with the program Autodock4[23] (http://autodock.scripps.edu/references) (see supplementary material for details). The coordinates of β-3´-galactosyl lactose were obtained with GLYCAM (http://www.glycam.org).

RESULTS

Structural modelling and mutant design.

A model of TmLac on which the design of mutations that could affect transglycosylation would be based upon, was constructed. The structure of most GH2 β-galactosidases is composed by four β-sandwich domains surrounding the central TIM barrel that harbours the catalytic residues. The template with highest sequence identity (ca. 40 %) was the β-galactosidase encoded by the *lac*Z gene of *E. coli* (PDB code 3CZJ), hereafter EcLac. Interestingly, TmLac is very similar to EcLac except that it shows an extra C-terminal domain composed of about 100 residues with a predicted β-sandwich fold (Figure 1A). This domain, which is present in β-galactosidases from *Thermotoga* species but not in other homologous enzymes (our own observation), was modeled using as template a CARDB-like domain from *Pyrococcus furiosus*, with a significant (ca. 20%) sequence identity. The C-score

183 value in both cases was higher than -1.5, which is considered the threshold for
184 structural models with correct toplology[22].

185 Inspection of the active site revealed that residues involved in catalysis or in
186 substrate binding in EcLac are highly conserved (70% identity) in TmLac,
187 rendering a good quality model for this region. In order to identify potential
188 residues involved in transgalactosylation, a docking analysis was carried out using
189 β-3´-galactosyl-lactose as ligand, since this is the major GOS synthesized by wild-
190 type TmLac[19]. The complex of TmLac with β-3´-galactosyl-lactose (Figure 1B) may
191 be considered as an analogue of the reaction intermediate where the enzyme is
192 covalently linked to a galactosyl moiety and a new lactose molecule binds to the
193 active site to act as the galactosyl acceptor (Figure S1). According to this model,
194 subsite -1, comprising residues that bind the sugar moiety at the non-reducing
195 end[24], would be conformed by D191, H369, N440, E441, E507, H510, W538 and
196 N574. Most of these residues are highly conserved among GH2 β-galactosidases
197 and are associated to essential roles in catalysis[25-30]. Moreover, mutations at many
198 of these positions, including the putative nucleophile and acid/base catalyst (E507
199 and E441, respectively), rendered inactive versions of EcLac[25-29]. Therefore, these
200 residues were discarded as mutagenesis targets. Only N574 showed some
201 sequence divergence among GH2 enzymes, with Asp at the equivalent position in
202 some cases (Figure 1C). The possible functionality of this substitution, either for
203 transglycosylation efficiency or product specificity was investigated with mutant
204 N574D. Furthermore, mutants N574A and N574S were generated. The rationale
205 behind this was that the disruption of a hydrogen bond with the galactosyl moiety
206 covalently linked to the enzyme in the reaction intermediate may increase the
207 rotational freedom of this group to be coupled to the acceptor lactose. Subsites +1

208 and +2, binding the galactosyl and glucosyl moiety of the acceptor lactose,
209 respectively, would involve fewer residues than subsite -1 and should not be as
210 critical for β-galactosidase activity. W959 and N95 are in contact to both the
211 galactosyl and glucosyl groups of the acceptor lactose, whereas F571 and D568
212 may conform subsite +2. Interestingly, W959 is conserved in bicistronic β-
213 galactosidases composed by two different polypeptides, where this residue is
214 located in subunit LacM, different from LacL, which contains the catalytic residues
215 (Figure 1C). In some monocistronic β-galactosidases, like that from *Kluyveromyces*
216 *lactis*, the equivalent position is occupied by a cysteine (Figure 1C). The
217 functionality of W959, F571 and D568 in the transgalactosylation reaction was
218 tested with mutants W959C, W959F, W959A, F571L and D568A. Finally, V93 and
219 V94 are located close (around 4 and 7 Å, respectively) to the terminal glucosyl
220 moiety (Figure 1B). These residues are substituted by Thr or Gln in other β-
221 galactosidases with high transgalactosylating activity (Figure 1C). These more
222 extended side chains may establish additional hydrogen bonds to the acceptor
223 lactose or facilitate the binding of a trisaccharide as alternative acceptor to
224 generate a tetrasaccharide. Mutants V93T, V94T and V94N were generated to
225 explore this hypothesis.
226
227 **Analysis of heterologous gene expression and thermal stability of mutant**
228 **enzymes.**
229
230 Extracts of *E. coli* transformants expressing different TmLac mutants were
231 analysed by SDS-PAGE to evaluate the production of enzyme. Thermal stability of
232 the mutant versions was assessed by a parallel analysis of the remaining soluble

233 protein after heat treatment. All TmLac mutants had a expression efficiency and
234 thermal resistance similar to the wild-type, with a few exceptions (supplementary
235 Figure S2). Mutants N574S and W959F were produced in significantly lower
236 amounts than the wild-type enzyme. Despite of this, the percentage of enzyme
237 recovered after heat treatment was similar to that obtained with the wild-type.
238 Therefore, none of the introduced mutations seem to cause drastic structural
239 changes altering the overall stability of the enzyme.
240
241 **Activity of transformants expressing mutant TmLac versions.**
242
243 Glucose release by TmLac activity is concomitant to both hydrolysis and
244 transglycosylation (Figure S1). Therefore, initial velocity of glucose production can
245 be used as an estimation of the overall activity of the enzyme. Substitution of Asn
246 574 by Ala or Ser caused a significant decrease in the global activity of the enzyme
247 per unit of total protein in the extract (to 31 % and 8 % of the wild-type,
248 respectively), whereas replacement by Asp did not have a significant effect (Table
249 1). The drastic reduction of activity in the N574S mutant seems to be correlated
250 with the decreased synthesis of enzyme observed in Figure S2. The three
251 mutations tested at W959 were associated with a decrease in the overall specific
252 activity of the enzyme (Table 1), although the decrease of W959F activity may be
253 simply the result of the lower expression of this mutant enzyme. On the contrary,
254 substitution of W959 by non-aromatic residues (Cys or Ala) caused a significant
255 decrease in activity (to 21% and 45% of the wild-type, respectively) despite
256 showing a similar enzyme amount. Activity of the rest of mutants was similar to
257 that of the wild-type. Therefore, comparing relative activity values (Table 1) with

relative enzyme amounts (Figure S2), we conclude that only substitutions N574A, W959C and W959A cause a significant detriment in TmLac activity per unit of enzyme mass.

Analysis of transgalactosylation efficiency and GOS profile of mutant enzymes.

The kinetics of β-3'-galactosyl-lactose and β-6'-galactosyl-lactose synthesis was evaluated for the wild-type enzyme (supplementary Figure S3A). In parallel, glucose release was determined as an estimation of enzyme activity, revealing that TmLac is progressively inactivated (Figure S3B). After 3 hours of reaction, the rate of glucose release was around 20 % of the initial velocity. Since neither substrate availability nor enzyme stability is compromised after this incubation time[19], this may be a consequence of enzyme inhibition by product accumulation, as reported for other β-galactosidases[31,32]. Consequently, GOS synthesis was also slowed down as reaction progressed, with a much slower rate after 3 hours of reaction (Figure S3A). As previously reported[19], β-3´-galactosyl-lactose was preferentially synthesized over β-6´-galactosyl-lactose. In order to compare GOS yields from different enzymatic variants, the same units of enzyme (i.e. initial glucose release per unit of time) were used. GOS production was determined after 5 hours of reaction and the kinetics of glucose release was monitored in parallel as a control. All the enzymatic versions synthesized the same GOS types as the wild-type enzyme (β-3´-galactosyl-lactose and β-6´-galactosyl-lactose). GOS with higher degree of polymerization were not detected in any case.

Substitutions N574S and N574A, but not N574D yielded an increase in the synthesis of β-3′-galactosyl-lactose, compared to the wild-type enzyme (around 30% and 40% higher, respectively), whereas β-6′-galactosyl-lactose synthesis was not significantly affected (Figure 2A). Similarly, the mutant F571L also showed an increased β-3′-galactosyl-lactose synthesis (around 40%) (Figure 2B).

Mutations at Asp 568, Val 93 and Val 94 did not affect the transgalactosylation efficiency or product specificity of the enzyme (Figure 3A and 3C). In contrast, substitutions W959C and W959A caused a drastic reduction (around 80%) in β-3′-galactosyl-lactose synthesis (Figure 3B). This effect was not so remarkable in β-6′-galactosyl-lactose synthesis. Therefore, β-6′-galactosyl-lactose to β-3′-galactosyl-lactose ratio was reduced from 2.6, obtained with the wild-type enzyme, to 1 and 0.7, for W959C and W959A, respectively. The W959F mutation did not cause a significant change in the transgalactosylating properties of the enzyme.

Substitutions with higher increase in transgalactosylating efficiency were combined generating mutants F571L/N574S and F571L/N574A. The resulting enzymes showed a significant reduction in global activity (Table 1), despite their expression was similar to that of the wild-type (Figure S2). The synthesis of β-3′-galactosyl-lactose by F571L/N574S and F571L/N574A was increased around 90% and 70% compared to the wild-type enzyme, respectively (Figure 2C). This represented a significant increment compared to the single mutants F571L and N574S.

DISCUSSION

GOS synthesis has become an active research topic due to the increasing use of these compounds in functional food and nutraceuticals. GOS with β(1,3), β(1,4) or β(1,6) linkages can be obtained with enzymes from different microbial sources[5]. Production of commercial GOS is carried out with enzymes from *Bifidobacterium bifidum* for β(1,3)-GOS, *Cryptococcus laurentii* or *Bacillus circulans* for β(1,4)-GOS and a mixture of *Aspegillus oryzae* and *Streptococcus thermophilus* for β(1,6)-GOS[33-39]. Production of GOS in high yield and with a broad chemical repertoire, are two major biotechnological challenges. A few cases of enzyme engineering to increase GOS yield, using GH1 and GH42 enzymes, have been reported[40-42]. So far GH2 enzymes have not been manipulated with this purpose. Therefore, the aim of this study was to find out the role that specific residues in the active site of a GH2 glycosidase might have in the transglycosylating activity and specificity of the enzyme. For this purpose we chose a thermoresistant β-galactosidase from the bacterium *Thermotoga maritima*. We undertook rational design of mutations based on the analysis of homologous sequences of GH2 enzymes with transglycosylating activity, and docking analysis of a structural model of TmLac with the transglycosylation product β-3′-galactosyl-lactose.

Aminoacid substitutions within the -1 subsite may disrupt a hydrogen bond with the galactosyl moiety in the covalent complex formed in the reaction intermediate. Accordingly, mutations N574S and N574A, caused a significant decrease in overall activity (Table 1), but in contrast transgalactosylation activity was increased with a substantial increment (30-40%) in the synthesis of β -3′-galactosyl-lactose. In the mutants, the loss of a stabilising interaction would decrease the affinity of the enzyme by the substrate and the stability of the

covalent intermediate, explaining the decrease in activity. However, a higher rotational freedom of the galactosyl moiety in the covalent complex could favour its transfer to an acceptor lactose molecule (Figure 4A). Our results also suggest that this higher flexibility would specifically facilitate coupling to the acceptor lactose through a β(1,3) linkage (Figure 2A). Mutant N574D, which involves the substitution of a carboxy group by an amide with the same side chain length, may still be able to keep such polar contact with the galactosyl moiety. In agreement with this, no changes were found in the protein expression (Figure S2), activity (Table 1) or transglycosylating efficiency (Figure 2A) of this mutant.

Docking analysis suggests that residues W959 and F571 may conform an aromatic residue platform for binding the acceptor lactose (Figure 1B). In agreement with this, substitutions of W959 by non-aromatic residues had a drastic effect in transgalactosylation. The homologous residue in EcLac (W999) has been assigned a double role in the catalysis of the enzyme, for lactose binding in the so-called "shallow" mode, and for glucose binding in the galactosyl covalent intermediate for allolactose synthesis[43-45]. Our results suggest that this residue is also involved in binding the lactose molecule that acts as an acceptor in the transgalactosylating reaction. The effect of W959 substitutions was specially remarkable for β-3´-galactosyl-lactose synthesis (Figure 3B). The β(1,3) to β(1,6) synthesis specificity decreased from 2.6 in the wild-type enzyme to 1 and 0.7 in W959C and W959A, respectively. Interestingly, β-galactosidase from *K. lactis*, with a Cys residue in the homologous position synthesizes preferentially β-6´-galactosyl-lactose[46], suggesting that this residue may be critical for the specific synthesis of β-3´-galactosyl-lactose.

Substitution of highly conserved F571 by a non-aromatic residue (F571L), resulted in a significant increase in the synthesis of β-3´-galactosyl-lactose (Figure 2B). Modelling studies suggest that substitution F571L may allow the rotation of W959, which seems to be sterically impeded in the wild-type enzyme (Figure 4B). Higher flexibility may facilitate a better orientation of W959 as a platform for binding the acceptor lactose molecule, favouring the synthesis of a β(1,3) linkage (Figure 2B). According to our assumptions, double mutation at F571 and N574 (F571L/N574S or F571L/N574A) would simultaneously confer higher flexibility of the galactosyl group at the covalent intermediate (by disrupting the putative hydrogen bond with N574) and a better docking of the acceptor lactose by reorienting W959, resulting in a more favourable coupling of both molecules. This double mutation has increased the transgalactosylating efficiency of the wild-type enzyme up to two-fold.

GOS yields obtained with β-galactosidases range between 20-200 g/L, with different chemical profiles[5]. Among thermoresistant β-galactosidases, GH1 enzymes from *S. solfactaricus* and *P. furiosus* synthesize around 40-50 g/L of trisaccharides consisting of β-3´-galactosyl-lactose and β-6´-galactosyl-lactose in a 2:1 ratio[15] whereas GH2 β-galactosidase from *S. thermophilus* produced galactose disaccharides but no trisaccharides[14]. Site-directed mutagenesis of the GH2 TmLac enzyme reported in this communication increased GOS yield in about 50% (from 30 g/L to 45 g/L) with a final β-3´-galactosyl-lactose to β-6´-galactosyl-lactose ratio of 4:1. Due to the high degree of conservation of F571 and N574 among the GH2 β-galactosidases, our results could be extrapolated to other enzymes, including those with different product specificity. Equivalent mutations may be attempted to increase the yield of different GOS products.

ACKNOWLEDGMENT

We thank F. J. Cañada for supplying us with β-3'-galactosyl-lactose and β-6'-galactosyl-lactose, used as standards.

SUPPORTING INFORMATION

Methodological details for docking analysis; Table S1, Oligonucleotides used as primers for site-directed mutagenesis; Figure S1, Reaction scheme of hydrolysis and transgalactosylation; Figure S2, Analysis of expression of TmLac mutants in *E. coli* by SDS-PAGE; Figure S3, Kinetics of glucose release and GOS synthesis by wild-type TmLac. This material is available free of charge via the Internet at http://pubs.acs.org

REFERENCES

1. Hernot, D. C.; Boileau, T. W.; Bauer, L. L.; Middelbos, I. S.; Murphy, M. R.; Swanson, K. S.; Fahey, G. C., Jr. In vitro fermentation profiles, gas production rates, and microbiota modulation as affected by certain fructans, galactooligosaccharides, and polydextrose. *J. Agric. Food Chem.* **2009,** *57*, 1354-1361.

2. Zhong, Y.; Cai, D.; Cai, W.; Geng, S.; Chen, L.; Han, T. Protective effect of galactooligosaccharide-supplemented enteral nutrition on intestinal barrier function in rats with severe acute pancreatitis. *Clin. Nutr.* **2009,** *28*, 575-580.

3. Macfarlane, G. T.; Steed, H.; Macfarlane, S. Bacterial metabolism and health-related effects of galacto-oligosaccharides and other prebiotics. *J. Appl. Microbiol.* **2008,** *104*, 305-344.

4. Roberfroid, M. Prebiotics: the concept revisited. *J. Nutr.* **2007,** *137*, 830S-837S.

5. Torres, D. P. M.; Gon alves, M. D. P.; Teixeira, J. A.; Rodrigues, L. G. R. Galacto-oligosaccharides: production, properties, applications, and significance as prebiotics. *Compr. Rev. Food Sci. Food Saf.* **2010,** *9*, 438-454.

6. Hernández-Hernández, O.; Marín-Manzano, M. C.; Rubio, L. A.; Moreno, F. J.; Sanz, M. L.; Clemente, A. Monomer and linkage type of galacto-oligosaccharides affect their resistance to ileal digestion and prebiotic properties in rats. *J. Nutr.* **2012,** *142*, 1232-1239.

7. Lombard, V.; Golaconda, R. H.; Drula, E.; Coutinho, P. M.; Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **2014,** *42*, D490-D495.

423 8. Reardon, D.; Farber, G. K. The structure and evolution of alpha/beta barrel
424 proteins. *FASEB J.* **1995,** *9*, 497-503.
425 9. Faijes, M.; Planas, A. In vitro synthesis of artificial polysaccharides by
426 glycosidases and glycosynthases. *Carbohydr. Res.* **2007,** *342*, 1581-1594.
427 10. Wang, L. X.; Huang, W. Enzymatic transglycosylation for glycoconjugate
428 synthesis. *Curr. Opin. Chem. Biol.* **2009,** *13*, 592-600.
429 11. Lafraya, A.; Sanz-Aparicio, J.; Polaina, J.; Marin-Navarro, J. Fructo-
430 oligosaccharide synthesis by mutant versions of *Saccharomyces cerevisiae*
431 invertase. *Appl. Environ. Microbiol.* **2011,** *77*, 6148-6157.
432 12. Marin-Navarro, J.; Talens-Perales, D.; Polaina, J. One-pot production of
433 fructooligosaccharides by a *Saccharomyces cerevisiae* strain expressing an
434 engineered invertase. *Appl. Microbiol. Biotechnol.* **2015,** *99*, 2549-2555.
435 13. Gänzle, M. G. Enzymatic synthesis of galacto-oligosaccharides and other
436 lactose derivatives (hetero-oligosaccharides) from lactose. *Int. Dairy J.* **2012,** *22*,
437 116-122.
438 14. Greenberg, N. A.; Mahoney, R. R. Formation of oligosaccharides by β-
439 galactosidase from *Streptococcus thermophilus. Food Chem.* **1983,** *10*, 195-204.
440 15. Petzelbauer, I.; Reiter, A.; Splechtna, B.; Kosma, P.; Nidetzky, B.
441 Transgalactosylation by thermostable beta-glycosidases from *Pyrococcus furiosus*
442 and *Sulfolobus solfataricus*. Binding interactions of nucleophiles with the
443 galactosylated enzyme intermediate make major contributions to the formation of
444 new beta-glycosides during lactose conversion. *Eur. J. Biochem.* **2000,** *267*, 5055-
445 5066.

16. Akiyama, K.; Takase, M.; Horikoshi, K.; Okonogi, S. Production of galactooligosaccharides from lactose using a beta-glucosidase from *Thermus* sp. Z-1. *Biosci. Biotechnol. Biochem.* **2001,** *65*, 438-441.

17. Ji, E. S.; Park, N. H.; Oh, D. K. Galacto-oligosaccharide production by a thermostable recombinant β-galactosidase from *Thermotoga maritima*. *World J. Microb. Biot.* **2005,** *21*, 759-764.

18. Kim, C. S.; Ji, E. S.; Oh, D. K. Characterization of a thermostable recombinant beta-galactosidase from *Thermotoga maritima*. *J. Appl. Microbiol.* **2004,** *97*, 1006-1014.

19. Marin-Navarro, J.; Talens-Perales, D.; Oude-Vrielink, A.; Canada, F. J.; Polaina, J. Immobilization of thermostable beta-galactosidase on epoxy support and its use for lactose hydrolysis and galactooligosaccharides biosynthesis. *World J. Microbiol. Biotechnol.* **2014,** *30*, 989-998.

20. Hemsley, A.; Arnheim, N.; Toney, M. D.; Cortopassi, G.; Galas, D. J. A simple method for site-directed mutagenesis using the polymerase chain reaction. *Nucleic Acids Res.* **1989,** *17*, 6545-6551.

21. McWilliam, H.; Li, W.; Uludag, M.; Squizzato, S.; Park, Y. M.; Buso, N.; Cowley, A. P.; Lopez, R. Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res.* **2013,** *41*, W597-W600.

22. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC. Bioinformatics.* **2008,** *9:40. doi: 10.1186/1471-2105-9-40.*, 40-49.

23. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009,** *30*, 2785-2791.

24. Davies, G. J.; Wilson, K. S.; Henrissat, B. Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J.* **1997,** *321*, 557-559.

25. Xu, J.; McRae, M. A.; Harron, S.; Rob, B.; Huber, R. E. A study of the relationships of interactions between Asp-201, Na^+ or K^+, and galactosyl C6 hydroxyl and their effects on binding and reactivity of beta-galactosidase. *Biochem. Cell Biol.* **2004,** *82*, 275-284.

26. Huber, R. E.; Hlede, I. Y.; Roth, N. J.; McKenzie, K. C.; Ghumman, K. K. His-391 of beta-galactosidase (*Escherichia coli*) promotes catalyses by strong interactions with the transition state. *Biochem.Cell Biol.* **2001,** *79*, 183-193.

27. Wheatley, R. W.; Kappelhoff, J. C.; Hahn, J. N.; Dugdale, M. L.; Dutkoski, M. J.; Tamman, S. D.; Fraser, M. E.; Huber, R. E. Substitution for Asn460 cripples beta-galactosidase (*Escherichia coli*) by increasing substrate affinity and decreasing transition state stability. *Arch. Biochem. Biophys.* **2012,** *521*, 51-61.

28. Roth, N. J.; Huber, R. E. The beta-galactosidase (*Escherichia coli*) reaction is partly facilitated by interactions of His-540 with the C6 hydroxyl of galactose. *J. Biol. Chem.* **1996,** *271*, 14296-14301.

29. Juers, D. H.; Heightman, T. D.; Vasella, A.; McCarter, J. D.; Mackenzie, L.; Withers, S. G.; Matthews, B. W. A structural view of the action of *Escherichia coli* (lacZ) beta-galactosidase. *Biochemistry.* **2001,** *40*, 14781-14794.

30. Juers, D. H.; Matthews, B. W.; Huber, R. E. LacZ beta-galactosidase: structure and function of an enzyme of historical and molecular biological importance. *Protein Sci.* **2012,** *21*, 1792-1807.

31. Boon, M. A.; Janssen, A. E.; van der Padt, A. Modelling and parameter estimation of the enzymatic synthesis of oligosaccharides by beta-galactosidase from *Bacillus circulans*. *Biotechnol. Bioeng.* **1999,** *64*, 558-567.

495 32. Vera, C.; Guerrero, C.; Illanes, A.; Conejeros, R. A pseudo steady-state model
496 for galacto-oligosaccharides synthesis with beta-galactosidase from *Aspergillus*
497 *oryzae*. *Biotechnol. Bioeng.* **2011,** *108*, 2270-2279.
498 33. Playne, M.J.; Crittenden, R.G. Galacto-oligosaccharides and other products
499 derived from lactose. In *Lactose, water, salts and minor constituents*, edition 3;
500 McSweeney, P.L.H., Fox, P.F., Eds.; Springer: New York, 2009; pp. 121–201
501 34. Urrutia, P.; Rodriguez-Colinas, B.; Fernandez-Arrojo, L.; Ballesteros, A.O.;
502 Wilson, L.; Illanes, A.; Plou, F.J. Detailed analysis of galactooligosaccharides
503 synthesis with β-galactosidase from *Aspergillus oryzae*. *J Agric Food Chem*. **2013,**
504 *61*, 1081-1087.
505 35. Depeint, F.; Tzortzis, G.; Vulevic, J.; I'Anson, K.; Gibson, G.R. Prebiotic
506 evaluation of a novel galactooligosaccharide mixture produced by the enzymatic
507 activity of *Bifidobacterium bifidum* NCIMB 41171, in healthy humans: a
508 randomized, double-blind, crossover, placebo-controlled intervention study. *Am J*
509 *Clin Nutr*. **2008**, *87*, 785–791.
510 36. Goulas, T.; Goulas, A.; Tzortzis, G.; Gibson, GR. Expression of four beta-
511 galactosidases from *Bifidobacterium bifidum* NCIMB41171 and their contribution
512 on the hydrolysis and synthesis of galactooligosaccharides. *Appl Microbiol*
513 *Biotechnol*. **2009**, *84*, 899-907.
514 37. Warmerdam, A.; Zisopoulos, F.K.; Boom, R.M.; Janssen, A.E. Kinetic
515 characterization of galacto-oligosaccharide (GOS) synthesis by three commercially
516 important β-galactosidases. *Biotechnol Prog*, **2014**, *30*, 38-47.
517 38. Song, J.; Imanaka, H.; Imamura, K.; Minoda, M.; Katase, T.; Hoshi, Y.;
518 Yamaguchi, S.; Nakanishi, K. Cloning and expression of a β-galactosidase gene of
519 *Bacillus circulans*. *Biosci Biotechnol Biochem*, **2011**, *75*, 1194-1197.

520 39. Yanahira, S.; Kobayashi, T.; Suguri, T.; Nakakoshi, M.; Miura, S.; Ishikawa, H.;
521 Nakajima, I. Formation of oligosaccharides from lactose by *Bacillus circulans* beta-
522 galactosidase. *Biosci Biotechnol Biochem*, **1995**, *59*, 1021–1026.
523 40. Placier, G.; Watzlawick, H.; Rabiller, C.; Mattes, R. Evolved beta-
524 galactosidases from *Geobacillus stearothermophilus* with improved
525 transgalactosylation yield for galacto-oligosaccharide production. *Appl. Environ.*
526 *Microbiol.* **2009,** *75*, 6312-6321.
527 41. Hansson, T.; Kaper, T.; van Der, O. J.; de Vos, W. M.; Adlercreutz, P. Improved
528 oligosaccharide synthesis by protein engineering of beta-glucosidase CelB from
529 hyperthermophilic *Pyrococcus furiosus*. *Biotechnol. Bioeng.* **2001,** *73*, 203-210.
530 42. Wu, Y.; Yuan, S.; Chen, S.; Wu, D.; Chen, J.; Wu, J. Enhancing the production of
531 galacto-oligosaccharides by mutagenesis of *Sulfolobus solfataricus* beta-
532 galactosidase. *Food Chem.* **2013,** *138*, 1588-1595.
533 43. Huber, R. E.; Hakda, S.; Cheng, C.; Cupples, C. G.; Edwards, R. A. Trp-999 of
534 beta-galactosidase (*Escherichia coli*) is a key residue for binding, catalysis, and
535 synthesis of allolactose, the natural lac operon inducer. *Biochemistry.* **2003,** *42*,
536 1796-1803.
537 44. Wheatley, R. W.; Lo, S.; Jancewicz, L. J.; Dugdale, M. L.; Huber, R. E. Structural
538 explanation for allolactose (lac operon inducer) synthesis by lacZ beta-
539 galactosidase and the evolutionary relationship between allolactose synthesis and
540 the lac repressor. *J. Biol. Chem.* **2013,** *288*, 12993-13005.
541 45. Jacobson, R. H.; Zhang, X. J.; DuBose, R. F.; Matthews, B. W. Three-
542 dimensional structure of beta-galactosidase from *E. coli*. *Nature.* **1994,** *369*, 761-
543 766.

46. Rodriguez-Colinas, B.; de Abreu, M. A.; Fernandez-Arrojo, L.; de, B. R.; Poveda, A.; Jimenez-Barbero, J.; Haltrich, D.; Ballesteros Olmo, A. O.; Fernandez-Lobato, M.; Plou, F. J. Production of galacto-oligosaccharides by the beta-galactosidase from *Kluyveromyces lactis* : comparative analysis of permeabilized cells versus soluble enzyme. *J. Agric. Food Chem.* **2011,** *59*, 10477-10484.

FUNDING:

This work was funded by grant BIO2013-48779-C4-3-R, from Spain's 'Secretaría de Estado de Investigación, Desarrollo e Innovación'. D T-P was supported by a FPU fellowship from 'Ministerio de Economía y Competitividad'.

555 **FIGURE CAPTIONS**

556

557 **Figure 1.** Structural analysis of TmLac and design of mutations. **A.** A schematic
558 representation of the domain arrangement of TmLac is depicted on top, showing
559 the central α/β barrel (α/β) surrounded by 5 β-sandwich domains (β1 to β5). Left
560 panel: Model of the main body of TmLac (residues 1-983) (C-score: -1.11). Right
561 panel: Model of the C-terminal domain (residues 984-1084) (C-score: -1.30). **B.**
562 Structural detail of the active site highlighting the catalytic residues (red) and the
563 residues targeted for site-directed mutagenesis (orange). β-3´-galactosyl-lactose is
564 depicted with the galactosyl moiety at the non-reducing end in blue. The rest of the
565 molecule, coloured in violet, represents the lactosyl moiety, acting as acceptor of
566 the galactosyl group during transgalactosylation. Dashed lines indicate distances of
567 3-4 Å between non-aromatic residues subjected to mutagenesis and the modeled
568 ligand. **C.** Sequence alignment of TmLac with other GH2 β-galactosidases with
569 transgalactosylating activity. The sequence stretches around the residues targeted
570 for site-directed mutagenesis are shown. The relative position of these stretches
571 within their corresponding domain is indicated on top. Genebank codes for the
572 indicated proteins are: *T. maritima*: AAD36268.1, *E. coli*: AAA24053.1, *L.*
573 *delbrueckii*: CAI98003.1, *K. lactis*: AAA35265.1, *B.*
574 *longum_sub_infantis*:_AAL02052.1, *L. acidophilus* LacL:_ABK59934.1, *L.*
575 *acidophilus*_LacM:_ABK59935.1, *L. reuteri*_LacL:_ ABF72116.1, *L.*
576 *reuterii*_LacM:_ABF72117.2, *L. plantarum*_LacL:_CAZ66936.1, *L.*
577 *plantarum*_LacM:_CAD65570.1

578

Figure 2. Effect of substitution of residues F571 and N574 on the synthesis of galacto-oligosaccharides (right panels) β-3´-galactosyl-lactose (black bars) and β-6´-galactosyl-lactose (grey bars). Error bars represent standard deviation of triplicates. Asterisks indicate significant differences ($p<0.01$) compared to wild-type data. Kinetics of glucose release is shown on the left panels as a control.

Figure 3. Effect of substitution of residues D568, W959, V93 and V94 on the synthesis of galacto-oligosaccharides (right panels) β-3´-galactosyl-lactose (black bars) and β-6´-galactosyl-lactose (grey bars). Error bars represent standard deviation of triplicates. Asterisks indicate significant differences ($p<0.01$) compared to wild-type data. Kinetics of glucose release is shown on the left panels as a control.

Figure 4. Structural model of mutants with higher transgalactosylating efficiency. β-3´-Galactosyl-lactose is depicted with the galactosyl moiety at the non-reducing end in light brown, and the lactosyl group acting as acceptor coloured in violet. **A**. Residues interacting with the galactosyl moiety at the non-reducing end are highlighted in green. The residues substituting N574 in the N574D and N574S mutants are overlapped and depicted in blue and orange, respectively. Dashed lines indicate putative hydrogen bonds at this specific position. **B**. The position of W959 and F571 in the wild-type enzyme is shown in green. The structure F571L mutant (orange) is overlapped and shows that the side chain of W959 is not sterically impeded to rotate to a different position.

Table 1

Enzyme	Total activity [µmol Glc · min^{-1} · µg $protein^{-1}$] (% of wild type activity)
wt	4.4 ± 0.9 (100)
V93T	4.15 ± 0.08 (94)
V94T	3.8 ± 0.3 (86)
V94Q	4.3 ± 0.7 (98)
D568S	4.8 ± 0.2 (110)
D568A	5.17 ± 0.17 (120)
F571L	3.1 ± 0.2 (70)
N574A	1.4 ± 0.3 * (31)
N574S	0.36 ± 0.06 * (8)
N574D	4.3 ± 0.3 (97)
W959F	0.22 ± 0.06 * (5)
W959C	0.94 ± 0.05 * (21)
W959A	2.0 ± 0.2 * (45)
F571L/N574S	0.41 ± 0.03 * (9)
F571L/N574A	0.30 ± 0.02 * (7)

* statistical difference ($p<0.01$) with wild-type value.

Figure 1

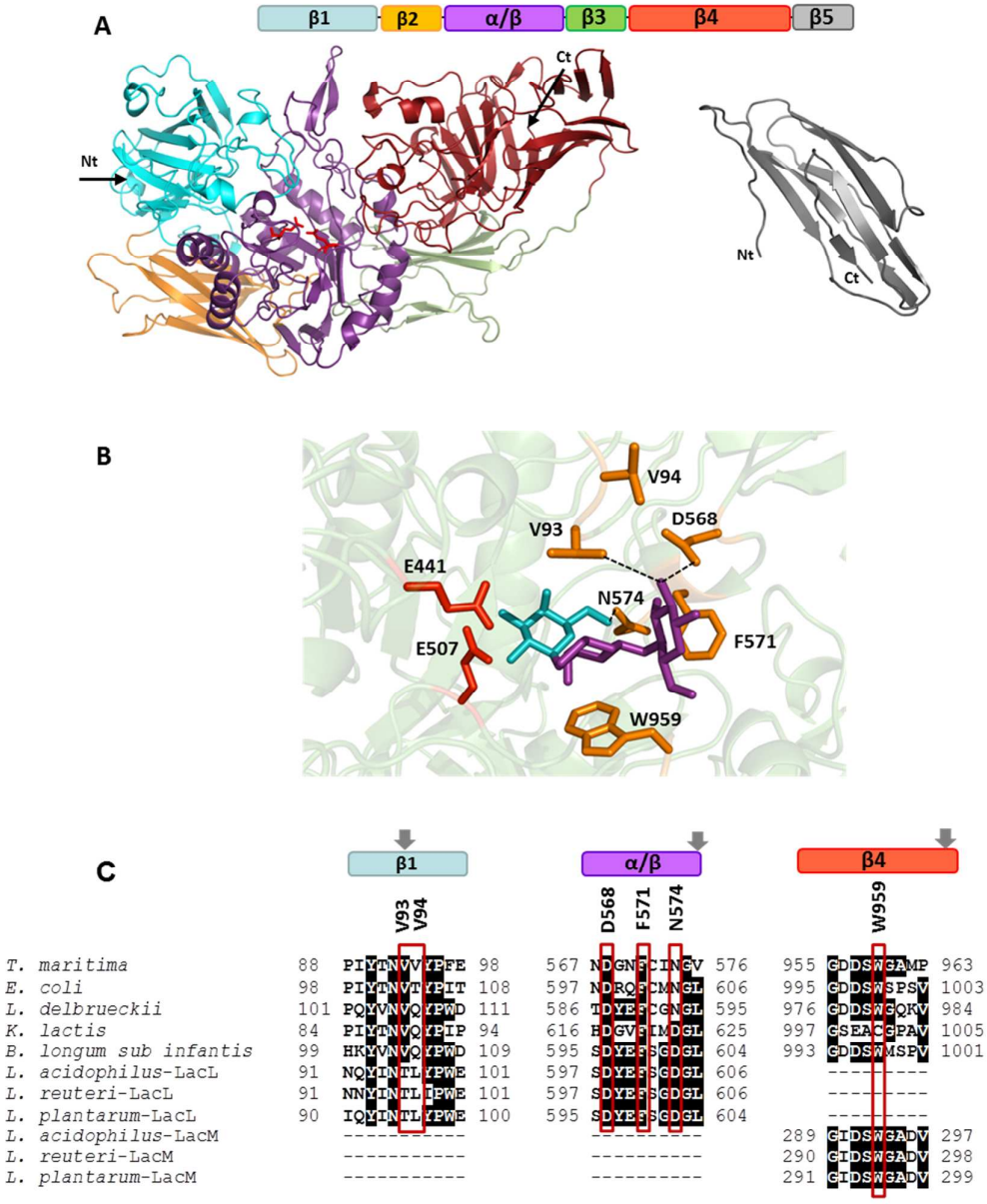

A
β1
β2
α/β
β3
β4
β5
Ct
Nt
Nt
Ct
B
V94
V93
D568
E441
N574
E507
F571
W959
C
β1
α/β
β4
V93
V94
D568
F571
N574
W959
T. maritima 88 PIYTNVVTYPFE 98 567 NDGNGCINGV 576 955 GDDSWGAMP 963
E. coli 98 PIYTNVTTYPIT 108 597 NDRQGCMNGL 606 995 GDDSWSPSV 1003
L. delbrueckii 101 PQYVNVQTYPWD 111 586 TDYEGCGNGL 595 976 GDDSWGQRV 984
K. lactis 84 PIYTNVQTYPIP 94 616 HDGVGIMDGL 625 997 GSEACGPAV 1005
B. longum sub infantis 99 HKYVNVQTYPWD 109 595 SDYEGSGDGL 604 993 GDDSWMSPV 1001
L. acidophilus-LacL 91 NQYTNTLTYPWE 101 597 SDYEGSGDGL 606
L. reuteri-LacL 91 NNYTNTLTIPWE 101 597 SDYEGSGDGL 606
L. plantarum-LacL 90 IQYTNTLTYPWE 100 595 SDYEGSGDGL 604
L. acidophilus-LacM
L. reuteri-LacM
L. plantarum-LacM
289 GIDSWGADV 297
290 GIDSWGADV 298
291 GIDSWGADV 299

Figure 2

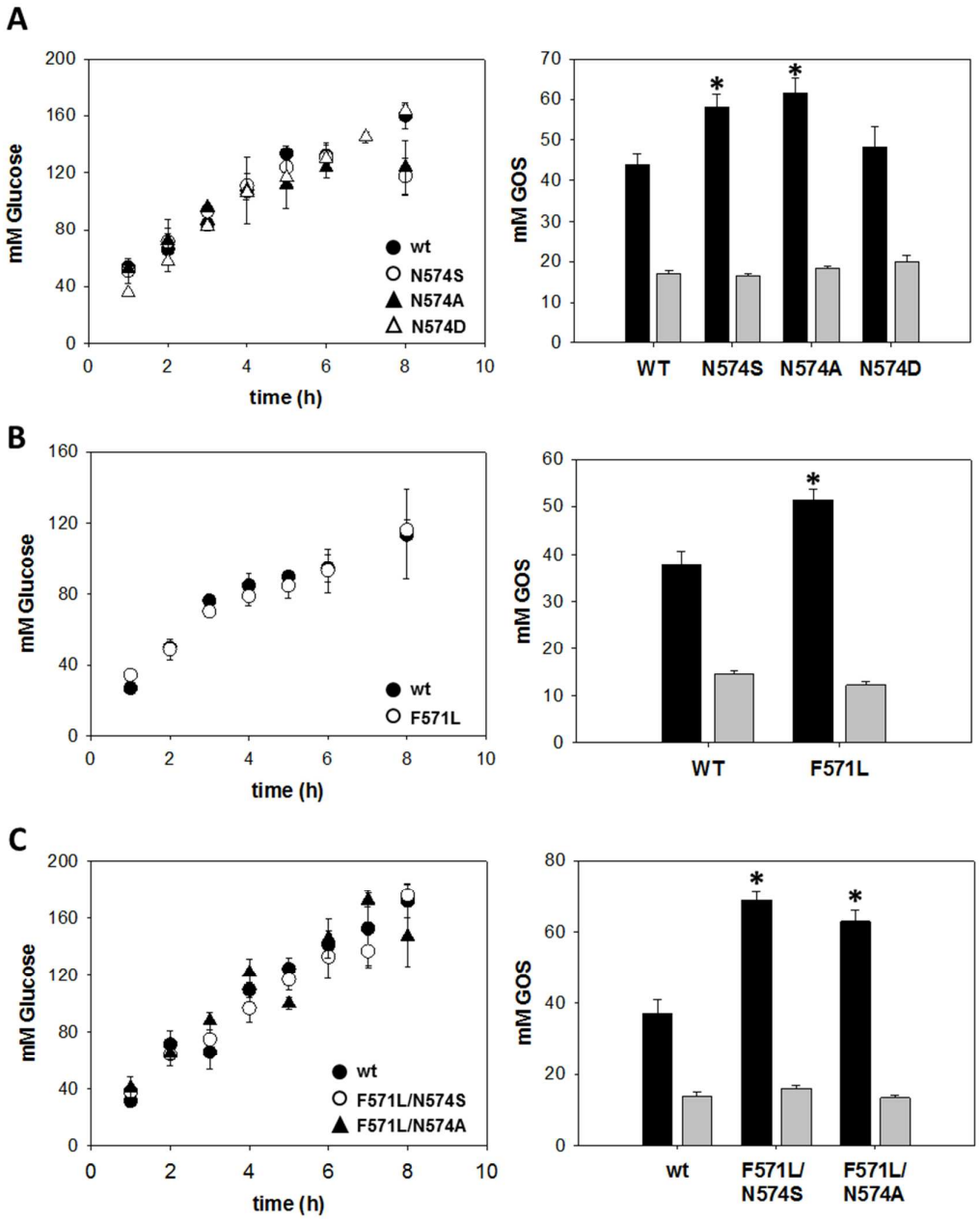

A
mM Glucose
time (h)
wt
N574S
N574A
N574D
mM GOS
WT
N574S
N574A
N574D
B
mM Glucose
time (h)
wt
F571L
mM GOS
WT
F571L
C
mM Glucose
time (h)
wt
F571L/N574S
F571L/N574A
mM GOS
wt
F571L/ N574S
F571L/ N574A

Figure 3

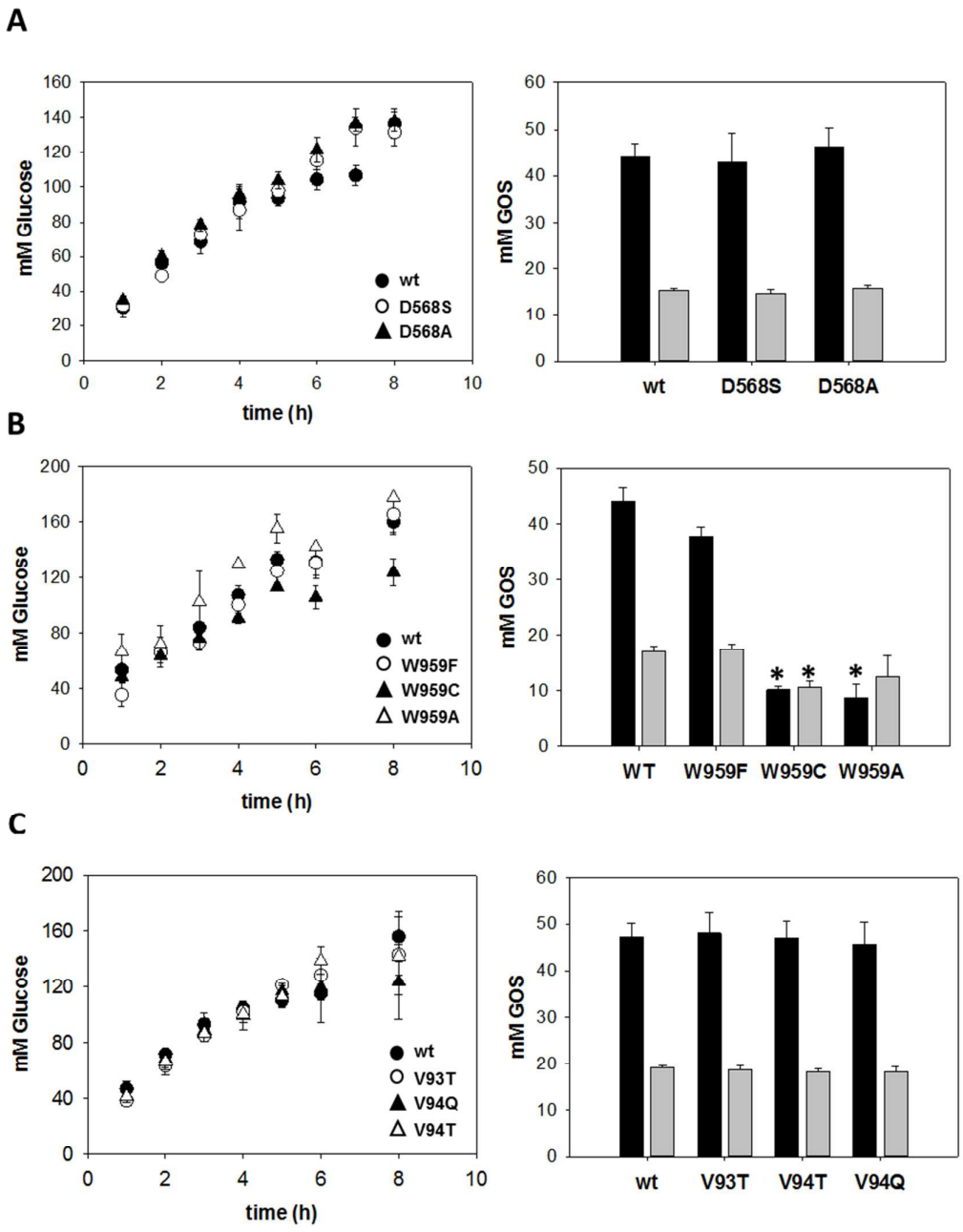
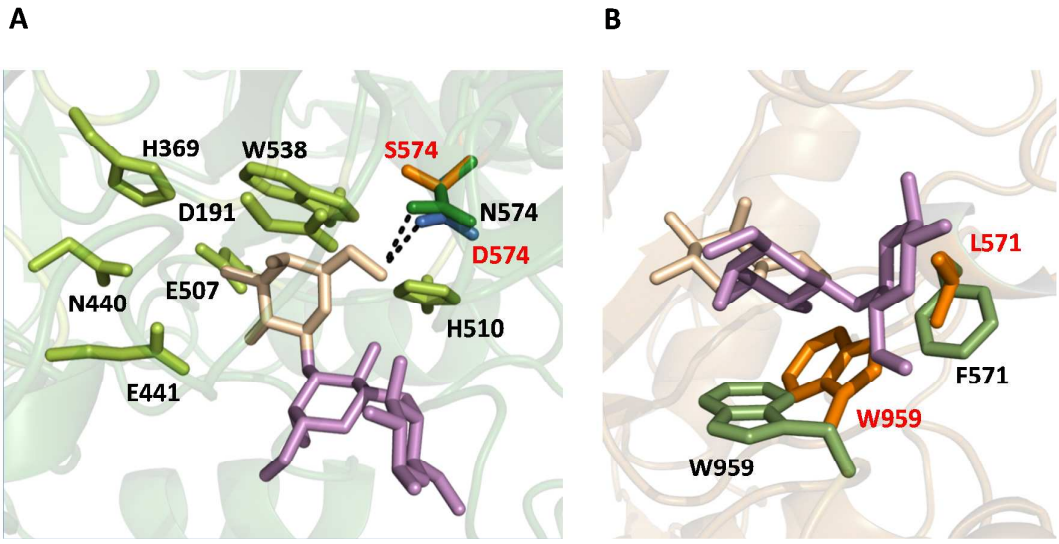

A
mM Glucose
time (h)
wt
D568S
D568A
mM GOS
wt
D568S
D568A
B
mM Glucose
time (h)
wt
W959F
W959C
W959A
mM GOS
WT
W959F
W959C
W959A
C
mM Glucose
time (h)
wt
V93T
V94Q
V94T
mM GOS
wt
V93T
V94T
V94Q

Figure 4


A
H369
W538
S574
D191
N574
D574
N440
E507
H510
E441
B
L571
F571
W959
W959

GRAPHIC FOR TABLE OF CONTENTS