Probing transglycosylation/hydrolysis equilibria in retaining glycoside hydrolases

Jiao Zhao

Université Fédérale
Toulouse Midi-Pyrénées

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Institut National des Sciences Appliquées de Toulouse

---

Présentée et soutenue par

Jiao ZHAO

Le 28 avril 2020

Investigation de l'équilibre transglycosylation/hydrolyse chez les glycoside hydrolases agissant avec rétention de configuration

---

Ecole doctorale : **SEVAB - Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingenieries**

Spécialité : **Ingénieries microbienne et enzymatique**

Unité de recherche :
TBI - Toulouse Biotechnology Institute, Bio & Chemical Engineering

Thèse dirigée par
Régis FAURÉ et Michael O'DONOHUE

Jury

Mme Caroline RÉMOND-ZILLIOX, Rapporteure
M. Marco MORACCI, Rapporteur
M. Sébastien FORT, Rapporteur
M. Antoni PLANAS, Président
Mme Isabelle ANDRÉ, Examinatrice
Mme Gerlind SULZENBACHER, Examinatrice
M. Régis FAURÉ, Directeur de thèse
M. Michael O'DONOHUE, Co-directeur de thèse

Thèse

Présentée devant

L'Institut National des Sciences Appliquées de Toulouse

En vue de l'obtention du

Doctorat

Spécialité : Ingénieries Microbienne et Enzymatique

Par **Jiao Zhao**

Probing transglycosylation/hydrolysis equilibria in retaining glycoside hydrolases

Directeurs de thèse

Dr. Régis FAURÉ et Dr. Michael J. O'DONOHUE

NOM: ZHAO **PRÉNOM**: Jiao

TITRE: **Investigation de l'équilibre transglycosylation/hydrolyse chez les glycoside hydrolases agissant avec rétention de configuration**

SPÉCIALITÉ: Ingénieries microbienne et enzymatique

ANNÉE: 2020 **LIEU**: INSA, TOULOUSE

DIRECTEURS DE THÈSE: Dr. Régis FAURÉ et Dr.Michael J. O'DONOHUE

Résumé:

La synthèse de sucres est importante et pertinente, car applicable au développement de diagnostics et de molécules thérapeutiques. Cependant, la glycochimie reste laborieuse, en particulier pour la synthèse de furanosides, ce type de synthèse figurant parmi les plus difficiles à réaliser. Néanmoins, les furanosides sont d'intérêt, étant présents dans de nombreux biopolymères et chez certains microorganismes pathogènes. Ainsi il est pertinent d'étudier les furanoside hydrolases en tant qu'outils glycosynthétiques potentiels. Dans ce cadre, l'α-L-arabinofuranosidase de *Thermobacillus xylanilyticus* (*Tx*Abf) de la famille GH51 a été étudiée et modifiée pour créer des transfuranosylases efficaces.

Ici, poursuivant nos études de *Tx*Abf, quatre aspects sont abordés: (i) Une étude structure-fonction porte sur deux déterminants du sous-sites -1. Cette étude révèle que le résidu L352 intervient dans les deux sous-sites -1 et +1, tandis que l'action de F26 reste limitée au sous-site -1. De plus, le triple mutant R69H-N214W-L352M provoque un effet domino, augmentant la flexibilité locale, créant ainsi de nouvelles interactions avec les molécules acceptrices. (ii) L'étude de l'autocondensation a permis d'augmenter les rendements globaux en homo-oligofuranosides. Fait significatif, la substitution N216W dans *Tx*Abf génère un sous-site de liaison supplémentaire et une modification de regiosélectivité. (iii) Concernant la synthèse de D-galactofuranosides, le mutant R69H-N216W s'avère être un outil de choix. Toutefois, sa regiosélectivité vers des liens (1,3) peut-être augmenté par l'introduction d'autres mutations. La synthèse de D-Gal*f*-(1,3)-α-D-Glc*p* et de β-D-Gal*f*-(1,n)-α-D-GlcpNAc sont également décrites. (iv) Travaillant avec d'autres équipes de recherche, il a été possible de démontrer l'utilité et la généricité d'une stratégie d'ingénierie des protéines simple permettant de rapidement concevoir de puissantes transglycosylases, à partir de glycoside hydrolases agissant par rétention.

Mots-clés: transglycosylation, α-L-arabinofuranosidase, promiscuité catalytique, interactions moléculaires, mutagénèse, furanoses

---

Ecole doctorale: SEVAB - Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingenieries

Unité de recherche: TBI - Toulouse Biotechnology Institute, Bio & Chemical Engineering

SURNAME: ZHAO **FIRST NAME**: Jiao

TITLE: Probing transglycosylation/hydrolysis equilibria in retaining glycoside hydrolases.

SPECIALITY: Enzymatic and Microbial Engineering

YEAR: 2020 **PLACE**: INSA, TOULOUSE

THESIS DIRECTORS: Dr. Régis FAURÉ and Dr.Michael J. O'DONOHUE

Summary:

Carbohydrate synthesis is important for a number of areas, including the development of diagnostics and chemotherapeutics. However, glycochemistry is challenging especially when furanose-based motifs are targeted. Indeed, furanosides figure prominently among the most difficult carbohydrates to synthesize. Nevertheless, they are of interest, because they are important components of many biopolymers and are present in many pathogenic microorganisms. For this reason, it is pertinent to study furanoside hydrolases as potential glycosynthetic tools. In pursuit of this goal, the family GH51 α-L-arabinofuranosidase from *Thermobacillus xylanilyticus* (*Tx*Abf) has been extensively studied and engineered to create efficient transfuranosylases.

Herein, work on *Tx*Abf is pursued, focusing on four aspects: (i) A structure-function study was used to probe two subsite -1 determinants. Study of F26L and L352M revealed that residue L352 intervenes in both subsites -1 and +1, whereas F26 is limited to subsite -1. Moreover, the triple mutant R69H-N216W-L352M causes a domino-like effect, increasing local flexibility, thus creating new interactions with acceptor molecules. (ii) Studying self-condensation using mutagenesis afforded increased overall yields of homo-oligofuranosides. Significantly, the substitution N216W in *Tx*Abf engenders an additional binding subsite that drives altered regioselectivity. (iii) Focusing on the synthesis of D-galactofuranosides, mutant R69H-N216W proved to be a tool of choice for transgalactofuranosylation. However, further mutagenesis afforded higher regioselectivity towards (1,3)-linkages. Likewise, the synthesis of β-D-Gal*f*-(1,3)-α-D-Glc*p* and Gal*f*-(1,n)-α-D-Glc*p*NAc are reported. (iv) Working within a wide consortium of research teams, it has been possible to demonstrate the usefulness and genericity of a simple protein engineering strategy that can be used to rapidly design potent transglycosylases, starting from retaining glycoside hydrolases.

Keywords: transglycosylation, α-L-arabinofuranosidase, catalytic promiscuity, molecular interactions, mutagenesis, furanoses

---

Doctoral school: SEVAB - Ecological, Veterinary, Agronomic Sciences and Bioingeneering

Research institute: TBI - Toulouse Biotechnology Institute, Bio & Chemical Engineering

Scientific production

Publications

- Teze, David, Jiao Zhao, Mathias Wiemann, Zubaida Gulshan Kazi, Rossana Lupo, Mette Errebo Rønne, Göran Carlström, et al. 2020. "Rational Enzyme Design Without Structural Knowledge: A Sequence-Based Approach for Efficient Generation of Glycosylation Catalysts." *ChemRxiv*. https://doi.org/10.26434/chemrxiv.11538708.v1.

- Jiao Zhao, Tobias Tandrup, Bastien Bissaro, Sophie Barbe, Isabelle André, Claire Dumon, Leila Lo Leggio, Michael J. O'Donohue and Régis Fauré[1]. "Probing the determinants of the transglycosylation/hydrolysis partition in a retaining α-L-arabinofuranosidase." *To be submitted*.

- Jiao Zhao, Jérémy Esque, Isabelle André, Michael J. O'Donohue, and Régis Fauré.. "Synthesis of α-L-Araf and β-D-Galf Series Difuranosides Using Mutants of a GH51 α-L-Arabinofuranosidase." *To be submitted*.

- Jiao Zhao, Jérémy Esque, Olivier Saurel, Isabelle André, Michael J. O'Donohue, and Régis Fauré. "Investigation of transgalactofuranosylation activity in an α-L-arabinofuranosidase" *In Preparation*.

Oral and poster communications

-13th Carbohydrate Bioengineering Meeting (19-22 May 2019, Toulouse, France) Molecular study of hydrolysis/transglycosylation modulation in retaining glycoside hydrolases. J. Zhao, T. Tandrup, B. Bissaro, S. Barbe, I. André, C. Dumon, L. Lo Leggio, M.J. O'Donohue, R. Fauré (oral communication)

-13th Carbohydrate Bioengineering Meeting (19-22 May 2019, Toulouse, France) Efficient conversion of GHs into transglycosylases: a conservation-based approach. D. Teze, J. Zhao, M. Wiemann, K.Z.G. Ara, T. Mazzarella, J. Duus, Y.-H. Sanejouand, C. André-Miral, C. Tellier, E. Nordberg-Karlsson, R. Fauré, H. Stålbrand, B. Svensson (oral communication)

-14th International Symposium on Biocatalysis and Biotransformations - BioTrans (7-11 July 2019, Groningue, the Netherlands) Exploiting catalytic promiscuity in glycoside hydrolases to modulate the hydrolysis/transglycosylation partition. J Zhao, T. Tandrup, B. Bissaro, S. Barbe, I. André, C. Dumon, L. Lo Leggio, M.J. O'Donohue, R. Fauré (Poster)

-15th European Training Course on Carbohydrates: Summer Course Glycosciences (24-28 June 2018, Wageningen, the Netherlands) Molecular study of

hydrolysis/transglycosylation modulation in retaining glycoside hydrolases. J. Zhao, T. Tandrup, B. Bissaro, L. Lo Leggio, C. Dumon, M.J. O'Donohue, R. Fauré (Poster)

Acknowledgement

Undertaking this three-and-half-year Ph.D has been a truly life-changing experience for me and it would not have been possible to accomplish without the support and guidance that I received from many people.

I would like to first say a very big thank you to my dear supervisors Régis and Mike. Both of them are charming, gentle and considerate. Régis has a rich sense of responsibility and gave me a lot of support and encouragement. He is always by my side whenever I need him, and never loses his patience to me. From Mike, I learned how to think more like a scientist. He is always encouraging me to find something interesting behind the phenomenon. And he is so thoughtful and makes me touched and feel reliable on him. It is my great fortune to meet these two lovely supervisors in my life.

Also, I sincerely thank my dear E-carb group, who gives me a home feeling. Thanks, Claire, Cédric, and Sophie. B for your valuable suggestions during each group meeting and in the routine lab work. And of course, I won't forget our former members Haiyang, Victor, Louise and Lénaïk who supplied me a lot of help especially at the hard beginning of my Ph.D.

Those my dear friends, Zhi, Manon. M, Awilda, Julien, Maher, Gema, Thomas, Manon. C, Jeremy, Ao, Emna, Mounir, Laurence, Marie, Emma, Akli and …, I can't even count how many times you helped me without any hesitation. You fulfilled my happiness for those periods. Recipe change, picnics, discovering new restaurants, parties, travels, skiing…I had so many wonderful memories with you, and those memories will be my happiness reservoir.

I am also very grateful to all those gave me a lot of technical supports, Guy, Edern, Lindsay, Neil, and Nelly, who were always trying their best to help me pave away the difficulties in the experiments.

I would like to thank my committee members, Toni and Isabelle, who provide great help for my thesis progress and inspired me a lot to boost my thesis work. Thanks Leila, Sophie. Barbe, Tobias, Olivier, David Teze, thanks to your excellent accomplishment of the collaboration parts, thus allowed me to have this complete thesis work. I also want to give a big thank to my jury members Dr. Rémond, Dr. Moracci, Dr. Fort and Dr. Sulzenbacherall, who are willing to spend time on reading and evaluating my thesis work, and even gathering 'online' to give me advises for further improvement.

I would also like to thank all current and past members of the Enzyme molecular engineering and catalysis group. Thanks Magali for leading this team into a nice working atmosphere and remove the language barrier for the foreign students. Thanks Yannick, Nathalie, Zhongpeng, Alex.T, Youness, Vincinane, Vincent, Maxant, Sarah, Katie, Audré, Julien. R, Pauline, Eli, Alex. G, all of you are so warm-hearted, make me feel not strange in the unacquainted country.

In the end, I am so grateful to my family, for the sustaining support from my father and my husband during my PhD study. Here I also want to express my memory of my dear mother who accompanied me as my spiritual pillar. In addition, I will say thanks to my dear friend HuiHui, the one we encouraged each other in these years.

I might forget some people. I hope you will not be mad at me. I do really enjoy the time in France, the time with all of you. It will be the treasure of my life. PS: We will defeat the coronavirus soon!

君子欲讷于言而敏于行。

—摘孔子《论语》

The Master said, "The superior man wishes to be slow in his speech and earnest in his conduct.”

- Confucius 《The Analects》

Table of Contents

General introduction

General introduction

The synthesis of carbohydrates is a major topic, because the biomedical applications for robust strategies are numerous and include the development of diagnostics, therapeutics, and vaccines. However, while Nature provides the demonstration of how an enormous diversity of carbohydrate-based structures can be made using sophisticated biological machinery, reproducing this prowess in the chemistry laboratory is not straightforward. Therefore, for many years now, synthetic chemists have been exploring the use of enzymes to generate various glycoconjugates. Among the enzymatic solutions that are studied, retaining glycoside hydrolases (rGHs) are attractive, because of their relative simplicity (no cofactors or expensive donor sugars are required) and diversity in terms of substrate specificity.

As their name implies, rGHs are hydrolytic enzymes, which operate via a double displacement mechanism and possess the potential to synthesize glycosidic bonds through transglycosylation, a reaction that competes with hydrolysis. Using protein engineering, it is possible to manipulate the transglycosylation/hydrolysis (T/H) partition, which is specific property of any given rGH in defined conditions.

Within the field of carbohydrate chemistry, furanosides are notoriously difficult to access, because of the thermodynamic instability of the 5-membered furanose ring. This is unfortunate, because furanosides are biologically-relevant, being part of carbohydrate structures found in plants and also in a variety of pathogenic microorganisms. Therefore, the use of furanose-specific rGHs is an attractive alternative to classical synthetic chemistry methods.

To address rGH-mediated furanoside synthesis, herein we pursue work on a family GH51 α-L-arabinofuranosidase, from *Thermobacillus xylanilyticus* (*Tx*Abf). In its wild-type form, this robust hydrolytic enzyme displays some ability to perform transglycosylation, especially using D-galactofuranosides as glycosyl sugar donor. Previous work using *in vitro* mutagenesis methods has demonstrated that it is possible to shift the T/H partition in favour of transglycosylation. In the present study we pursue this line of work, combining a variety of strategies, including enzymology, structural biology and *in silico* approaches to further our understanding of how the T/H partition can be altered, pinpoint molecular determinants of regioselectivity and explore how previous knowledge gained on the synthesis of L-arabinofuranosides can be transposed to the *Tx*Abf-mediated synthesis of D-galactofuranosides.

In Chapter II, during the course of this work on *Tx*Abf, we revealed the different impacts of two donor subsite substitutions F26L and L352M, revealing that the latter intervenes in both subsites -1 and +1. Importantly, when introduced into the mutated genotype R69H-N216W, L352M causes a domino-like effect that increases the flexibility of two loops bearing W248 and W302, residues that form important interactions with acceptors. In turn, this fundamental study suggested that introducing flexibility in either acceptor subsites loops or acid/base residues by subtle molecular modifications in active subsites is an alternative strategy to enhance T/H in rGHs. In Chapter III, working on self-condensation (a reaction where the donor and acceptor are the same molecular species) and using mutagenesis, it was possible to increase overall yields of oligo-L-arabino- and oligo-D-galactofuranosides when compared to the performance of the wild-type enzyme. Significantly, this study revealed that the substitution N216W leads to the creation of an additional binding subsite that provides the basis for altered regioselectivity. Moreover, in Chapter IV, focusing on the specific question of how to synthesize D-galactofuranosides, we found that the mutant R69H-N216W is a tool of choice for general transgalactofuranosylation, but other mutants catalyze more regioselective reactions, providing access to the biologically relevant β-D-Gal*f*-(1,3)-α-D-Glc*p* motif. Likewise, the transfer of D-galactofuranosyl unit onto α-D-Glc*p*NAc moiety is demonstrated. Finally, in Chapter V, working within a wide consortium of research teams, it has been possible to demonstrate the usefulness and genericity of a simple protein engineering strategy that can be used to rapidly procure potent transglycosylases, starting from rGHs.

Overall, this work contributes considerable a number of results to a growing corpus of knowledge that increasingly provides the means to supply potent enzymes to the field of synthetic glycochemistry, particularly for the synthesis of difficult to access furanosides.

Chapter I.

Literature review

Abbreviations

α-L-Ara*f*O*p*NP: 4-nitrophenyl α-L-arabinofuranoside

Abfs : α-L-arabinofuranosidases

AfT : arabinofuranosyltransferase

AG : arabinogalactan

AS : amylosucrases

AX : arabinoxylan

AXOS : arabinoxylo-oligosaccharides

A^2XX : α-L-Ara*f*-(1,2)-β-D-Xyl*p*-(1,4)-β-D-Xyl*p*-(1,4)-D-Xyl*p*

A^3XX : α-L-Ara*f*-(1,3)-β-D-Xyl*p*-(1,4)-β-D-Xyl*p*-(1,4)-D-Xyl*p*

β-D-Fuc*f*O*p*NP: 4-nitrophenyl β-D-fucofuranoside

β-D-Gal*f*O*p*NP: 4-nitrophenyl β-D-galactofuranoside

CAZy : Carbohydrate-Active Enzymes

CGTases : cyclodextrin glucanotransferases

DP : degree of polymerization

D-Xyl*p* : D-Xylopyranosyl unit

FTs : fructosyltransferases

FS : fructose-specific sucrases

iGHs : inverting glycoside hydrolases

D-Gal*f* : D-Galactofuranosyl unit

GHs : glycoside hydrolases

GlfT : galactofuranosyltransferase

GS : glucansucrases

GPs : glycoside phosphorylases

GTs : glycosyltransferases

LAM : lipoarabinomannan

LG : leaving group

L-Ara*f* : L-arabinofuranosyl unit

mAG : mycolyl arabinogalactan

QM/MM : quantum mechanics/molecular mechanics

rGHs : retaining glycoside hydrolases

SAs : sialidases

TcTS : *Trypanosoma. cruzi* trans-sialidase

trS : trans-sialidases

TrSA : *Trypanosoma rangeli* sialidase

TGs : transglycosylases

*Tx*Abf : α-L-arabinofuranosidases from *Thermobacillus xylanilyticus*

T/H : transglycosylation/hydrolysis ratio

VI : vacuolar invertases

X_3 : (1,4)-β-D-xylotriose

XA^3XX : β-D-Xyl*p*-(1,4)-[α-L-Araf-(1,3)]-β-D-Xyl*p*-(1,4)-β-D-Xyl*p*-(1,4)-D-Xyl*p*

XA^3X : β-D-Xyl*p*-(1,4)-[α-L-Araf-(1,3)]-β-D-Xyl*p*-(1,4)-D-Xyl*p*

XEHs : xyloglucan *endo*-hydrolases

XETs : xyloglucan *endo*-transglycosylases

1. Glycosynthesis and interests of the carbohydrates

Carbohydrates are the most abundant biopolymers and the predominant component of biomass on earth, being an integral part of animals, plants, and bacteria (Seeberger 2015). Within living organisms, carbohydrates often play vital roles in biological processes. Therefore the study and synthesis of carbohydrates is of considerable relevance for the development of strategies for diagnosis, therapeutics, and vaccine development (Muthana, Cao, and Chen 2009). However, carbohydrates are characterized by an impressive level of chemical and structural diversity. A large number of monosaccharide subunits displaying different chemical structures and configurations can be combined via a variety of glycosidic linkages (Laine 1994), forming a myriad of oligomeric and polymeric structures. For this reason, compared to the study of other biological polymers such as nucleic acids or proteins, the study of carbohydrates is less advanced, in particular because there is a lack of straightforward synthetic methods that can be readily automatized.

Unlike the synthesis of nucleic acids and proteins that are biosynthesized through a template-guided processes (Seeberger and Werz 2007), carbohydrates are built up stepwise in successive enzyme-catalyzed reactions. Therefore, the development of efficient and general synthetic routes for carbohydrates remains a major challenge. Fortunately, a number of current chemical and enzymatic methods already allow the construction of a range of structurally defined oligosaccharides, glycoproteins, and other glycoconjugates (Kiessling and Splain 2010; Krasnova and Wong 2016). However, in spite of intensive efforts, many questions remain unclear, notably regarding the underlying mechanisms that control the formation of glycosidic bonds (i.e. stereo- and regioselectivity).

1.1. Chemical glycosylation

Glycochemistry is an old discipline that traces its origins back to the pioneering work performed by Arthur Michael and Emil Fischer, 140 years ago. These were the first scientists to perform chemical glycosylation (M. M. Nielsen and Pedersen 2018). A general approach for oligosaccharide assembly requires a glycosyl donor, a glycoside acceptor, and a promotor. The glycosylation process generates a new stereocenter and is characterized by the realization of protection/deprotection of hydroxyl groups, this being required to introduce regioselectivity. Accordingly, the selective exposure of unique hydroxyl groups provides the means to control the glycosylation site on the acceptor. Frequently, the donor is activated with a promoter (*e.g.* a Lewis acid) and its anomeric carbon is attacked by the free hydroxyl group from the acceptor, forming a new glycosidic linkage (Wen et al. 2018). For example, one of the most famous glycosylation reactions still in common use is the Koenigs-Knorr reaction, where glycosyl halides are used as donors under activation by silver carbonate (Koenigs and Knorr 1901).

In chemical glycosylation, the control of chemo-, regio- and stereo-selectivity depends on several factors. For example, one needs to carefully consider the choice of leaving group at the anomeric position of the donor, the protection/deprotection strategy applied to both the donor and acceptor, the solvent system and the choice of promoter (Park et al. 2007). Consequently，chemical glycosylation is relatively labor-intensive and time-consuming, the choice of suitable orthogonal protecting groups and their selective manipulation being the most difficult tasks to accomplish (Krasnova and Wong 2016).

Several convergent glycosylation strategies are proposed to replace stepwise glycosylation. These can be one-pot (Panza et al. 2018; Huang et al. 2004), automated (Kinnaert et al. 2017; Seeberger and Werz 2007; Wen et al. 2018) or chemo-enzymatic synthesis (Muthana, Cao, and Chen 2009), and all aim to economize on time and increase efficiency. One-pot strategies streamline glycan synthesis, because all glycosylations are performed in a single reaction vessel (pot) and do not require purification of intermediates (Huang et al. 2004). Automated synthesis, also known as Automated Glycan Assembly (AGA), has progressed over the two last decades from an idea to a technology that produces glycans as molecular tools for various applications. The hardware (*i.e.* the automated synthesizer) and the relevant consumables (*i.e.* linker-functionalized resins and monosaccharide building blocks) are now commercially available, thus access to this methodology is increasingly facilitated (Guberman and

Seeberger 2019; Panza et al. 2018). Despite these advances, when compared to other classes of biomolecules, the use of automated methods for carbohydrate synthesis is still underdeveloped. However, the use of advanced technologies that incorporate enzymatic components are likely to overcome current bottlenecks. Therefore convergent chemo-enzymatic synthetic routes are interesting targets for R&D aimed at the development of new ways to access carbohydrates.

1.2. Enzymatic glycosylation

In Nature, carbohydrate-active enzymes (CAZymes, which are classified within the CAZy database; www.cazy.org and www.cazypedia.org, (Lombard et al. 2014) catalyze the synthesis and hydrolysis of glycosyl-containing compounds. Regarding, carbohydrate biosynthesis, this is mostly performed by so-called Leloir glycosyl transferases (GTs), glycoside phosphorylases (GPs), and glycoside hydrolases (GHs), the latter including naturally-occurring non-Leloir transglycosylases (TGs). In this respect, the synthesis of glycosidic bonds by CAZymes has been studied for over 60 years (Edelman 2006). Importantly, like most enzymes, CAZyme provide the means to control regioselectivity, stereospecificity and operate under mild physicochemical conditions. Consequently, when used as catalysts for the purposes of synthetic chemistry, CAZymes offer the possibility to dispense with protection/deprotection methods and the use of organic solvents, and thus constitute an attractive alternative (or complementarity) to traditional glycochemistry (Krasnova and Wong 2016; Filice and Marciello 2013). Among CAZymes, the main enzymes that catalyze glycosylation are shown in Fig. 1 (Benkoulouche et al. 2019). These are classified according to reaction itinerary along which glycosidic bond formation occurs. In section I-2 of Part I, discussion will be focused on the naturally-occurring GTs and TGs. The modified enzyme glycosynthase (GS) will be discussed in section 2.1.

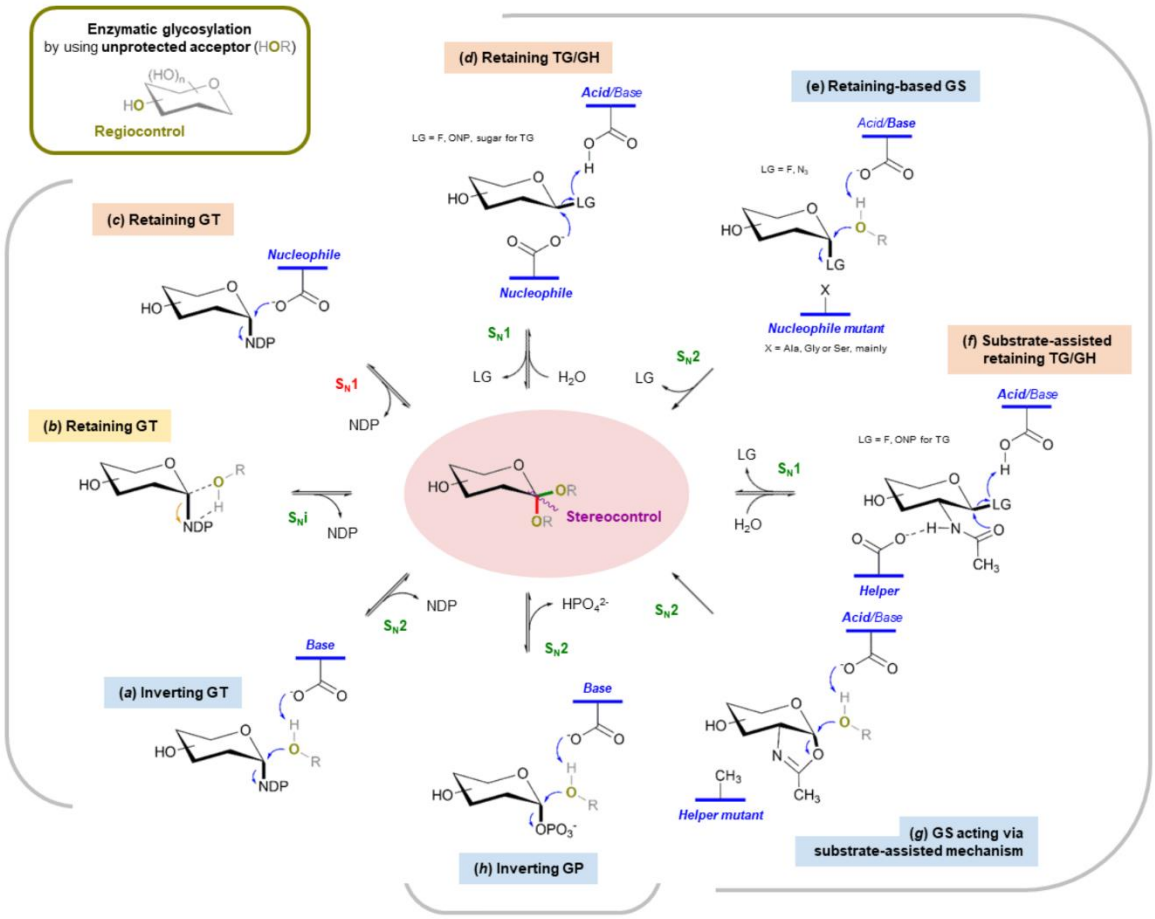


Fig. 1. Main enzymatic pathways for glycosidic bond formation by wild-type and engineered glycoenzymes mentioned in this chapter (GT, TG/GH and GS, and GP). HOR, unprotected acceptor; LG, leaving group; GH, glycoside hydrolase; GP, glycoside phosphorylase; GS, glycosynthase; GT, glycosyltransferase; TG, transglycosylase; NDP, nucleotide diphosphate; ONP, *O*-nitrophenyl derivative (Benkoulouche et al. 2019).

1.2.1. Glycosyltransferases (GTs)

The main group of glycosidic bond-forming enzymes are Leloir-type GTs, which were found by Leloir and his coworkers in 1950s (Cardini et al. 1950; Horton 2008). GTs are responsible for the synthesis of most cell-surface glycoconjugates in mammalian systems and cell-wall carbohydrates in plants, fungi and bacteria (Lairson et al. 2008). GTs catalyze the transfer of glycosyl moieties from activated glycosyl donors, typically a nucleotide sugar, onto glycoside and non-glycoside acceptors. Chemically, GT-catalyzed reactions are nucleophilic substitutions at the donor anomeric carbon and proceed with either inversion or retention of configuration (*a*, *b* and *c* in Fig. 1, Nidetzky, Gutmann, and Zhong 2018). Although GTs are very efficient, their industrial implementation is complicated by the high price of nucleotide-

activated donors. In addition, the catalysts themselves are difficult to work with, as many of them are membrane bound proteins that have proven to be difficult to express at high-level (Nidetzky, Gutmann, and Zhong 2018; Winkler et al. 2018). Enhancing enzyme expression to enable efficient whole cell biocatalysis is thus a major R&D pursuit. In addition to the selection of suitable expression hosts and genetic constructions, protein design is also an important factor in order to facilitate the expression of GTs in soluble form. To achieve this, several strategies have been tested, including targeted truncation to remove membrane-anchoring parts, fusion to solubility enhancing modules and removal or substitution of aggregation prone elements) (Ortiz-Soto and Seibel 2016; Choi et al. 2016; Wakarchuk et al. 1998). Moreover, to improve GTs as tools for the synthetic chemistry, focus has been given to nucleotide recycling cascades. This has allowed for the efficient production and reuse of nucleotide glycosyl sugar donors in robust one-pot multi-enzyme glycosylation cascades. Consequently, taking into account the most recent advances, Leloir GTs are now close to being considered as industrially amenable enzymes that can be deployed to achieve multi-enzyme, programmable cascade glycosylations (Mestrom et al. 2019; Schmölzer, Lemmerer, and Nidetzky 2018).

1.2.2. Glycoside phosphorylases (GPs)

All the GPs known so far are classified in both GT and GH families (Garron and Henrissat 2019). These enzymes catalyze phosphorolysis, transferring a glycosyl moiety from the non-reducing end of a poly- or oligosaccharide glycosyl donor to inorganic phosphate, thus generating glycosyl-1-phosphates. Practically, there could be two types of GPs in regard to anomeric configuration of the glycosyl donor and the resulting glycosyl phosphate, inverting (*h* in Fig. 1) and retaining GPs (sharing mechanisms *b-d* in Fig. 1) (Benkoulouche et al. 2019; Puchart 2015). Significantly, the high energy of the glycosyl phosphate provides the basis for reaction reversibility, meaning that GPs can be employed for synthetic purposes using appropriate glycosyl phosphates and acceptors (Puchart 2015; Suzuki et al. 2009; Nakai et al. 2013). Although the use of GPs for the synthesis of oligosaccharides is so far limited (due to the rather small diversity of reported phosphorylases), this shortcoming is gradually being overcome thanks to recent discoveries of novel GPs (Macdonald et al. 2019, 2018; Franceus et al. 2019).

1.2.3. Glycoside hydrolases (GHs)

GHs are enzymes responsible for cleaving glycosidic bounds that connect carbohydrates to either other carbohydrates or other bio-molecules. They are attractive enzymes because they are readily available, cheap, robust and display a very extensive range of substrate specificities (Bissaro, Monsan, et al. 2015). As of the 4th of March, 2020, GHs are grouped within the CAZy database into 161 families and 18 different clans. Based on the catalytic mechanism, GHs are classified either as inverting or retaining GHs, the latter being more abundant (67% of all 755 496 classified GH modules among 83 of the 161 GH families (data recorded on February 24th, 2020).

1.2.3.1. Catalytic mechanisms

Like GTs, GH-mediated hydrolysis of glycosidic bonds results in one of two stereochemical outcomes, either retention or inversion of the anomeric configuration. Therefore, GHs are classified as either retaining (rGHs) or inverting (iGHs) (S. G. Withers 1999). Elementary mechanisms for these two classes of glycosidases were initially proposed by Koshland (Koshland 1953). Catalysis by rGHs and iGHs relies on the intervention of two catalytic residues, which are usually carboxylic acid groups. Sometimes this dyad is completed by a third residue that plays a catalytic "helper" role, providing nucleophile stabilization and thus facilitating catalysis (S. Withers 2001; Sinnott 1990; Debeche et al. 2002).

The iGHs use a direct displacement mechanism. This means that the two catalytic carboxylic acids are suitably positioned to perform catalysis without the need for a formation of a covalent catalytic intermediate. In this reaction, one catalytic carboxylate acting as base activates an incoming water molecule (when R=H, Fig. 2), which in turn launches a nucleophilic attack on the anomeric carbon of the glycosyl moiety bound in subsite -1 (Davies, Wilson, and Henrissat 1997). Simultaneously, the second catalytic carboxylate acting as an acid donates a proton to the interglycosidic oxygen of the scissile glycosidic bond. During the reaction catalyzed by iGHs, a transient oxocarbenium ion-like transition state is formed. A distinctive feature of iGHs concerns the distance that separates the two catalytic carboxylates. Generally this can be up to 10 Å, which is greater than that observed in GHs that operate via a retaining mechanism (S. G. Withers 1999; G. Davies and Henrissat 1995).

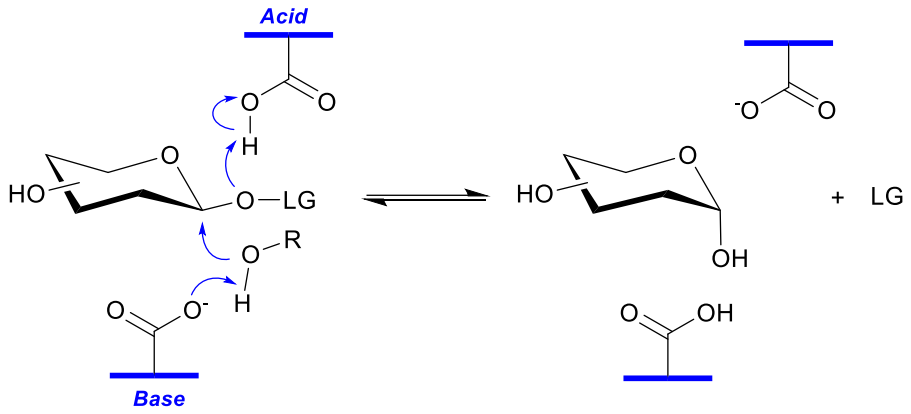


Fig. 2. Inverting GHs catalytic mechanism

The rGHs catalyze the reaction by a double-displacement mechanism that involves the formation of a covalent catalytic intermediate. Regarding the two catalytic carboxylates, one adopts nucleophilic character, while the other acts successively as an acid and then a base during catalysis (Fig. 3). In the first step (designated glycosylation), the acid/base catalysts as an acid, donating its proton to the interglycosidic oxygen of the targeted scissile glycosidic bond. Simultaneously, the nucleophile catalyst attacks the anomeric carbon causing inversion of the anomeric configuration and the formation of a covalent glycosyl-enzyme intermediate. This first step is accompanied by the departure from the glycoside substrate of the leaving group (LG). In the second step (designated deglycosylation), glycosyl-enzyme intermediate is deglycosylated through an attack by an acceptor molecule (R-OH) on the anomeric centre (Koshland 1953). The acceptor is concomitantly activated by the acid/base catalyst, which acts as a base. Once again, this step leads to inversion of the anomeric configuration, thus generating a product whose anomeric configuration is identical to that of the substrate (*i.e.* the net result is conservation of the substrate configuration). In this mechanism, the distance between the two catalytic residues is shorter, being around 5.5 Å (S. G. Withers 1999; G. Davies and Henrissat 1995).

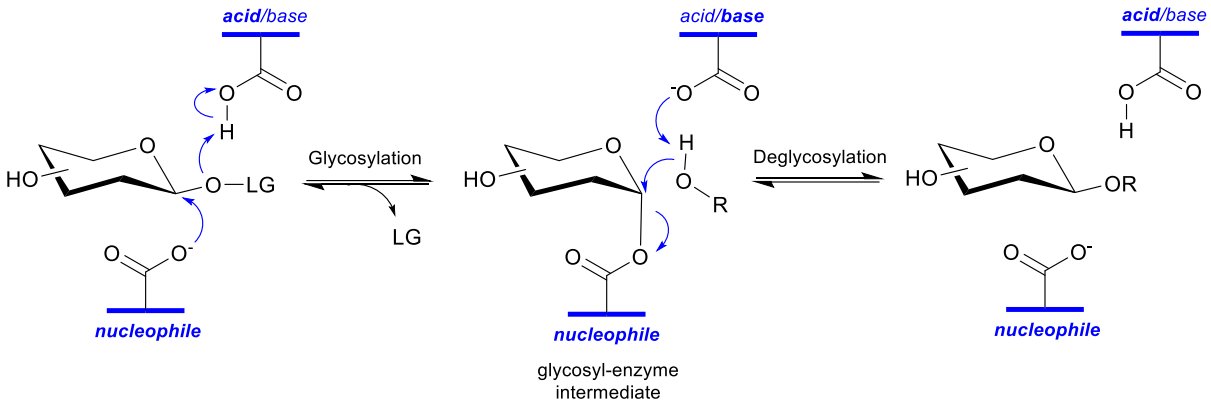


Fig. 3. Retaining GHs catalytic mechanism

In aqueous medium, hydrolysis (i.e. water molecule acts as the acceptor) is generally the principal outcome of rGH-catalyzed reactions. However, intrinsically all rGHs hold the potential to perform another reaction, known as transglycosylation, which occurs when a glycoside acts as acceptor (*i.e.* R = glycoside acceptor) to deglycosylate the glycosyl-enzyme intermediate. Generally, in aqueous medium, this alternative reaction competes with hydrolysis. However, it is only perceptible in the case of enzymes that display a sufficiently high transglycosylation (T): hydrolysis (H) ratio. In rarer cases, where the T/ H is very significantly shifted in favor of transglycosylation rGHs are designated using the alternative name, transglycosylases (TGs).

It is noteworthy that there is an exception to the classical Koshland description of the retaining mechanism. This is provided by rGHs that use substrates bearing an *N*-acetyl group at *C*-2 position and which do not possess a catalytic nucleophile. In this case, the substrate assists catalysis (*f* in Fig. 1), with a 2-acetamido group acting as an intramolecular nucleophile that is stabilized by a helper residue.

1.2.3.2. Nomenclature for sugar-binding subsites

In addition to the classification of GHs according to their catalytic mechanism, it is also possible to classify them according to whether they cleave internal glycosidic bonds (*i.e.* bonds linking glycosyl subunits that are themselves bound to other glycosyl subunits in a chain) or external ones (*i.e.* bonds that link a terminal glycosyl moiety to its neighbor). The former are called *endo*-acting GHs, while the latter are *exo*-acting GHs (Bissaro, Monsan, et al. 2015). Regarding *exo*-enzymes, these mostly remove terminal non-reducing glycosyl moieties, but rarer cases are those that remove reducing glycosyl subunits (Honda et al. 2008).

Considering that GHs in general, and especially *endo*-acting GHs, are active with a variety of carbohydrates, including oligomers and polymers, enzyme binding can involve interactions with several glycosyl moieties in a chain. Therefore, to facilitate the description of enzyme-substrate interactions and to describe reactions it is useful to use a standard nomenclature to describe active site interactions between a GH and its substrate. In the most widely used nomenclature, each monosaccharidyl moiety is identified by a number, which is mirrored by the same number to designate the enzyme subsite at which the moiety binds (Fig. 4). The scissile bond represents the zero point, such that glycosyl moieties that lie upstream (*i.e.* on the non-reducing side of the bond) are numbered using negative values starting from -1, while these that lie downstream (*i.e.* on the reducing side of the bond) are numbered with positive values starting from +1. Similarly, the subsites are numbered accordingly, with subsites -1 and +1 fixing the two glycosyl moieties that are linked by the scissile bond (Davies and Henrissat 1995; Davies, Wilson, and Henrissat 1997).

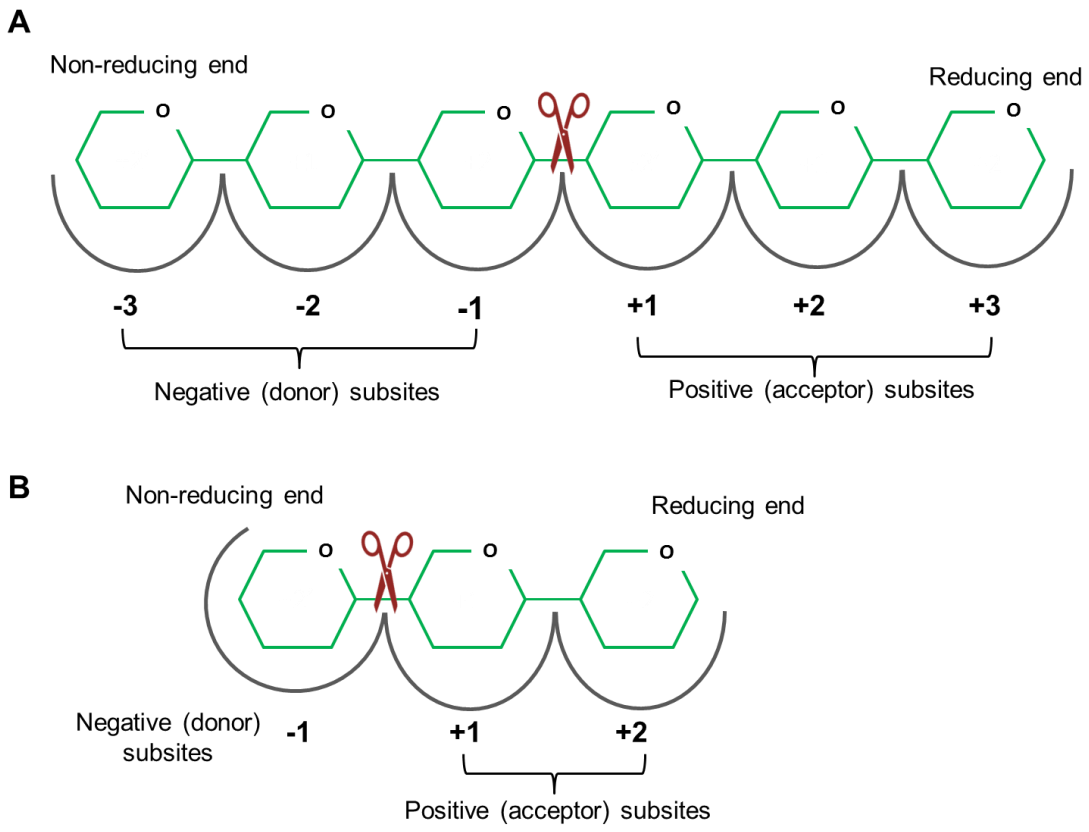


Fig. 4. Nomenclature for sugar–binding subsites in GHs. (**A**) *endo*-acting GHs cleave internal glycosidic bonds; (**B**) *exo*-GHs remove terminal glycosyl moieties, acting generally (but not exclusively) on non-reducing sugars.

1.2.4. Transglycosylase (TGs)

As mentioned in section 1.2.4, TGs are special rGHs that, while using the same mechanism as rGHs, mainly catalyze transglycosylation, even in dilute conditions and aqueous media (Sinnott 1990). Intriguingly, TGs are highly related to their hydrolytic GH counterparts, with any single TG being more related to the other members of its GH family than to TGs from other families (Bissaro, Monsan, et al. 2015). However, only relatively few naturally-occurring TGs have been described so far, these being mainly grouped in families rGH2, 13, 16, 31, 70, 77, 23, 102, 103 and 104 (Williams 2015). TGs have been the subject of extensive studies and have intrigued researchers for many years. Furthermore, the existence of naturally-occurring TGs provides inspiration for work aimed at engineering new TGs, starting from hydrolytic rGHs.

1.3. Furanosides in Nature and their biological importance

Furanoses are carbohydrates displaying a five-membered ring conformation, which when compared to the 6-membered pyranose ring, is a thermodynamically unfavorable conformation. Nevertheless, furanoses are widespread in Nature, the main example being D-(deoxy)ribofuranose, which forms an integral part of nucleic acids. Other furanoses are also widespread, with arabinofuranose (Ara*f*), galactofuranose (Gal*f*) and fructofuranose (Fru*f*) being frequently found in naturally occurring oligo- and polysaccharides (Richards and Lowary 2009; Lowary 2003). Significantly, with the exception of ribofuranose, furanoses are not present in mammalian glycoconjugates, but are widespread in the glycans produced by many bacterial pathogens and in plant cell wall polysaccharides. In bacteria, furanose sugars are often found in cell surface glycoconjugates, and are essential for the viability or virulence of these organisms (Poulin and Lowary 2010). As a result, the enzymes involved in the biosynthesis of bacterial furanosides are attractive targets as potential selective antimicrobial agents. However, before such chemotherapeutics can be designed, synthesized, and evaluated, more information about the activity and specificity of these enzymes is required (Peltier et al. 2008; Pedersen and Turco 2003; Houseknecht and Lowary 2001).

1.3.1. Arabinofuranosides

1.3.1.1. In plants

In Nature, both D- and L-Ara*f* are found in glycomotifs. However, L-Ara*f* is a plant glyco-structure sugar that is not found in animals or microorganisms (Kotake et al. 2016), being the second most abundant pentose after D-Xyl*p* in the plant polysaccharides (Seiboth and Metz 2011). In grasses, including rice (*Oryza sativa*) and wheat (*Triticum aestivum*), arabinoxylans (AXs) form the major L-Ara*f*-containing polysaccharide in cell walls. Arabinoxylans (AXs) are the second most abundant polysaccharide in the plant kingdom after cellulose. AXs are polymers found in primary and secondary plant cell walls that are constituted by a backbone of β-(1,4)-linked D-Xyl*p* units. Depending on the botanical and histochemical origin of AXs, individual backbone xylosyl moieties unit are either unsubstituted, or bear mono- or di-substitutions. Regarding these substitutions, they are often α-L-Ara*f* moieties that link to xylosyl moieties through (1,2) and/or (1,3)- linkage(s) (Fig. 5) (Pastell et al. 2009).

In addition to AXs, α-L-Ara*f* is also present in arabinans, which are composed of a backbone of (1,5)-linked α-L-Ara*f* that can be decorated at position *O*-3 and/or position *O*-2 by α-L-Ara*f* side-chain substitutions (Kawabata et al. 1995; Joseleau et al. 1977; Seiboth and Metz 2011; Classen, Baumann, and Utermoehlen 2019). Additionally, β-L-Ara*f* is also present in plant carbohydrate structures (e.g. arabinans, arabinogalactans) (Kotake et al. 2016; Voragen et al. 2009).

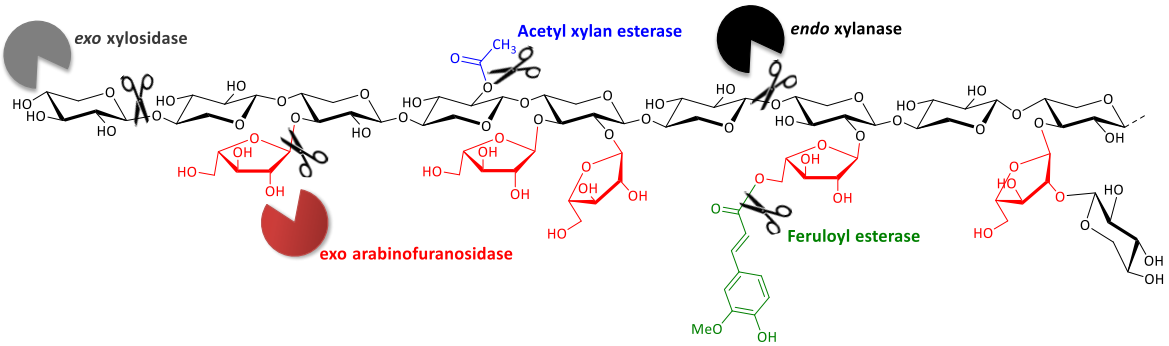


Fig. 5. Arabinoxylan structure and associated degrading enzymes

In order to describe the complexity of arabinoxylo-oligosaccharides (AXOS), an efficient nomenclature system was created (Fauré et al. 2009). This rich-name nomenclature is a one-letter code system allowing an accurate description of diverse oligosaccharides derived from

heteroxylans (e.g., X = unsubstituted D-Xyl*p*, A^2 or A^3 = D-Xyl*p* unit substituted by an α-L-Ara*f* moiety with (1,2) or (1,3) linkage, respectively; Fig. 6). AXOS with different degrees of polymerization (DPs 2-5) were characterized and their chemical shifts determined by NMR are listed in Table 1.

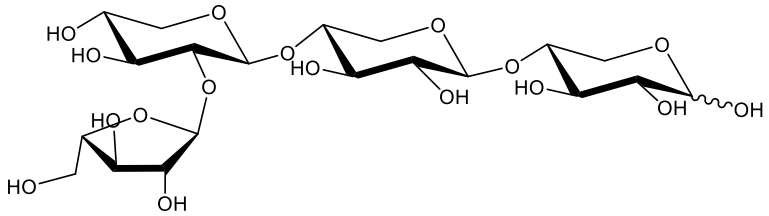


Fig. 6. Structure of A^2XX (α-L-Ara*f*-(1,2)-β-D-Xyl*p*-(1,4)-β-D-Xyl*p*-(1,4)-D-Xyl*p*): a representative of arabinoxylo-tetrasaccharides.

Table 1. ^{1}H-NMR characterization of AXOS

AXOS	*δ* (ppm) of anomeric proton from α-L-Ara*f* moieties			Reference
	23°C	**27**°C	**45°C**	
A^2X	5.28		5.20	(Pastell et al. 2008) (Zhao et al. 2020)
A^3X		5.33-5.34		(Gruppen et al. 1992; Viëtor et al. 1994)
$A^{2+3}X$		[5.24/5.25]		(Pastell et al. 2008)
A^2XX	5.26	5.29	5.20	(Pastell et al. 2008) (Viëtor et al. 1994; Ferré et al. 2000) (Zhao et al. 2020)
A^3XX		5.33	5.25	(Ferré et al. 2000) (Zhao et al. 2020)
$A^{2+3}XX$	[5.24/5.25]	[5.24/5.25]		(Pastell et al. 2008) (Gruppen et al. 1992; Viëtor et al. 1994)
XA^3X		5.40-5.40	5.32	(Gruppen et al. 1992; Viëtor et al. 1994) (Zhao et al. 2020)
A^2A^3X		5.29/5.43-5.42		(Viëtor et al. 1994)
A^3A^3X		5.33/5.40		(Gruppen et al. 1992)

A^{2+3} A^{3}X	[5.24/5.24]/5.43-5.42	(Viëtor et al. 1994)
$A^{3}A^{2+3}$X	5.33/[5.27/5.23]	(Gruppen et al. 1992)
XA^{2}XX	5.29	(Ferré et al. 2000)
XA^{3}XX	5.40	(Hoffmann et al. 1991) (Ferré et al. 2000)
XA^{2+3}XX	[5.27/5.23]	(Hoffmann et al. 1991; Viëtor et al. 1994)
X$A^{3}A^{3}$X	5.40/5.39	(Viëtor et al. 1994)
$A^{3}A^{2+3}$XX	5.33/[5.27/5.23]	(Gruppen et al. 1992)
$A^{2+3}A^{2+3}$XX	[5.25/5.25]/ [5.22/5.29]	(Gruppen et al. 1992)

The range of chemical shifts are shown for anomeric protons of α-L-**Ara*f*** units that are influenced by the anomery mixture of the reducing D-Xyl*p* unit. When a D-Xyl*p* unit is di-substituted, the substitution on *O*-2 position is given first followed by that one on *O*-3 position and values are shown in square brackets. This Table is updated from the summarized data of the thesis (Bissaro 2014).

1.3.1.2. In microorganisms

The D-Ara*f* enantiomer is present in bacteria. However, probably the most well-studied and notable occurrence of D-Ara*f* is in cell wall arabinogalactans (AG) and lipoarabinomannans (LAM) of mycobacteria. The arabinan portions of mycobacterial AG and LAM comprise (1,5)-linked chains of α-D-Ara*f* that are terminated with a branched hexasaccharide motif that contains two β-D-Ara*f* residues. (Lowary 2003; Ayers et al. 1998; Lee, Brennan, and Besra 1997). Besides, D-Ara*f* is also present in other pathogenic bacteria such as *Pseudomonas aeruginosa* (Kus et al. 2008; Harvey et al. 2018) and *Corynebacterium diphtheriae* (Wesener, Levengood, and Kiessling 2017).

1.3.2. D-galactofuranosides

Unlike Ara*f*, D-Gal*f* is the only enantiomer present in Nature. Both α- and β-D-Gal*f* residues are present in various lower organisms. However, β-D-Gal*f* linkages are most prevalent (Lowary 2003). D-Gal*f*-containing polysaccharides are widespread in naturally occurring glycoconjugates from bacteria, protozoa, fungi, plants and archaebacterial, many of which are pathogenic for humans, such as *Mycobacterium tuberculosis*, *Leishmania major* and

Trypanosoma cruzi (Peltier et al. 2008; Pedersen and Turco 2003; Houseknecht and Lowary 2001).

In many pathogenic bacteria, D-Gal*f* forms part of key cell surface glycoconjugates, including the lipopolysaccharide (LPS) *O*-antigens of *Escherichia coli* (Stevenson et al. 1994), the capsular polysaccharide (CPS) of *Campylobacter jejuni* (Hanniffy et al. 1999) and mycolyl arabinogalactan (mAG) complex and lipoarabinomannan (LAM) of *Mycobacterium tuberculosis*, which is the causative agent of tuberculosis, (Lee, Brennan, and Besra 1997). Notably, the mAG of mycobacteria cell wall is formed by a linear galactan chain of alternating (1,5)- and (1,6)-linked β-D-Gal*f* residues and branched with arabinan chain of (1,5)-linked α-D-Ara*f* units (Fig. 7) (Lowary 2003; Umesiri et al. 2010).

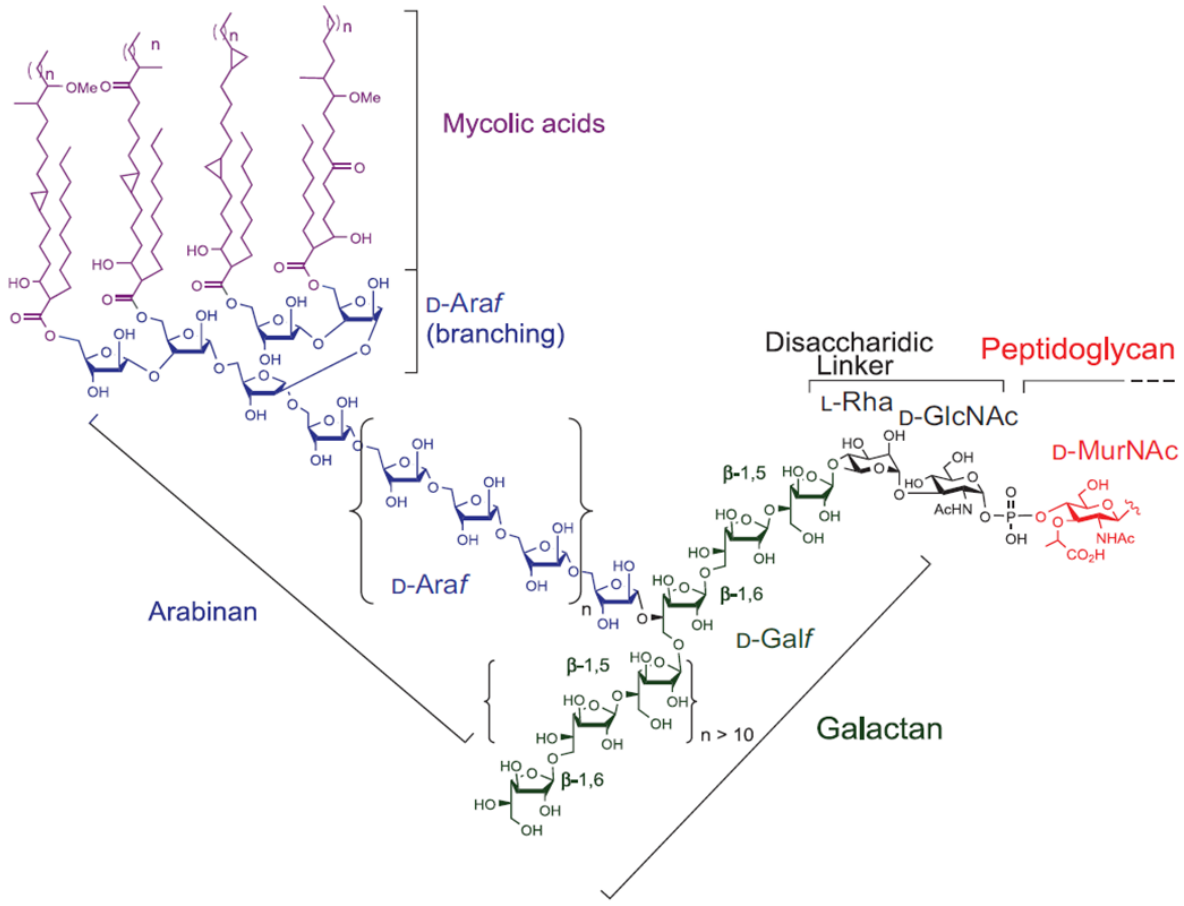


Fig. 7 Molecular structure of mAG complex in mycobacterial cell-wall (Eppe, Bkassiny, and Vincent 2015).

The protozoan *T. cruzi* (causative agent of Chagas' disease) possesses well-characterized β-D-Gal*f* conjugates. In the mucins of *T. cruzi*, two core motifs have been identified (Fig. 8), β-D-Gal*f*-(1,2)-β-D-Gal*f* and β-D-Gal*f*-(1,4)-α-D-Glc*p*NAc (Jones et al. 2004).

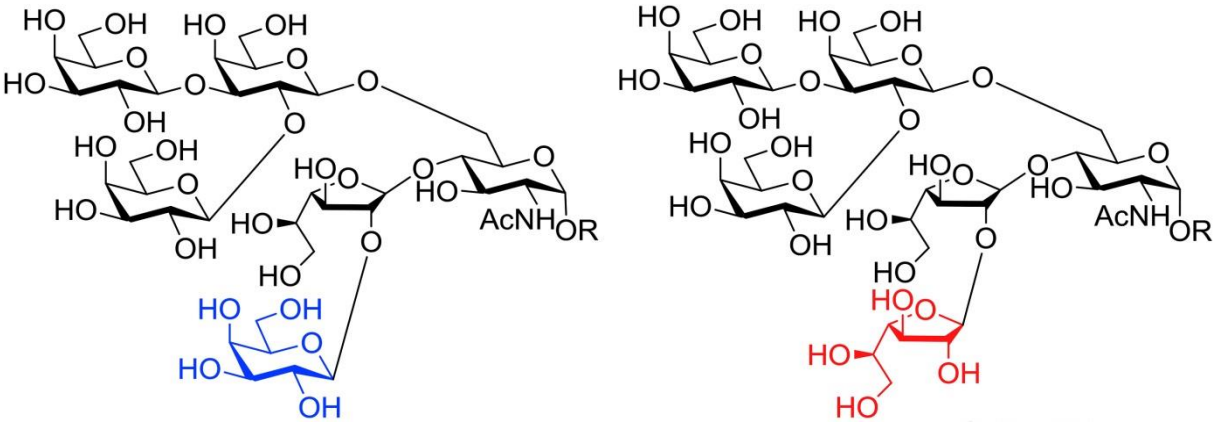


Fig. 8 Molecular structure of two types of β-D-Gal*f* glycoconjugates in mucin-like from *T. cruzi* (Kashiwagi et al. 2019).

D-Gal*f* is also part of the major antigen circulating in patients suffering from invasive aspergillosis, caused by the fungus *Aspergillus fumigatus* (Tefsen et al. 2012). At least four different molecules that include β-D-Gal*f* units in their structure are found in *A. fumigatus* (Marino, Rinflerch, and De Lederkremer 2017). A galactomannan secreted by *A. fumigatus* is also principally constituted by a core chain composed of (1-2)- and (1-6)-linked α-D-mannopyranosyl units with sidechains consisting in around 4-10 units of (1-5)-linked β-D-Gal*f* residues (Latge 2009). Other glycomotifs, such as β-(1,3)-linked D-Gal*f* moieties are present in oligosaccharidic structures of *Talaromyces* species (Prieto, Bernabé, and Leal 1995; Leitão et al. 2003).

1.3.3. Biological importance

1.3.3.1. Prebiotics

In principle, human cells do not produce enzymes to degrade AXOS. Therefore, these plant-derived oligosaccharides can be considered as prebiotics, since upon ingestion they are mostly digested in the colon, being the target of the microorganisms that form the intestinal microbiome. Several *in vitro* studies reveal that microorganisms from genera such as *Bifidobacteria* or *Lactobacilli* are able to degrade AXOS, thus boosting their growth (Grootaert, Verstraete, and Van de Wiele 2007; Mathew et al. 2018; Sanchez et al. 2009).

Moreover, a study of different *Bifidobacteria* strains showed that the prebiotic potential of AXOS is dependent on the exact bacterial strain under study and that this variability is related to the L-Ara*f* substitution pattern (Pastell et al. 2009). Besides their prebiotic effect, AX and AXOS play a vital role in human health by lowering cholesterol level and the risk of type II diabetes and obesity by reducing postprandial glucose level (Amrein et al. 2003; Möhlig et al. 2005).

1.3.3.2. Enzymatic substrates for the study of hemicellulases

The availability of well-defined AXOS is of crucial importance for the study of arabinoxylan-acting enzymes. This is particularly relevant for research purposes, but also for food and biofuels industry applications (McCleary et al. 2015; Thakur, Sharma, and Goyal 2019; Poria et al. 2020). Currently, the substrates available in the Megazyme arabinoxylan portal are particularly representative of commercially available compounds (https://www.megazyme.com/focus-areas/arabinoxylan-portal). Nevertheless, other examples of specific, arabinose-based substrates developed for enzyme screening purposes populate scientific literature (Borsenberger et al. 2012). Moreover, the availability of AXOS presenting different degrees of polymerization (DP) are particularly valuable as standards to monitor both hydrolytic and synthetic reactions for degrading or biosynthetic enzymes (Trincone 2015; Bissaro et al. 2014).

1.3.3.3. Therapeutic interests

In many organisms (e.g. mycobacteria), furanose components of glycans (e.g. arabinogalactan) have been demonstrated to be essential for cell viability, or to play a critical role in cell physiology (Pan et al. 2001; Pedersen and Turco 2003). As key components of cell wall glycoconjugates of pathogenic microorganisms, D-Gal*f*-bearing chains are regarded as antigenic epitopes that trigger immunogenic responses (Arruda, Colli, and Zingales 1989; Marino, Rinflerch, and De Lederkremer 2017). Considering that D-Gal*f* is absent in mammals, there has been a great deal of interest in developing inhibitors of enzymes involved in galactofuranoside biosynthesis, the aim being to develop medical applications strategies to fight bacterial infection (Poulin and Lowary 2010). Accordingly, using detailed knowledge of the structures of target glycomotifs, considerable work has been performed on the inhibition of arabinosyl- and galacto-furanosyl transferases (Houseknecht and Lowary 2002). Likewise, synthetic oligosaccharides have been prepared for use as precursors of artificial antigens that can be used for diagnostic purposes or as the basis of synthetic carbohydrate-based vaccines (Marino, Rinflerch, and De Lederkremer 2017).

1.4. Furanosides biosynthesis

In biological systems, furanose-containing carbohydrates are synthesized by arabino- and galactofuranosyltransferases, enzymes that use nucleotide glycosyl sugar donors (Rose et al. 2006; Belanova et al. 2008; Marino, Rinflerch, and De Lederkremer 2017). However, considering the difficulties related to the use of such enzymes for *in vitro* synthesis (see GTs in Section 1.2.1), so far conventional chemistry has mostly been used to synthesize furanosides. Unfortunately, this is also fraught with difficulties, because sugars in furanose configuration display less thermodynamic stability than their pyranose counterparts (Taha, Richards, and Lowary 2013). Some progress has nonetheless been made by applying different strategies developed for the preparation of partially protected derivatives of donor and using suitable acceptors for the construction of (1,2)-, (1,3)-, (1,5)-, and (1,6)-linkages (Marino and Baldoni 2014; Imamura and Lowary 2011). However, alternative combinatorial strategies that use both chemistry and enzymes have also been used to tackle the challenge of preparing furanoside-containing compounds (Rose et al. 2006).

1.4.1. Arabinofuranosyltransferases

So far, the vast majority of scientific reports have focused on mycobacterial arabinofuranosyltransferases (AfT) involved in the biosynthesis of cell wall LAM and mAG assemblies (Alderwick et al. 2006; Rose et al. 2006; Alderwick et al. 2018; Jankute et al. 2018), with less being known about plant AfTs (Kotake et al. 2016). Some examples of AfTs that are involved in synthesizing arabinofuranosides (mainly for plants) are listed in Table 2.

AfTs play a part in the assembly of mycobacterial cell wall. In mAG synthesis, AfT connects three α-D-Ara*f*-(1,5) units to the galactan in order to initiate the biosynthesis of the arabinan part of mAG (Belanova et al. 2008). A set of AfTs related to the AG biosynthesis of the mycobacterial cell wall core were recently reviewed by Alderwick et. al (Alderwick et al. 2018). Fig. 9 illustrates the intricate mechanism used by this pathogen to assemble arabinogalactan. To be brief, AftA adds D-Ara*f* residues to the galactan chain via an α-(1,5) linkage. AftC plays a role of branching the arabinan chain by adding an α-(1,3)-D-Ara*f* unit. AftD is in charge of elongating the arabinan chain via α-(1,5) linkage and AftB processes a "capping role" by adding a β-(1, 2)-Ara*f* unit at the terminal position (Alderwick et al. 2006, 2018).

Table 2. AfTs involved in the synthesis of L-Ara*f* structures in plants (Kotake et al. 2016)

Enzyme	GT family	Product	Polysaccharide	References
			Cell wall polysaccharide	(Kotake et al. 2016)
ARAD	47	α-L-Ara*f*-(1,5)-α-L-Ara*f*	Rhamnogalacturonan I (RG-I) arabinan	(Harholt et al. 2006) (Konishi et al. 2006)
XAT	61	α-L-Ara*f*-(1,3)-β-D-Xyl*p*	Arabinoxylan	(Anders et al. 2012)
XST	47	α-L-Ara*f*-(1,2)-α-D-Xyl*p*	Xyloglucan	(Schultink et al. 2013)
			Signaling peptide	
HPAT	95	β-L-Ara*f*-*O*-Hyp	CLV3 peptide	(Ogawa-Ohnishi, Matsushita, and Matsubayashi 2013)
RRA3a	77	β-L-Ara*f*-(1,2)-β-L-Ara*f*	CLV3 peptide	(Xu et al. 2015)

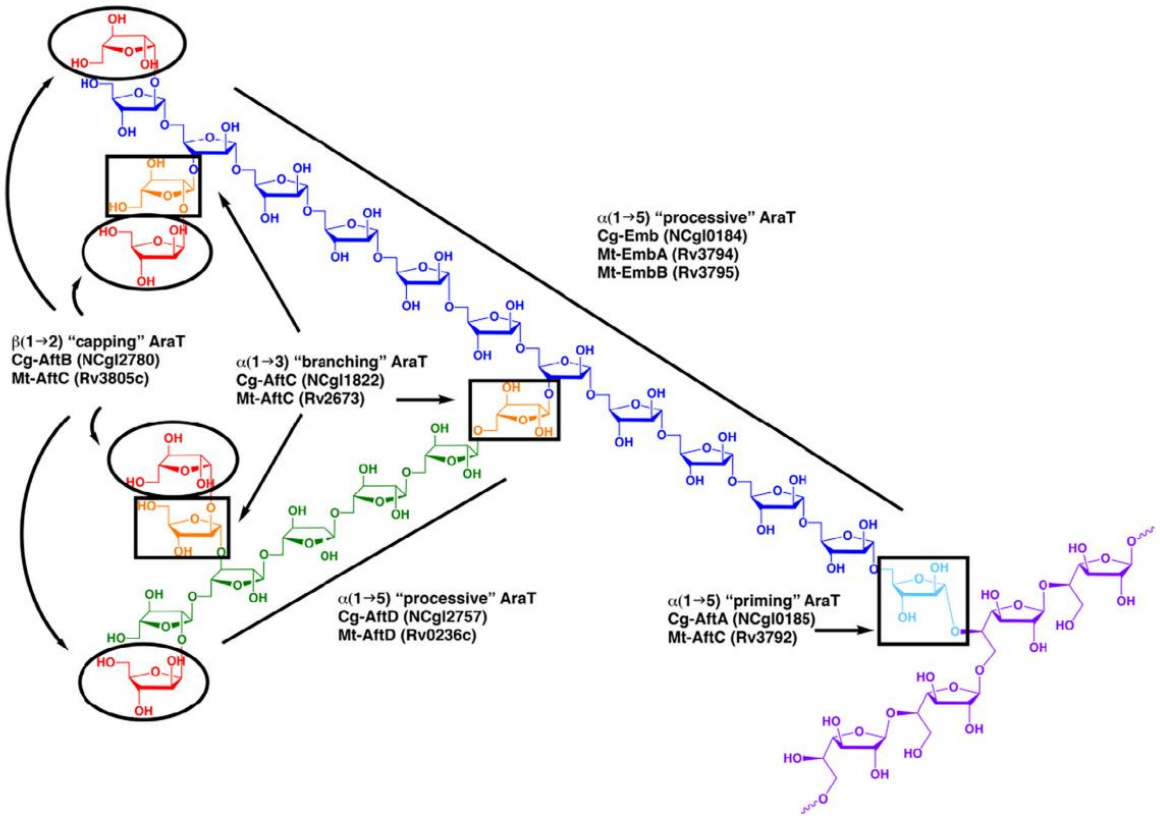


Fig. 9. Revised pathway of arabinan biosynthesis illustrating the individual roles of Emb, AftA, AftB, AftC and AftD arabinofuranosyltransferases (Alderwick et al. 2018).

1.4.2. Galactofuranosyltransferase

For synthetic purposes, several galactofuranosyltransferases have been cloned. These include those from *E. coli* (Wing et al. 2006), *Aspergillus nidulans* (Komachi et al. 2013) and *Trypanosoma rangelis* (Stoco et al. 2012). However, the two most studied galactofuranosyltransferases remain those of *M. tuberculosis*, GlfT1 and GlfT2 (Berg et al. 2007; Tam and Lowary 2009). GlfT1 and GlfT2 are sequentially involved in the biosynthesis of galactan using UDP-Gal*f* as donor, this being synthesized by UDP-Gal*p* mutase (UGM) starting from from UDP-Gal*p* (Fig. 10). GlfT1 initiates galactan biosynthesis by catalyzing the successive transfer of (1,4)- and (1,5)-β-D-Gal*f* donors onto the L-rhamnopyranosyl unit of the decaprenyl diphosphate α-L-Rha*p*-(1,3)-α-L-Glc*p*NAc mAG linker as acceptor. Galactan is then polymerized by GlfT2 that performs the synthesis of (1,6)- and (1,5)-β-D-Gal*f* alternating transfers (Belanova et al. 2008). Recently, a similar biosynthesis pathway for mAG has been described in *Corynebacterium diphtheria*, which also includes two GlfTs (GlfT1 and GlfT2) that are comparable to those from *M. tuberculosis* (Wesener, Levengood, and Kiessling 2017).

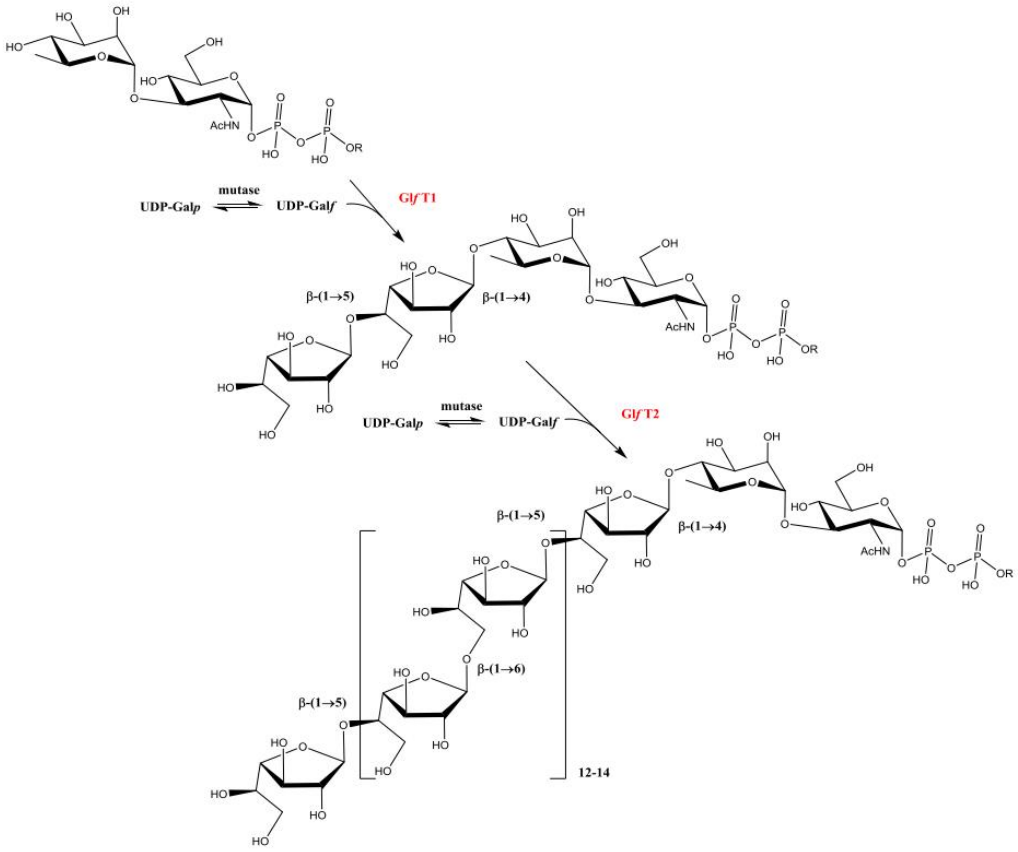


Fig. 10 Mycobacterial galactan biosynthesis catalyzed by GlfT1 and GlfT2. (R = decaprenyl) (Belanova et al. 2008).

Using GlfTs, both Lowary et al. (Vembaiyan et al. 2011; Poulin, Zhou, and Lowary 2012; Szczepina et al. 2010) and Kiessling et al. (Splain and Kiessling 2010; Levengood, Splain, and Kiessling 2011) have reported the preparation of mycobacterial galactans, the aim of both teams being to elucidate the mechanism of action of these enzymes and evaluate potential inhibitors.

1.4.3. β-D-galactofuranosidase

An alternative strategy to access furanose-bearing glycoconjugates is the use of transfuranosylases. As mentioned earlier (section 1.2.1), transfuranosylases are enzymes that possess the ability to synthesize furanose-containing compounds, using furanosyl donors. Unlike GTs, transfuranosylases are actually rGHs and thus and do not require the use of costly nucleotide sugars as donors. Ideally for the synthesis of galactofuranosides it would be convenient to use a D-galactofuranosidase (EC 3.2.1.146). However, up to now, only few such enzymes have been reported (Matsunaga, Higuchi, Mori, Yairo, et al. 2017). Back in 1977, the first *exo*-β-D-galactofuranosidase was purified from the culture medium of *Penicillium fellutanum* (Rietschel-Berst et al. 1977), and later several *exo*- and *endo*-β-galactofuranosidases were purified from the culture supernatants and cell lysates of fungi (Van Bruggen-Van der Lugt et al. 1992; Matsunaga, Higuchi, Mori, Tashiro, et al. 2017; Cousin et al. 1989), bacteria (Ramli et al. 1995) and protozoa (Miletti et al. 2003). The most recent reports of β-D-galactofuranosidases belonging to GH5 and GH43 subfamilies appeared in 2019 (Helbert et al. 2019). In the majority of cases, these studies reported D-galactofuranosidase-like activity, but did not unequivocally demonstrate that the enzyme responsible was a *bona fide* β-D-galactofuranosidase. Recently, a β-D-galactofuranosidase encoding gene from *Streptomyces* species has been described (Matsunaga et al. 2015; Matsunaga, Higuchi, Mori, Yairo, et al. 2017) and the most recent study by Helbert *et al.* does appear to confirm that such enzymes exist and are distinct from other D-galactofuranose-acting enzymes, such as arabinofuranosidases (Remond et al. 2005; Chlubnova et al. 2012). Despite this excellent news, the lack of knowledge about β-D-galactofuranosidases constitutes a bottleneck for the application of such enzymes in chemosynthetic strategies aimed at the preparation of D-Gal*f*-containing oligosaccharides. Therefore, for many years α-L-arabinofuranosidase (Abfs), enzymes that are quite abundant in Nature, have been targeted for this purpose.

1.4.4. α-L-arabinofuranosidase

As mentioned above, work on some Abfs has revealed catalytic promiscuity that allows certain enzymes to act on both L-Ara*f* and its *C*-5-hydroxymethyled analog, D-Gal*f* (Remond et al. 2005; Chlubnová et al. 2010, 2014). Accordingly, using GH51 α-L-arabinofuranosidases from *Ruminiclostridium thermocellum* (*Ct*Abf) or *Thermobacillus xylanilyticus* (*Tx*Abf) and β-D-Gal*f*O*p*NP, chemo-enzymatic syntheses have procured disaccharides such as *p*-nitrophenyl β-D-Gal*f*-(1,2)-β-D-Gal*f* and β-D-Gal*f*-(1,6)-β-D-Gal*f*, as well as other D-galactofuranosides (DP 2 to 5，Table 5). The acceptor specificities of these enzymes have also been investigated, showing that D-Gal*f* can be transferred onto a variety of pyranoside acceptors and ethanol (Pavic et al. 2019). Hence, although the aforementioned Abfs are intrinsically hydrolytic GHs, they are also clearly ideal candidates for the development of potent synthetic enzymes for the preparation of L-Ara*f* and D-Gal*f*-containing saccharides and conjugates.

2. Towards transglycosylation in rGHs

The synthesis of glycosidic bonds by GHs (i.e. transglycosylation) has been studied for over 60 years (Edelman 2006). Unfortunately, application of GTs, the most common class of carbohydrate-active biosynthetic enzymes, has so far been hampered by the challenges discussed in section 1.2.1 (Field 2011; Chang et al. 2011; Nidetzky, Gutmann, and Zhong 2018). The development of alternative strategies employing other classes of enzymes, particularly non-Leloir transglycosylases (TGs) has thus been an area of considerable research, particularly because these are readily expressed as recombinant enzymes and use relatively inexpensive often commercially available donor substrates, such as *p*NP-derivatives.

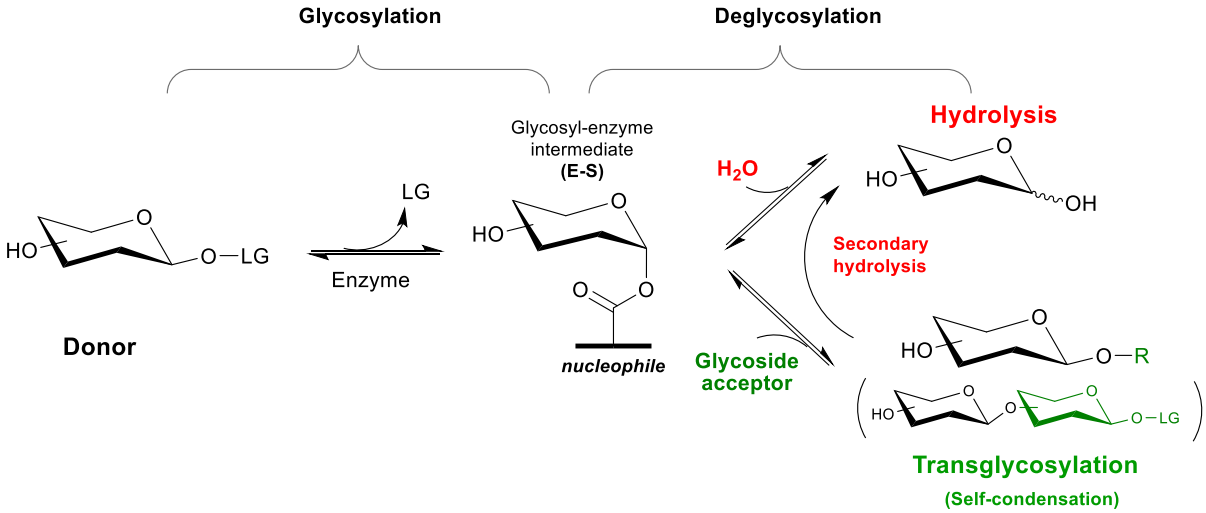


Fig. 11. Reactions catalyzed by rGHs regarding different incoming acceptors (water or carbohydrate) that lead to either hydrolysis or transglycosylation, including self-condensation.

Importantly, as described in section 1.2.3 and 1.2.4, TGs are directly related to rGHs, and are in fact a subclass of rGHs that preferentially perform transglycosylation. Regarding rGHs, these catalyze the hydrolysis and synthesis of glycosidic bonds using a double displacement mechanism, performing transglycosylation when the deglycosylation step involves transfer of the glycosyl sugar donor onto a glycoside acceptor. Similarly, self-condensation occurs when transglycosylation involves a single sugar species that acts as both the donor and acceptor (Fig. 11). In aqueous media, mainly for thermodynamic reasons (*i.e.* the concentration of water is >55 M) rGHs favor water-mediated deglycosylation that leads to hydrolysis.

However, transglycosylation always remains a plausible reaction outcome, the extent to which this is occurs being highly dependent on the exact nature of the rGH under study. Any given rGH is thus characterized by a transglycosylation/hydrolysis (T/H) partition the value of which is constant in defined reaction conditions. The fundamental understanding of how this partition is established, particularly the identification of the amino residues that determine T/H, is thus of considerable value if one wishes to design new TGs, starting from hydrolytic rGHs (Bissaro, Monsan, et al. 2015).

2.1. Glycosynthases

An important development in the use of GHs as synthetic tools is the glycosynthase concept. This is based on the creation of a mutant rGH whose nucleophile is replaced by a catalytically inert residue (Mackenzie et al. 1998; Malet and Planas 1998; Moracci et al. 1998). Glycosynthases are thus catalytically impotent, being unable to perform hydrolysis or transglycosylation in the presence of most standard substrates. However, when glycosynthases are fed highly activated donors (usually glycosyl fluorides) displaying the opposite anomeric configuration to that of the innate substrate configuration, synthesis activity is restored. The fluoride moiety fits into the cavity left by the ablated nucleophile, thus allowing the glycosyl fluoride donor to mimic the reactive glycosyl-enzyme intermediate. The reactivity of the glycosyl fluoride is sufficient to allow departure of the fluoride group (HF is formed) and transfer of the glycosyl moiety onto a suitable acceptor, all of this occurring in a one-step inverting-like mechanism (Fig. 2) (Danby and Withers 2016). Since glycosynthases are catalytically impotent, once the oligosaccharide is formed it cannot be hydrolyzed by the enzyme (*i.e*. secondary hydrolysis is impossible). This is a practical feature that allows product to accumulate. To date, the glycosynthase strategy has been applied to over 16 GH families (Aragunde et al. 2014), including inverting GHs and has allowed the efficient production of a variety of oligo- and polysaccharides (Honda and Kitaoka 2006).

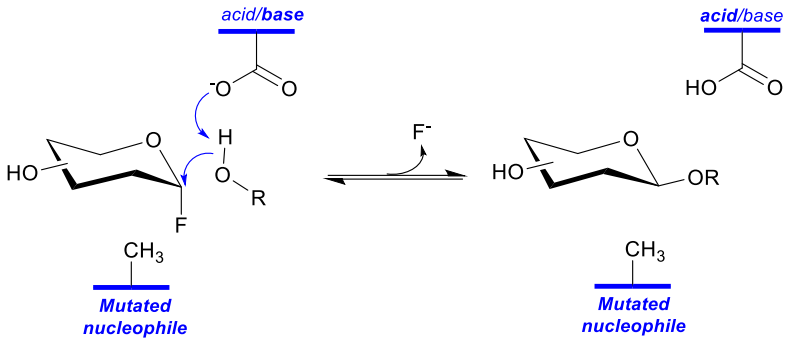


Fig. 12. Mechanism catalyzed by glycosynthase

It is noteworthy that glycosynthases have also been designed using rGHs that operate via a substrate-assisted mechanism. When converted into glycosynthases, these enzymes require the deployment of oxazoline-activated donors. In this way (Fig. 13), they are able to perform transglycosylation (Wada et al. 2008).

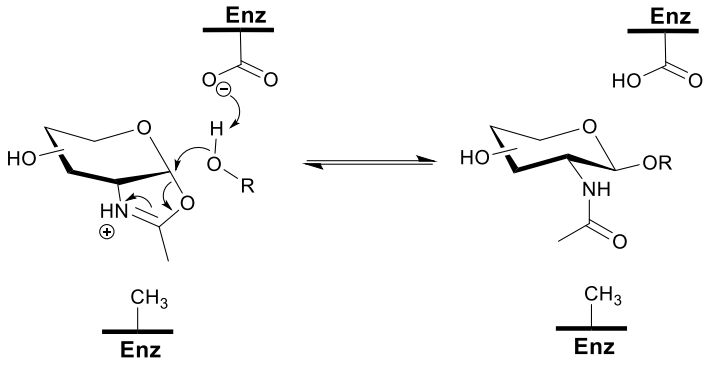


Fig. 13. Mutated enzyme catalyze transglycosylation using oxazoline-activated donors.

As mentioned above, since the introduction of the concept glycosynthases have been regarded as an excellent way to adapt enzymes for glycosynthesis. However, two caveats are associated with glycosynthases. First, glycosynthases are intrinsically inactive, thus their ability to turnover reactions in the presence of activated substrates remains low. This limit has been partially alleviated by performing directed evolution on nucleophile-less GH mutants (Mayer et al. 2000). Glycosynthases must also be fed highly activated glycosyl sugar donor, such as glycosyl fluorides. While the obtention of such substrates is feasible in the case of hexopyranoses, such as glucose, this is infinitely more difficult in the case of thermodynamically less stable pentofuranoses. Indeed, to date the synthesis of such

compounds has never been reported. Therefore, the glycosynthase strategy has so far never been applied to furanosyl hydrolases.

2.2. Naturally-occurring transglycosylases

As described in Part I, only a few rGHs have so far been classified as TGs in the CAZy database. However, known TGs are scattered across several GH families, belong to different clans and display diverse structural folds. This implies that the intrinsic determinants of transglycosidase activity are not exclusive to any one family, nor associated with a particular molecular architecture. Notable examples of TGs include xyloglucan *endo*-transglycosylases (XETs), glucansucrases (GS), cyclodextrin glucanotransferases (CGTases) and trans-sialidases (trS) (Williams 2015). A more complete list of naturally occurring TGs present in the CAZy database is shown in Table 3. XETs are involved in the remodeling of the cell walls of higher plants and yeasts. They are mainly grouped in family GH16 and display a typical β-jellyroll fold that is shared with their hydrolytic counterparts xyloglucan *endo*-hydrolases (XEHs, Saura-Valls et al. 2008; Baumann et al. 2007). Sucrases are *exo*-enzymes that include glucose-specific sucrases (i.e. glucansucrases) and fructose-specific sucrases (FS). GS are classified in both GH13 and GH70 families, and amylosucrases (AS) in GH13 are designated as GS, being characterized by a $(\beta/\alpha)_8$-barrel architecture (Stam et al. 2006). Regarding FS (e.g. levansucrases and inulosucrases), they are classified in GH68 and GH32 (clan GH-J and 5-fold β-propeller structure). In GH68, FS members display more hydrolytic activity than transglycosylation. In GH32, sucrases or invertases that are able to transfer the D-fructosyl moiety onto a sucrose acceptor are designated as fructosyltransferases (FTs, Lammens et al. 2009). In GH13, CGTases and their hydrolytic counterparts α-amylases are involved in starch depolymerization and modification. All of these enzymes display a $(\beta/\alpha)_8$ TIM-barrel fold (Moulis, André, and Remaud-Simeon 2016). Sialidases (SAs) and *tr*S, that catalyze either hydrolysis or synthesis of sialyl-glycoconjugates, are members of family GH33 and belong to the clan GH-E (6-fold β-propeller). trSs exhibit both activities (hydrolysis and transglycosylation) (Lipničanová et al. 2020).

Table 3. Examples of TGs classified in CAZy database

Clan	GH family	Enzyme	Reference
A	1	Rice OsBGlu31 (*exo*)	(Luang et al. 2013)
		AA5GT, AA7GT	(Matsuba et al. 2010)
	5	Mannan transglycosylase	(Schröder et al. 2004)
	10	Xylan endotransglycosylases	(Aspeborg et al. 2005)
B	16	XET Crh1, Crh2	(Baumann et al. 2007) (Blanco et al. 2015)
C	12	XET	(R. I. Nielsen 2002)
D	31	α-Glucosyltransferase	(Larsbrink et al. 2012)
E	33	Trans-sialidases	(Ribeirão et al. 1997; Lipničanová et al. 2020)
H	13	CGTases	(Kelly et al. 2008; Leemhuis, Dijkstra, and Dijkhuizen 2002)
		AS	(Moulis, André, and Remaud-Simeon 2016; Potocki De Montalk et al. 2000)
	70	GS	(Monsan, Remaud-Siméon, and André 2010; Moulis, André, and Remaud-Simeon 2016; van Hijum et al. 2006)
	77	Amylomaltase	(van der Maarel and Leemhuis 2013)
J	32	FTs	(Van den Ende et al. 2009)
	68	FS (levansucrases, inulosucrases)	(Olvera et al. 2012; Lammens et al. 2009)

Overall, each TG is generally related to homologous hydrolytic GHs that are classified within the same GH family. Moreover, TGs are much more closely related to the hydrolytic GHs within their family than with other TGs belonging to other GH families. This observation clearly indicates that the specific properties of TGs are the result of very subtle molecular level modifications that differentiate a TG from a GH.

2.3. Lessons learned from Nature and from engineered TGs

One way to explore differences between TGs and their homologous hydrolytic counterparts is to simply perform sequence alignment analysis. In this way, it is possible to detect putative determinants of hydrolysis or transglycosylation that can be probed using site-directed mutagenesis (Baumann et al. 2007; Frutuoso and Marana 2012; Madhuprakash et al. 2013; Florindo et al. 2018). Alternatively, advanced protein engineering methods can be used to explore the enhancement of TG activity. So far, the use of random or semi-random combinatorial approaches has procured several impressive examples of engineered TGs (Feng et al. 2005; Sun et al. 2016; Lundemo, Adlercreutz, and Karlsson 2013). Together these approaches provide several clues regarding the differences between GHs and TGs, even though our knowledge is still incomplete. One key finding from these studies is that to engineer a TG it is important to be able to detect transglycosylation in the parental enzyme. This is because the complete absence of detectable transglycosylation activity would suggest that the parental enzyme is highly evolved and exquisitely specialized for hydrolysis. (Demetrius 1998). Therefore, to reverse this situation, extensive engineering would be necessary to introduce transglycosylation. In this regard, it has been suggested that ancestral enzymes that form the evolutionary basis for modern day enzymes may have been less catalytically efficient, performing both hydrolysis and transglycosylation. Similarly, nowaday hydrolases are supposed to be the result of extensive evolution. One piece of evidence to support this postulate is the observation that TGs are generally rather 'lazy' enzymes that display low catalytic efficiency (Baumann et al. 2007; Luang et al. 2013).

A certain number of studies have been performed to elucidate the differences between TGs and their hydrolytic counterparts, working with naturally-occurring and engineered TG/GH pairs. These studies have used a series of techniques including biochemical characterization, structural studies and *in silico* simulations (molecular dynamic and quantum mechanics/molecular mechanics, QM/MM). Hereafter, several findings from these studies are summarized.

2.3.1. Kinetic analysis of TG/GH pairs

As mentioned above, TGs usually display low catalytic efficiency (k_{cat}/K_M), meaning that they are overall less efficient catalysts than their hydrolytic counterparts (Luang et al. 2013; Bissaro, Monsan, et al. 2015; Potocki De Montalk et al. 2000). This is also the case of engineered TGs that display drastically impaired catalytic efficiency compared to that of the

corresponding wild-type GHs (Teze et al. 2015; Bissaro, Durand, et al. 2015). Work on FT illustrates this (Schroeven et al. 2008). Comparison of the kinetic parameters of FT and its hydrolytic counterpart, vacuolar invertases (VI), indicated that VI is a more efficient catalyst (Table 4) Indeed, FT displays an extremely high K_M value (>500 mM) coupled with a low or moderate turnover constant k_{cat}. Therefore, the catalytic efficiency of FT is significantly impaired compared to VI. Nevertheless, FT displays higher transglycosylation activity than that of VI. Another example of a TG/GH pair is that of AS from *Neisseria polysaccharea* (*Np*AS) and its hydrolytic counterpart, sucrose hydrolase from *Xanthomonas axonopodis pv. glycines* (*Xag*SUH) (Potocki De Montalk et al. 2000; Kim et al. 2008). *Xag*SUH is unable to catalyze transglucosylation and displays a 120-fold higher k_{cat} value on sucrose than that of *Np*AS. A similar trend is observed when TGs are created in the laboratory, starting from rGHs. Generally, the resultant mutants display impaired catalytic efficiency coupled to significantly increased transglycosylation ability (Bissaro, Durand, et al. 2015; Teze et al. 2014).

Table 4. Comparison of catalytic constants for TG/GH pairs

Family	Enzyme	Donor	Acceptor	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($s^{-1}.mM^{-1}$)	Synthesis yield (%)
Naturally-occuring TG/GH pairs							
32	Wheat FT[a]	Sucrose		551	0.78	-	70-80
	Wheat VI[a]			15	608	-	2-5
13	*Np*AS[b]			1.9	0.55	3.45	>70
	*Xag*SUH[c]			2.24	66.5	0.03	-
Engineered TG/GH pairs							
51	*Tx*Abf[d]	α-L-Ara*f*O*p*NP	Xylotriose	0.25	139	556	10
	*Tx*Abf R69H[d]			0.09	2.29	25.41	31
	*Tx*Abf R69H-L352M[d]			1.23	0.97	0.79	27
	*Tx*Abf R69H-N216W-L352M2[d]			0.48	0.58	1.21	80
1	*Tt*β-gly[e]	β-D-Fuc*p*O*p*NP	β-D-Glc*p*N(Me)OBn	-	-	28	36
	*Tt*β-gly Y284F[e]			-	-	3.83	90
	*Tt*β-gly N282T[e]			-	-	0.97	86

References are indicated as follows: [a] (Schroeven et al. 2008); [b] (Potocki De Montalk et al. 2000); [c] (Kim et al. 2008); [d] (Bissaro, Durand, et al. 2015); [e] (Teze et al. 2014).

2.3.2. Exploring the hydrogen bond network in the TG/GH active sites

FT and its hydrolytic counterpart VI provide insight into the Hydrogen bond network differences that govern the T/H equilibria. A hydrogen bond network of VI is displayed in Fig. 14. The sucrose binding box (WMNDPNG) is conserved among different VIs (Lammens et al. 2009; Van den Ende et al. 2009). Interestingly, W19 and N21 (4EQV numbering), which form part of the sucrose binding box and are H-bonded to the nucleophile, are quite often substituted in FTs, belonging to the same GH32 subgroup, by Tyr and Ser respectively. In this case, the H-bond network is disrupted. When these substitutions (W23Y and N25S) were introduced into VI from *Triticum astivum*, transglucosylation was increased 16-fold compared to wild-type VI (Schroeven et al. 2008). Similar conclusions were also found for VI from yeast (Lafraya et al. 2011).

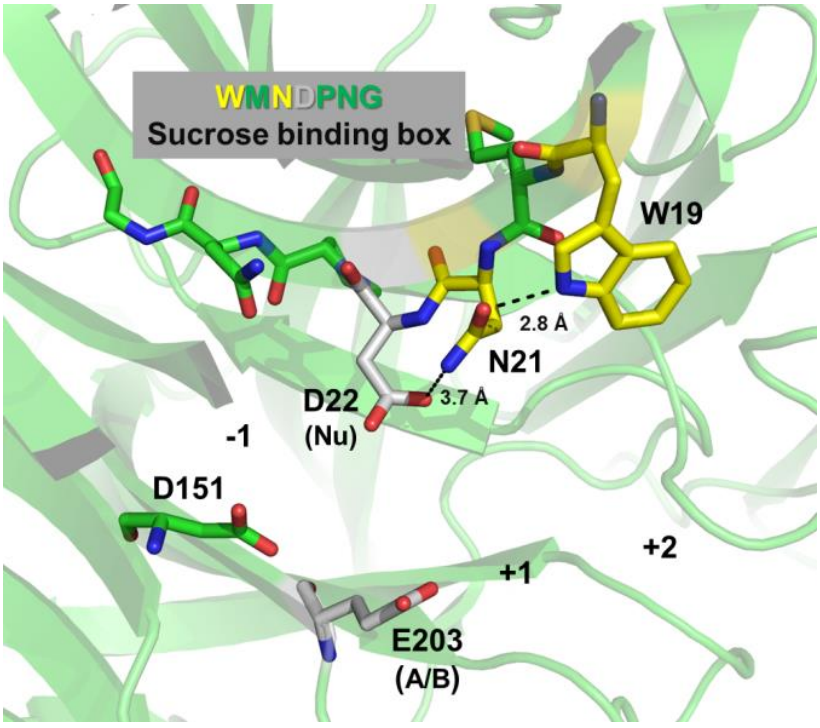


Fig. 14. Active site of the vacuolar invertase (VI) from *Saccharomyces* (PDB: 4EQV, homologous to VI from *Triticum aestivum*) (Sainz-Polo et al. 2013). D151 plays a role in TS-stabilizing (Lammens et al. 2009; Van den Ende et al. 2009).

2.3.3. Enzymes/substrates interactions and active site flexibility

Comparing the hydrogen binding network between XETs and XEHs (TG/GH pair) reveals that XEHs possess a greater number of hydrogen bonds in the negative subsites. However, inversely the number of hydrogen bonds formed between the XETs and the acceptor substrate is greater than that in XEHs (Mark et al. 2009). This suggests that subsite -1 binding in hydrolytic enzymes is obviousely important, perhaps to create a very specific transition state configuration that favours water-mediated deglycosylation. On the other hand, it appears logical that TGs display a greater ability to bind glycoside acceptors.

Active site flexibility (i.e., plasticity and adaptation) has been examined in order to explain the different T/H partitions observed in TG/rGH pairs. The study of the TcTS/TrSA (*T. cruzi* trans-sialidase/*Trypanosoma rangeli* sialidase) pair has provided evidence that trans-sialidase displays greater active site flexibility than sialidase (Paris et al. 2005; Amaya et al. 2003; Demir and Roitberg 2009). The nucleophile side-chain displays two distinct orientations in crystal structures of TcTS, while a rigid topology in this region was found in TrSA, suggesting that nucleophile flexibility might be important for the transglycosylation reaction (Paris et al. 2005). Similar conclusion was also reported that increased flexibility in the -1 subsite increased the transglycosylation activity of the GH13 *Np*AS (Champion et al. 2012).

In the early time it was also reported that the sidechain of the acid/base residue is much more flexible in the case of mutants that display enhanced transglycosylation in GH1 transglycosidases. The authors propose that, in addition to the extraction of a proton from the catalytic water molecule during the deglycosylation step, the acid/base residue retains the catalytic water molecule. Presumably, higher flexibility will reduce its ability to properly fulfill this role (Zechel and Withers 2000; Davies and Henrissat 1995) In thermodynamic terms, this implies that flexibility at the acid/base position adds entropic cost to the free energy barrier of the reaction. Consequently, hydrolysis is slowed down and the probability of glycosylation by glycoside acceptors is increased, thus favoring the transglycosylation (David et al. 2019b).

Loop flexibility and enzyme dynamics are likely to be the determinants of altered substrate recognition that in turn will determine the ability of the enzyme to create a catalytically productive transition state. In this regard, a dynamics study performed on an amylosucrase suggested that molecular plasticity in subsite +1 forms the basis for acceptor promiscuity and improved transglucosylation (Daudé et al. 2013). Moreover, a structural study on

transglycosylating *endo*-(1,4)-β-glucanases in the GH5_5 subfamily revealed that increased binding affinity and flexibility of the +1 and +2 subsites appeared to improve sugar recognition and subsequently favor transglycosylation (Dutoit et al. 2019).

Nevertheless, an opposite conclusion was found with cyclic tetrasaccharide cycloalternan (CA) in GH31. This suggested that by rigidly retaining a deeper conformation into the active site, the α4-α5 loop facilitates glycoside acceptor binding and sterically disfavors hydrolysis, thereby promoting transglycosylation (Light et al. 2017).

2.3.4. Transition states in T/H partition

As mentioned in section 1.2.3.1, rGHs operate via a two-step mechanism involving glycosylation and deglycosylation (of the enzyme). Along the reaction pathway, two transition states (oxocarbenium-like TS) occur, denoted TS1 and TS2 respectively (Fig. 15, Sinnott 1990; Davies et al. 2003). When the reaction involves a *p*NP-bearing donor (*i.e.* a good leaving group), the ratio k_{cat}/K_M and the rate constant k_{cat} reflect the glycosylation and the rate limiting deglycosylation steps, respectively. Therefore, considering the cumulative knowledge related to TGs, it is possible to postulate that TGs display low catalytic efficiency, thus a low k_{cat}/K_M value, combined with a low k_{cat} value, which reflects slow deglycosylation of the glycosyl-enzyme intermediate by water. Overall, this translates into higher TS energy barriers for both glycosylation and deglycosylation. Therefore, to investigate the properties of TS and the extent of bond breaking at each individual step, QM/MM approaches have been employed to study sialidases (SA) and trans-sialidases (trS). The assessment of the free energy profiles for the conversion of the Michaelis complex to the covalent glycosyl-enzyme intermediate (*i.e.* the glycosylation step, TS1) revealed that hydrolytic SA is less energy-demanding than trS. Conversely, the trans-sialylation reaction (i.e. deglycosylation step, TS2) appeared more favorable for trS (Pierdominici-Sottile, Horenstein, and Roitberg 2011; Pierdominici-Sottile, Palma, and Roitberg 2014). Likewise, the study of GH3 β-glucosidases has revealed that the conformational itineraries in TS2 are different for hydrolysis and transglycosylation. These differences are completed by the observation that the TS is more positively charged in the case of transglycosylation (Geronimo, Payne, and Sandgren 2018b).

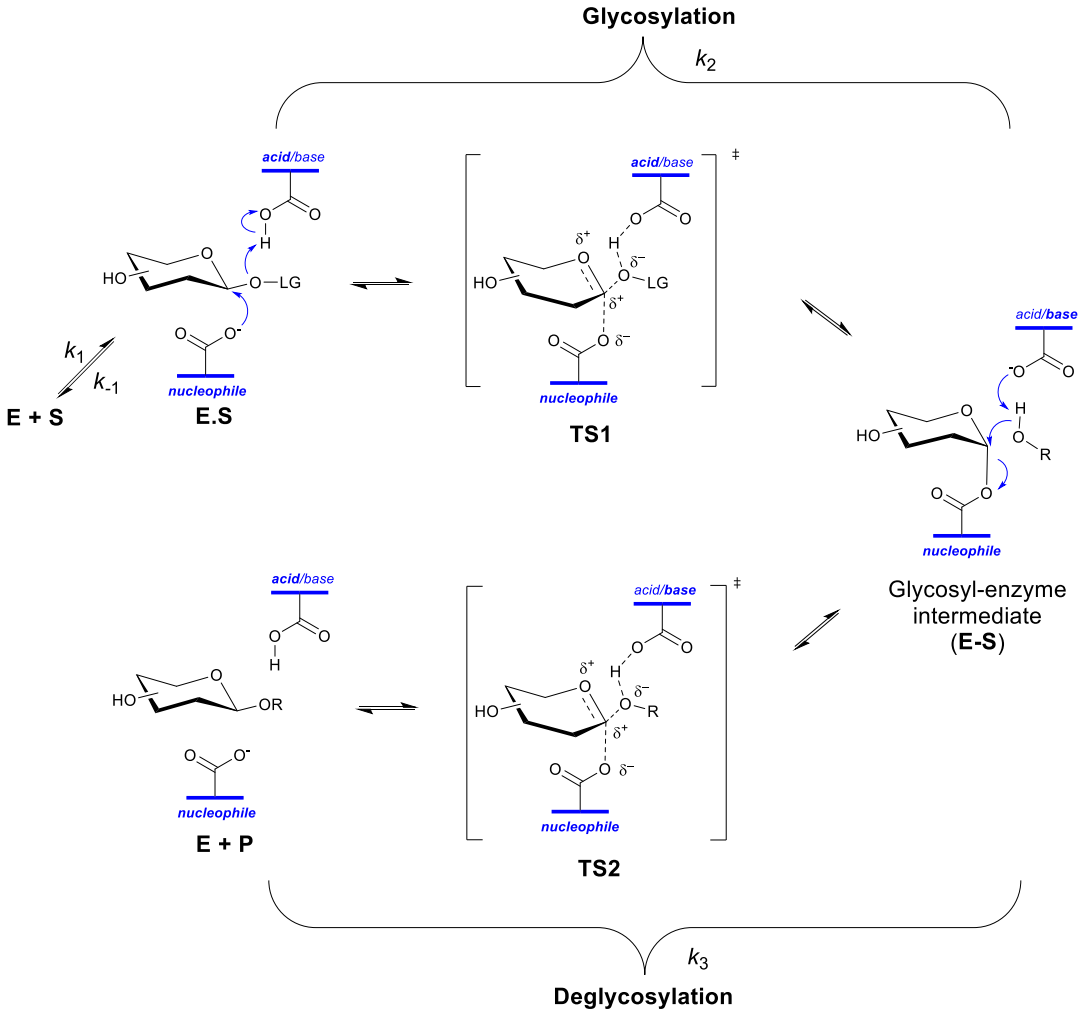


Fig. 15. Double-displacement mechanism of rGHs involving oxocarbenium-like transition states.

2.3.5. Comparing the p*K*a of catalytic residues in TGs and rGHs

Several studies have investigated the importance of the pKa of the catalytic residues for hydrolysis and transglycosylation. For example, the substitution of the catalytic nucleophile in a GH13 dextran glucosidase from *Streptococcus mutans* by a cysteinesulfinate (i.e. SOO^- function) provoked a drastic drop in k_{cat} value, combined with an acidic pK_a shift (from 3.9 to 1.5), and an increase in transglucosylation yields (Saburi et al. 2013). Similarly, shortening the nucleophile residue (E78D) in the GH11 *Bacillus circulans* xylanase (*Bc*Xyn) reduced global catalytic efficiency (1600- to 5000-fold decrease compared to the wild-type). Carboxymethylation of the *Bc*Xyn mutant E78C, which is associated with both lengthening of the nucleophile and an acidic pK_a shift from 4.6 to 3.3, partially restored the activity. Unfortunately, information concerning the transglycosylation activity was not reported

(Lawson, Wakarchuk, and Withers 1996). Likewise, shortening the length of the nucleophile's sidechain (E134D) in a GH16 (1,3)-(1,4)-β-glucanase from *Bacillus licheniformis* enhanced glycosynthase activity. The resulting hydrolase-glycosynthase enzyme, which retained 2% residual hydrolytic activity, displayed a lower pK_a value (from 7.2 to 5.8) for its acid/base catalytic residue (Andrés, Aragunde, and Planas 2014). Moreover, a mutant GH51 α-L-arabinofuranosidase (*Tx*Abf R69H) exhibiting a T/H partition shift in favor of transglycosylation also displays a basic pK_a shift (from 4.6 to 5.3) of its nucleophile residue (E298) and an acidic pK_a shift (from 7.6 to 6.9) of its acid/base residue (E176) compared to the parental enzyme (Bissaro, Durand, et al. 2015). Finally, an *in silico* study performed on GH3 β-glucosidase from *Hypocrea jecorina* revealed that the pK_a of the catalytic acid/base residue (E441) is low (2.0-2.6), irrespective of whether an acceptor molecule is present or not. The authors postulate that the low basicity of E441 reduces its ability to deprotonate acceptor molecules, including water. However, the various hydrogen bond networks present in the active site of the apo enzyme and in its bound state, notably that involving R169 within subsite +1, were thought to overcome this limitation (Geronimo, Payne, and Sandgren 2018a). Therefore, one can conclude that while the pK_a of the catalytic residues is an important consideration to take into account, this factor alone cannot be used to predict the T/H partition. Hydrogen bond networks, in particular in the acceptor subsite(s), also need be explored, because these can also enhance transglycosylation activity.

2.4. Engineering strategies towards artificial transglycosylases

Although the glycosynthase strategy has proven to be a quite efficient way to convert GHs into transglycosylases, it is not universally applicable for the reasons discussed earlier. Therefore, a large innovation space remains for the development of alternative strategies. In this regard, a considerable amount of work focusing on the modulation of the T/H partition in rGH has been published. The idea is to better understand the molecular determinants of this partition in order to establish rules that will guide the conversion of rGHs into TGs.

2.4.1. Negative subsite interactions

There is now a considerable amount of literature reporting the fact that the conversion of rGHs into TGs requires modifications in the so-called "negative" subsites of rGHs. The modification of enzyme-substrate interactions involving the sugar(s) lying upstream of the scissile bond appears to be a prerequisite for transglcosylation. This can be achieved in two ways: either by modifying key amino acids that define the negative subsites, or by using a

glycosyl donor whose structure does not properly fulfill the specificity requirements of the enzyme. In either case, such modifications can lead to transition state (TS) destabilization, meaning that the overall catalytic mechanism is penalized, irrespective of whether deglycosylation is water- or sugar-mediated. The consequences of TS destabilization are increased energy costs both to form the glycosyl-enzyme intermediate and, because of TS symmetry, to break it down (*i.e.* deglycosylation). Overall this means that although the covalent glycosyl-enzyme intermediate is more difficult to establish (TS1), once formed its half-life is prolonged (Bissaro, Monsan, et al. 2015). As an energy diagram shows (Fig. 16), the net effect of these modifications is a bigger penalty for water-mediated deglycosylation than for sugar-mediated deglycosylation, which is explained by the fact that the latter is less energy demanding. In other terms, hydrolysis is reduced more than transglycosylation, a conclusion that can be deduced from the comparison of naturally-occurring TGs with their hydrolytic counterparts (Kim et al. 2008). From a kinetic point of view, assuming that deglycosylation is the rate-limiting step (*i.e.* generally the case when sufficiently activated donors are employed), transglycosylases are characterized by rather low k_{cat}/K_M values when compared to hydrolytic rGH counterparts (Teze et al. 2014, 2015).

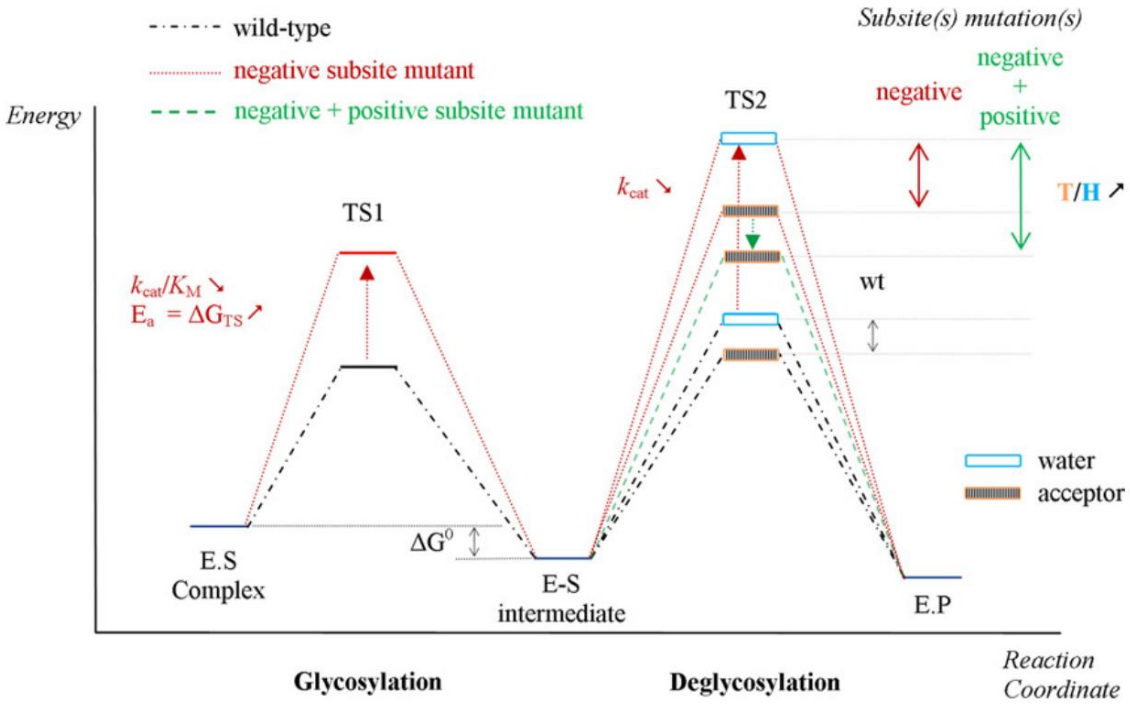


Fig. 16. Energy diagram of the two-step displacement mechanism of rGHs (black dash-dot) and alternative energetic pathways for evolved TGs (red dot or green dash for negative and negative plus positive subsite mutants respectively. From (Bissaro, Monsan, et al. 2015)

Experimentally, the relevance of mutating donor subsite amino acids to increase transglycosylation has been demonstrated several times. For example, galacto-oligosaccharide (GOS) production using a β-glucosidase (from *Thermotoga naphthophila* RKU-10) was enhanced by 50% simply by introducing a single (F414S) in subsite -1. The kinetic consequences of the mutation were a 22-fold decrease in catalytic efficiency, or k_{cat}/K_M (Yang et al. 2017). Similarly, the mutation of residue D142 in the GH18 chitinase from *Serratia marcescens* (*Sm*ChiB) led to better transglycosylation. This is significant, because this enzyme operates via a substrate-assisted mechanism. Subsequent QM/MM calculations performed on the hyper-transglycosylating D142N mutant, which forms part of a highly conserved DxDxE motif, concluded that the amino acid substitution affects both TS stabilization and the catalytic water molecule (Zakariassen et al. 2011).

The use of sub-optimal donor substrates (*i.e.* ones that are do not exactly fit the specificity requirements of an enzyme) has also been experimentally exemplified. Notably, the GH51 *Tx*Abf, which optimally recognizes α-L-Ara*f* moieties in its subsite -1, displays higher transglycosylation in the presence of β-D-Gal*f*O*p*NP. The transglycosylation yield can attain 75% when α-D-Xyl*p*OBn is used as the acceptor, while the equivalent reaction using α-L-Ara*f*O*p*NP only procures 7% transglycosylation product yield (Rémond et al. 2004; Rémond et al. 2005). Consistently, the catalytic efficiency (k_{cat}/K_M) of the reaction involving β-D-Gal*f*O*p*NP is low and mainly caused by an extremely high K_M value.

2.4.2. Acceptor subsite modification

Positive subsite interactions are also known to be important for transglycosylation, because these can contribute higher binding free energy to interactions with the incoming acceptor (Toshima et al. 2003; Kawamura et al. 2004). Moreover, stacking interactions and hydrogen bonding with the acceptor moiety appear to contribute most to the enhancement of transglycosylation (Rosengren et al. 2012; Dilokpimol et al. 2011). Therefore, in the quest to convert rGHs into TGs, it is important to address enzyme affinity for the acceptor.

Several studies have highlighted the importance of aromatic residues in the positive subsites (Taira et al. 2010; Johansson et al. 2004). The mutation Y217F in the *Mucor hiemalis* ENGase increased both positive subsite hydrophobicity and transglycosylation. Similarly, mutation of so-called ‘gatekeeper’ amino acids (Trp216 and Trp244) in the *Arthrobacter protophormiae* ENGase also increased transglycosylation, an effect that was partially attributed to alterations in active site dynamics (Yin et al. 2009; Umekawa et al. 2008).

Quite often tryptophan residues located in positive subsites are regioselectivity and transglycosylation determinants. Accordingly, in some cases the introduction of tryptophan in positive subsites will switch regioselectivity (Tran et al. 2010), while in others it will provide new opportunities for ring stacking, either with the acceptor sugar ring, or with other aromatic moieties (e.g. *p*NP) borne by the acceptor compound (Bayón et al. 2013). Computational simulation for AfChiB1 (chitinases found in *Aspergillus fumigatus*) suggested that the complete loss of transglycosylation coupled to the maintenance of hydrolysis in the mutant W137E is due to the loss of a stacking interaction between Trp137 and the glycoside acceptor (Lü et al. 2009). It was also speculated that this interaction would be necessary for the efficient attack of the oxazolinium ion intermediate (Lü et al. 2009). In a tetrasaccharide cycloalternan-catalyzing enzyme (CA), W430 appears to be poised to act as a hydrophobic shield, sterically hampering the optimal approach of nucleophilic water, thus reducing hydrolysis (Light et al. 2017). Jamek et al. made a similar conclusion, suggesting that a hydrophobic shield protects the glycosyl-enzyme intermediate from water molecules, thus reducing the probability of water-mediated deglycosylation (Jamek et al. 2018).

Overall, it is possible to conclude that the modification of positive subsite determinants can be used to improve acceptor recognition and positioning for transglycosylation. However, knowledge gleaned from the literature suggests that simply modifying acceptor subsites is insufficient to achieve significant gains in transglycosylation yield. Nevertheless, when combined with perturbations in the negative subsites, positive subsite modifications can tip the balance, compensating for TS destabilization and possibly providing a situation where $TS2_{Sugar}$ energy is lower than that of $TS2_{water}$ (Fig. 16). To achieve this, it is much better to target the proton network in the negative subsite, if possible, combined with the positive site engineering (Lafraya et al. 2011; Álvaro-Benito et al. 2012)

2.4.3. Water activation and channels

It is widely held that hydrolytic enzymes are characterized by channels that specifically guide substrates, including water molecules into the active site (Pravda et al. 2014). Therefore, intuitively one way to diminish hydrolysis might be to obstruct the way in which catalytic water accesses the active site and/or the way it is positioned to perform catalysis.

In some studies, tyrosine residues have been identified as water-binding determinants. For example, the introduction of a tyrosine close to the active site (V286Y) of the *Bacillus licheniformis* α-amylase led to a 5-fold increase in hydrolytic activity compared to the wild-

type enzyme. This observation is consistent with the hypothesis that the tyrosine hydroxyl group favors water access to the active site (Rivera et al. 2003) and also provides catalytic assistance to activate the water molecule (Umekawa et al. 2008). Conversely, removing a hydrogen bond that putatively fixes catalytic water has been successfully used to increase transglycosylation in xylose-releasing exo-oligoxylanase (Rex) (Honda et al. 2008).

Regarding water access, molecular dynamics simulations and crystallography have revealed details of how water is organized in enzymes. Both water clusters and ordered water chains have been observed (Teze et al. 2013; David et al. 2017). Putative water channels that potentially transport catalytic water to the catalytic site have been demonstrated in GH1 (Teze et al. 2013) and GH117 (Rebuffet et al. 2011). Similarly a channel that leads from the active site of a GH68 β-fructofuranosidase from *Arthrobacter sp*. was suspected to play a role in the drainage of water out of the catalytic site (Tonozuka et al. 2012). Although this function was not experimentally demonstrated, it would limit the solvation of the catalytic site and thus explain the predominance of transglycosylation compared to hydrolysis.

It is noteworthy to emphasize the functional importance of water channels in hydrolases. When water channel amino acids are mutated, for example in *Candida antartica* lipase B (Wittrup Larsen et al. 2010) or in *Tenebrio molitor* β-glycosidase (Frutuoso and Marana 2012), reduced hydrolysis is obtained. Regarding transglycosylation, the mutation of three amino acids (D341, Q342, and S351) in *endo*-β-agarase (AgaD), suspected to be involved in the control of water access to the active site (Fig. 17), led to a 50-fold reduction in hydrolysis, while the transglycosylase activity was maintained and even slightly improved.

All of these studies converge towards the conclusion that disturbing water binding (*e.g*. mutation of water-binding tyrosine residues), or water delivery systems (*e.g.* blocking channels by altering internal polarity) can affect hydrolytic potency. Therefore, these approaches should be taken into consideration when altering the T/H balance to improve glycosynthesis. However, as demonstrated, a prerequisite to achieve this is to use suitable *in silico* predictive methods and high resolution experimental structural biology approaches to precisely locate the water molecules within GHs.

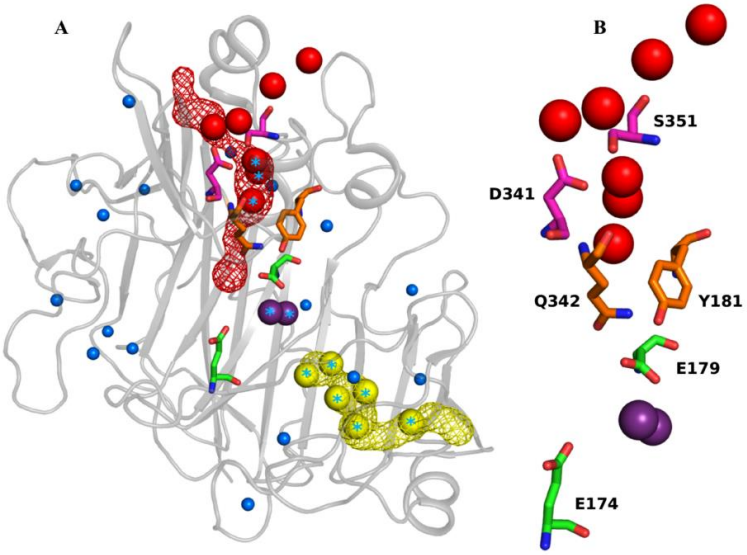


Fig. 17. (A) Internal water molecules (blue spheres and stars) within the crystal structure of AgaD (PDB 4ASM) with putative water channels 1 (red mesh) and 2 (yellow mesh) as predicted by Caver 3.0. (B) Water chain corresponding to channel 1 and the amino acids lining the water channel 1 (in magenta and orange). The catalytic residues are shown in green, and well-ordered water molecules closest to the catalytic acid base residue (E179) are depicted in purple. (David et al. 2017)

2.4.4. Conservation-based strategy

Some of the key determinants of the T/H partition in rGHs appear to be conserved residues that have thus persisted throughout evolution (Feng et al. 2005; Koné et al. 2009; Bissaro, Durand, et al. 2015). Therefore, modifying conserved residues in the first and second shells[1], particularly in the negative subsite(s), has been proposed as a way to alter the T/H equilibrium. Experimental data show that such as strategy often leads to increased transglycosylation yield, although this is usually at the expense of catalytic efficiency caused by the modification of the TS stabilization. Based on sequence-based identity, Teze *et al.* studied the conserved residues in a GH1 *Thermus thermophilus* β-glycosidase (Ttb-gly). Targeting seven first shell residues in −1 subsite for mutation procured significant enhancement of transglycosylation, with a disaccharide being produced at 65–82% yield, compared to 36% for the wild-type enzyme (Teze et al. 2014). This conservation–based strategy was further refined and applied to a GH36 *Geobacillus stearothermophilus* α-galactosidase (AgaB). In this study, second-shell conserved residues, forming part of subsite -1, were successfully targeted using mutagenesis

[1] First-shell residues are those that residues interact directly with the substrate, while those that directly interact with one or more first-shell residues are called second-shell residues.

(Teze et al. 2015). The advantage of this conservation-based strategy compared to fully random approaches is that it reduces screening and thus intensive lab work. However, not all of the mutations improve transglycosylation. Another advantage of this strategy is that once hotspots are detected, the knowledge can be transferred to other GHs that belong to the same clan (Zeuner et al. 2019). Overall, the conservation-based approach is powerful, but will necessarily miss key residues when these are not conserved. This was the case for residues located in the +1 and +2 sites of enzymes targeted in the aforementioned studies (Teze et al. 2014, 2015).

3. Evolution of *Tx*Abf

3.1. Distribution of α-L-arabinofuranosidase in CAZy

The α-L-arabinofuranosidases (Abfs) function as debranching enzymes that remove L-arabinofuranosyl substituents from L-arabinose-containing polymers, including arabinoxylan and arabinan that are plant cell wall structural components (Dumon et al. 2012; Poria et al. 2020). From a chemical reaction standpoint, Abfs are grouped into the class of enzymes designated glycoside hydrolases (EC 3.2.1.-). However, there are two types of Abfs, *endo*-Abfs (EC 3.2.1.99) and *exo*-Abfs (EC 3.2.1.55). *Endo*-Abfs cleave internal α-L-Ara*f*-(1,5)-linkages in arabinans, whereas *exo*-Abfs remove terminal α-L-Ara*f* moieties from oligomers and polymers cleave, cleaving (1,2)-, (1,3)-, and/or (1,5)-bonds (Wilkens et al. 2017; Thakur, Sharma, and Goyal 2019). Depending on the sequence similarity, *exo*-Abfs have been assigned to different families in the CAZy classification. These include GH2, GH3, GH43, GH51, GH54, and GH62, with families GH2 and GH51 belonging to the large Clan-GHA, while GH62 belongs to Clan-GHF (Shi et al. 2014). It is noteworthy that *endo*-Abfs have also been identified in GH43 (Kaji and Saheki 1975; Leal and Sá-Nogueira 2004).

3.2. Features of *Tx*Abf

3.2.1. Biochemical properties

The GH51 *Tx*Abf has been the subject of research for two decades. The bacterial strain *Thermobacillus xylanilyticus* that naturally produces this enzyme also produces xylanases and is thus considered to be competent for bioconversion of plant cell wall hemicelluloses (Touzel et al. 2000). *Tx*Abf displays a molecular weight of 56 kDa, is optimally active at 75 °C and between pH 5.6 and 6.2 and retains 50% of its maximum activity after two hours incubation at 90 °C (Debeche et al. 2000). Regarding its catalytic properties, *Tx*Abf cleaves both (1,2)- and (1,3)-bonds that link terminal α-L-Ara*f* moieties to D-Xyl*p* residues contained with AXs, which appear to be the class of substrates for which it is best-adapted (C. Rémond et al. 2008). Within the simplified context of laboratory experiments, *Tx*Abf is most active on α-L-Ara*f*O*p*NP, which is consistent with its putative ‘natural’ glycone specificity (Borsenberger et al. 2014).

3.2.2. 3D structure

The 3D structures of *Tx*Abf in both apo and complexed (with XA^3XX) states (PDB: 2VRK and 2VRQ) have been solved (Paës et al. 2008). These reveal that *Tx*Abf consists of two domains: the catalytic domain that displays a $(\beta/\alpha)_8$ barrel structure (also known as TIM barrel architecture) and a C-terminal domain that adopts a jelly-roll fold. The function of the latter is unknown, but it is likened to the widespread fibronectin III (FN3) domain. Like all members of clan GH-A, the catalytic domain of *Tx*Abf contains two glutamate catalytic residues that fulfill acid/base (E176) and nucleophile (E298) functions. In the complexed enzyme structure, position 176 is occupied by a glutamine (Fig. 18), because it was necessary to inactivate the enzyme to obtain a stable enzyme-substrate complex (Debeche et al. 2002). A molecular dynamics study reveals that two residues, H98 and W99, are crucial for donor substrate recognition and impact mobility of loop β2α2 that constitutes a key feature of substrate binding and catalysis (Fig. 18) (Arab-Jaziri et al. 2012).

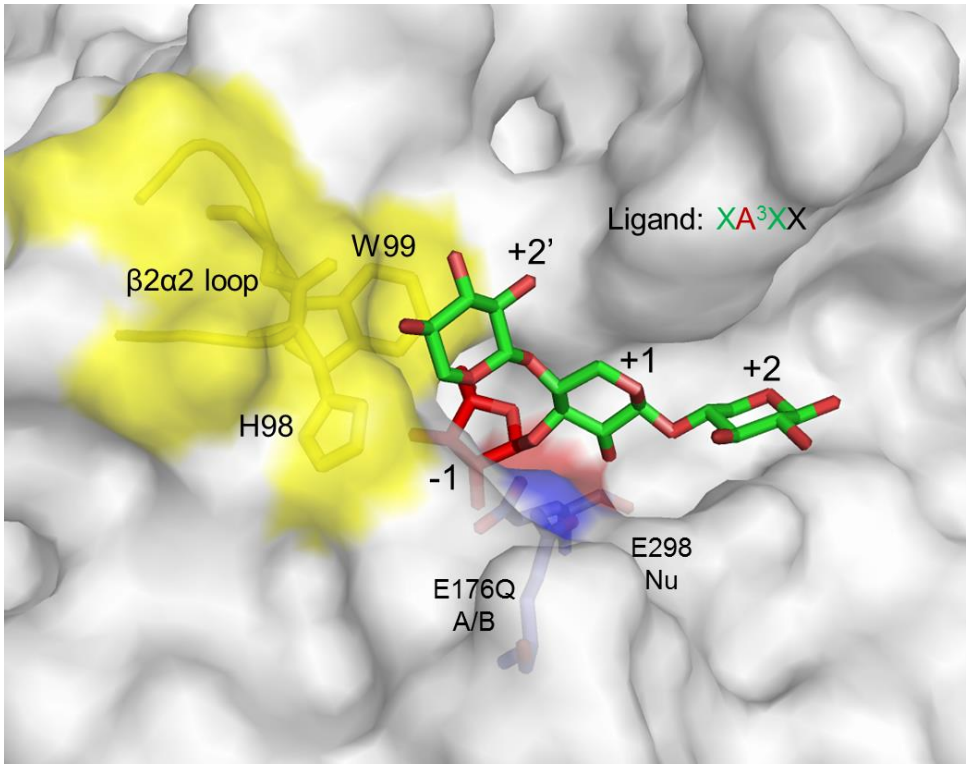


Fig. 18 Active site topography of *Tx*Abf co-crystallized with XA^3XX (α-L-Ara*f* in red and D-Xyl*p* in green) (PDB: 2VRQ). Catalytic residues E298 and E176 (mutated into Gln) are shown in purple. Residues H98, W99 and the relevant β2α2 loop are shown in yellow.

Examination of the *Tx*Abf structure reveals that the active site adopts a pocket like topology, typical of an *exo*-enzyme. In addition to the two catalytic residues, the active site is composed

of ten other amino acids (E28, C74, H98, W99, N175, N216, Y242, W248, W302 and Q347) that are located on different loops. The pocket-like structure receives the L-Ara*f* moiety from the substrate and constitutes subsite -1. Conversely, subsites +1 and +2 are defined by a surface exposed groove that accommodates the reducing side elements of the substrate (Fig. 18). Although it is possible that *Tx*Abf bears a subsite +3, this has so far not been evidenced, because in the original crystal structures the reducing end D-Xyl*p* unit was ill-defined. Regarding the L-Ara*f* moiety in subsite -1, it forms seven direct hydrogen bond interactions with residues E28, C74, N175, E176, Y242, E298 and Q347 (Fig. 19), whereas D-Xyl*p* in subsite +1 forms vdW interactions with Y242, W248 and W302, a direct hydrogen bond with E176 and another indirect one, via a water molecule, with W216. In subsite +2, D-Xyl*p* interacts with N216 and W248. Importantly, because *Tx*Abf recognizes L-Ara*f* moieties that decorate xylan chains, the enzyme also displays other subsites that are designated using the (') symbol. Accordingly, the structure of the complexed enzyme reveals a subsite +2' in which a non-reducing D-Xyl*p* moiety interacts with H98 and W99 (Paës et al. 2008).

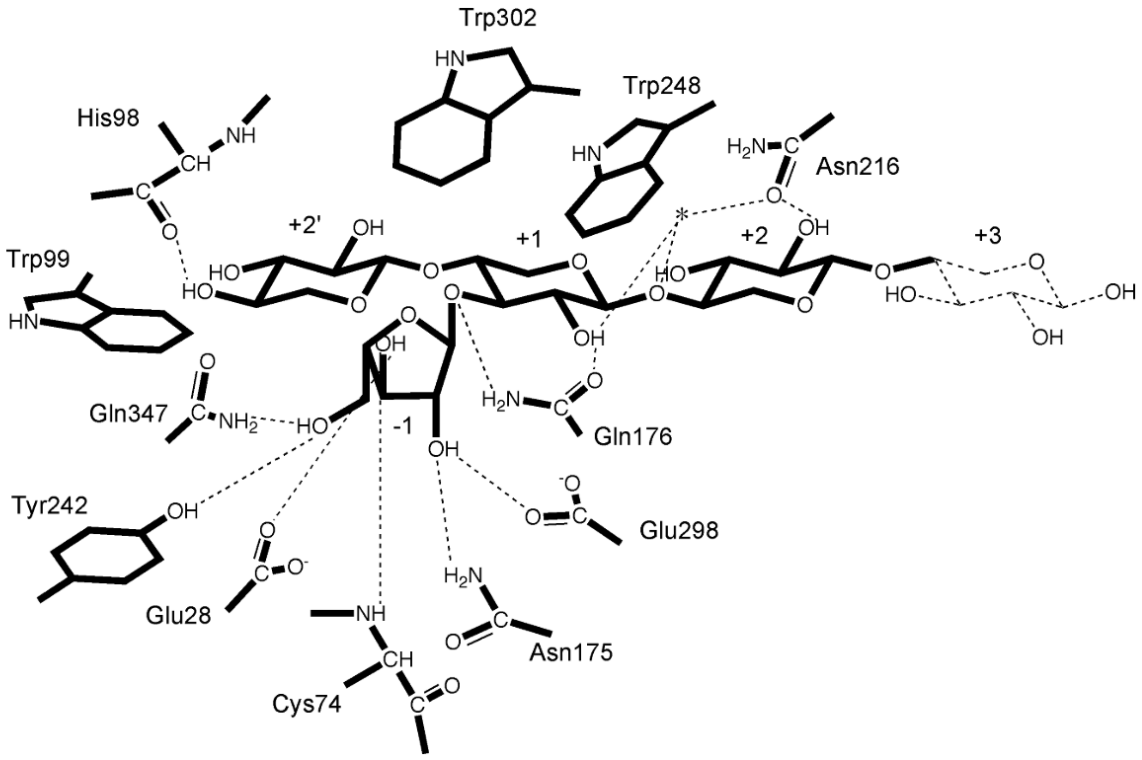


Fig. 19. Schematic representation of the XA^{3}XX interactions. Hydrogen bonds are shown as dotted lines. Water molecules are represented with an asterisk. Subsite +3 D-Xyl*p* is shown with dashed lines since it is not observed in the structures. (Paës et al. 2008)

3.3. The ability of *Tx*Abf to perform transglycosylation

3.3.1. On protein engineering

For nearly a decade, directed evolution has been used to improve the glycosynthetic potential of *Tx*Abf. In the early period, random mutagenesis was applied and mutants displaying potentially improved transglycosylation activity were screened using a two-step screening strategy (Arab-Jaziri et al. 2013). Briefly, in the first step mutants displaying decreased hydrolytic activity were detected using a chromogenic donor substrate, while in the second step the activity of mutants were screened in the presence of both donor and acceptor in order to detect transglycosylation (Arab-Jaziri et al. 2013, 2015). This strategy procured the mutant L352M, which was subsequently used as a template for further random mutagenesis. This ultimately led to the detection of the mutant R69H and the creation of the double mutant R69H-L352M, in which both modified residues are located in subsite -1. More recent work focused on acceptor subsite mutants, using techniques such as site-saturation mutagenesis and *in silico* prediction (*i.e.* the as yet unpublished Bindscan approach). This work revealed the mutant N216W. Finally, the study of the triple mutant R69H-N216W-L352M provided evidence to claim that this is a *bona fide* transarabinofuranosylase that reaches transglycosylation yields of up to 80% when using 15 mM of α-L-Ara*f*O*p*NP as donor and 30 mM of xylotriose as acceptor (compared to 11% for wild-type *Tx*Abf) (Arab-Jaziri et al. 2015; Bissaro, Durand, et al. 2015).

3.3.2. On different reactions

To probe the synthetic versatility of *Tx*Abf, different reactions were performed using different substrates, that are compared herein with the catalytic ability of a related Abf from *Ruminiclostridium thermocellum* (Table 5). Among the notable successes, β-D-Gal*f*O*p*NP was found to be a useful donor compound for transglycosylation, since yields surpass 20% and even reach 75% when using β-D-Gal*f*-(1,2)-α-D-Xyl*p*OBn as acceptor with wild-type *Tx*Abf.

Table 5 Synthetic reactions catalyzed by *Tx*Abf and *Ct*Abf

***Tx*Abf**				
Donor (5mM)	**Acceptor**	**Products**	**Yield (%)**	**Reference**
α-L-Ara*f*O*p*NP	Share with donor	Major: α-L-Ara*f*-(1,2)-α-L-Ara*f*O*p*NP Minor: (1,3) and (1,5)	8	(Caroline Rémond et al. 2004)
	α-D-Xyl*p*OBn 5mM	α-L-Ara*f*-(1,2)-α-D-Xyl*p*OBn	7	
	Alcohol (23%, v/v) (methanol, ethanol, *n*-propanol, etc.)	α-L-Ara*f*OMe, α-L-Ara*f*OEt,etc.	24-54	
	Xylotriose 10 mM	XA^3X	3	(Arab-Jaziri et al. 2015)
		A^3XX	4	
		A^2XX	6	
β-D-Xyl*p*O*p*NP	Share with donor	β-D-Xyl*p*-(1,2)-β-D-Xyl*p*O*p*NP	6	(Remond et al. 2005)
		β-D-Xyl*p*-(1,3)-β-D-Xyl*p*O*p*NP	6	
β-D-Gal*f*O*p*NP	Share with donor	β-D-Gal*f*-(1,2)-β-D-Gal*f*O*p*NP	21	
	α-D-Xyl*p*OBn 5 mM	β-D-Gal*f*-(1,2)-α-D-Xyl*p*OBn	75	
	α-D-Xyl*p*O*p*NP 5 mM	Disaccharides (TLC)	-	
	β-D-Xyl*f*O*p*NP 5 mM	Disaccharides (TLC)	-	
β-D-Fuc*f*O*p*NP	Share with donor	Disaccharides (HPLC)	-	(Euzen et al. 2005)
6F-β-D-Gal*f*O*p*NP	Share with donor	Disaccharides (HPLC)	-	
***Ct*Abf**				
Donor (5mM)	**Acceptor (50 mM)**	**Products**	**Yield (%)**	**Reference**
β-D-Gal*f*O*p*NP	α-D-Gal*p*O*p*NP	β-D-Gal*f*-(1,2)-α-D-Gal*p*O*p*NP	41	(Chlubnová et al. 2014)
		β-D-Gal*f*-(1,3)-α-D-Gal*p*O*p*NP	32	
	β-D-Gal*p*O*p*NP	β-D-Gal*f*-(1,2)-β-D-Gal*p*O*p*NP	27	
		β-D-Gal*f*-(1,6)-β-D-Gal*p*O*p*NP	25	

	α-D-Man*p*O*p*NP	β-D-Gal*f*-(1,2)-α-D-Man*p*O*p*NP, mixed with (1,3) and (1,4)	43	
		β-D-Gal*f*-(1,6)-α-D-Man*p*O*p*NP	19	
	β-D-Man*p*O*p*NP	β-D-Gal*f*-(1,4)-β-D-Man*p*O*p*NP	49	
		β-D-Gal*f*-(1,6)-β-D-Man*p*O*p*NP	16	
	α-L-Rha*p*O*p*NP	β-D-Gal*f*-(1,4)-α-L-Rha*p*O*p*NP	38	

4. Thesis objectives

Building on the abundant knowledge that has already been gathered, this doctoral study aimed to pursue the investigation of *Tx*Abf, particularly regarding its ability to catalyze transglycosylation reactions. Our aim in this work was to address some outstanding questions and uncover finer details of some of the mutants that are available. Specifically, we set out to:

1) Refine our understanding of the T/H partition in *Tx*Abf and increment our knowledge base.

2) Explore to what extent our hitherto cumulated knowledge can be transferred to other clan GH-A enzymes

3) Investigate how knowledge from the study of *Tx*Abf acting on α-L-Ara*f*-containing substrates can help to better understand and modify the activity of *Tx*Abf on β-D-Gal*f*-containing substrates, the aim here being to define mutants that will form a new generation of synthetic tools, thus providing access to a range of biologically-relevant glycoconjugates.

References

Alderwick, Luke J., Helen L. Birch, Karin Krumbach, Michael Bott, Lothar Eggeling, and Gurdyal S. Besra. 2018. 'AftD Functions as an α1→5 Arabinofuranosyltransferase Involved in the Biosynthesis of the Mycobacterial Cell Wall Core'. *The Cell Surface* 1 (March): 2–14. https://doi.org/10.1016/j.tcsw.2017.10.001.

Alderwick, Luke J., Mathias Seidel, Hermann Sahm, Gurdyal S. Besra, and Lothar Eggeling. 2006. 'Identification of a Novel Arabinofuranosyltransferase (AftA) Involved in Cell Wall Arabinan Biosynthesis in Mycobacterium Tuberculosis'. *Journal of Biological Chemistry* 281 (23): 15653–61. https://doi.org/10.1074/jbc.M600045200.

Álvaro-Benito, Miguel, M. Angela Sainz-Polo, David González-Pérez, Beatriz González, Francisco J. Plou, María Fernández-Lobato, and Julia Sanz-Aparicio. 2012. 'Structural and Kinetic Insights Reveal That the Amino Acid Pair Gln-228/Asn-254 Modulates the Transfructosylating Specificity of Schwanniomyces Occidentalis β-Fructofuranosidase, an Enzyme That Produces Prebiotics'. *Journal of Biological Chemistry* 287 (23): 19674–86. https://doi.org/10.1074/jbc.M112.355503.

Amaya, Maria Fernanda, Alejandro Buschiazzo, Tong Nguyen, and Pedro M. Alzari. 2003. 'The High Resolution Structures of Free and Inhibitor-Bound Trypanosoma Rangeli Sialidase and Its Comparison with T.CruziTrans-Sialidase'. *Journal of Molecular Biology* 325 (4): 773–84. https://doi.org/10.1016/S0022-2836(02)01306-2.

Amrein, Thomas M., Peter Gränicher, Eva Arrigoni, and Renato Amadò. 2003. 'In Vitro Digestibility and Colonic Fermentability of Aleurone Isolated from Wheat Bran'. *LWT - Food Science and Technology* 36 (4): 451–60. https://doi.org/10.1016/S0023-6438(03)00036-7.

Anders, N., M. D. Wilkinson, A. Lovegrove, J. Freeman, T. Tryfona, T. K. Pellny, T. Weimar, et al. 2012. 'Glycosyl Transferases in Family 61 Mediate Arabinofuranosyl Transfer onto Xylan in Grasses'. *Proceedings of the National Academy of Sciences* 109 (3): 989–93. https://doi.org/10.1073/pnas.1115858109.

Andrés, Eduardo, Hugo Aragunde, and Antoni Planas. 2014. 'Screening Glycosynthase Libraries with a Fluoride Chemosensor Assay Independently of Enzyme Specificity: Identification of a Transitional Hydrolase to Synthase Mutant'. *Biochemical Journal* 458 (2): 355–63. https://doi.org/10.1042/BJ20131057.

Arab-Jaziri, Faten, Bastien Bissaro, Sophie Barbe, Olivier Saurel, Hélène Débat, Claire Dumon, Virginie Gervais, et al. 2012. 'Functional Roles of H98 and W99 and β2α2 Loop Dynamics in the α-l-Arabinofuranosidase from Thermobacillus Xylanilyticus'. *The FEBS Journal* 279 (19): 3598–3611. https://doi.org/10.1111/j.1742-4658.2012.08720.x.

Arab-Jaziri, Faten, Bastien Bissaro, Michel Dion, Olivier Saurel, David Harrison, Fernando Ferreira, Alain Milon, Charles Tellier, Régis Fauré, and Michael J. O'Donohue. 2013. 'Engineering Transglycosidase Activity into a GH51 α-L-Arabinofuranosidase'. *New Biotechnology* 30 (5): 536–44. https://doi.org/10.1016/j.nbt.2013.04.002.

Arab-Jaziri, Faten, Bastien Bissaro, Charles Tellier, Michel Dion, Régis Fauré, and Michael J. O'Donohue. 2015. 'Enhancing the Chemoenzymatic Synthesis of Arabinosylated Xylo-Oligosaccharides by GH51 α-L-Arabinofuranosidase'. *Carbohydrate Research* 401: 64–72. https://doi.org/10.1016/j.carres.2014.10.029.

Aragunde, Hugo, Estela Castilla, Xevi Biarnés, Magda Faijes, and Antoni Planas. 2014. 'A

Transitional Hydrolase to Glycosynthase Mutant by Glu to Asp Substitution at the Catalytic Nucleophile in a Retaining Glycosidase'. *Carbohydrate Research* 389 (1): 85–92. https://doi.org/10.1016/j.carres.2014.02.003.

Arruda, Monika V., Walter Colli, and Bianca Zingales. 1989. 'Terminal Beta-d-Galactofuranosyl Epitopes Recognized by Antibodies That Inhibit Trypanosoma Cruzi Internalization into Mammalian Cells'. *European Journal of Biochemistry* 182 (2): 413–21. https://doi.org/10.1111/j.1432-1033.1989.tb14847.x.

Aspeborg, Henrik, Jarmo Schrader, Pedro M. Coutinho, Mark Stam, Åsa Kallas, Soraya Djerbi, Peter Nilsson, et al. 2005. 'Carbohydrate-Active Enzymes Involved in the Secondary Cell Wall Biogenesis in Hybrid Aspen'. *Plant Physiology* 137 (3): 983–97. https://doi.org/10.1104/pp.104.055087.

Ayers, Joseph D., Todd L. Lowary, Caroline B. Morehouse, and Gurdyal S. Besra. 1998. 'Synthetic Arabinofuranosyl Oligosaccharides as Mycobacterial Arabinosyltransferase Substrates'. *Bioorganic & Medicinal Chemistry Letters* 8 (5): 437–42. https://doi.org/10.1016/S0960-894X(98)00049-3.

Baumann, Martin J., Jens M. Eklöf, Gurvan Michel, Åsa M. Kallas, Tuula T. Teeri, Mirjam Czjzek, and Harry Brumer. 2007. 'Structural Evidence for the Evolution of Xyloglucanase Activity from Xyloglucan Endo -Transglycosylases: Biological Implications for Cell Wall Metabolism'. *The Plant Cell* 19 (6): 1947–63. https://doi.org/10.1105/tpc.107.051391.

Bayón, Carlos, Álvaro Cortés, José Berenguer, and María J. Hernáiz. 2013. 'Highly Efficient Enzymatic Synthesis of Galβ-(1→3)-GalNAc and Galβ-(1→3)-GlcNAc in Ionic Liquids'. *Tetrahedron* 69 (24): 4973–78. https://doi.org/10.1016/j.tet.2013.04.015.

Belanova, M., P. Dianiskova, P. J. Brennan, G. C. Completo, N. L. Rose, T. L. Lowary, and K. Mikusova. 2008. 'Galactosyl Transferases in Mycobacterial Cell Wall Synthesis'. *Journal of Bacteriology* 190 (3): 1141–45. https://doi.org/10.1128/JB.01326-07.

Benkoulouche, Mounir, Régis Fauré, Magali Remaud-Siméon, Claire Moulis, and Isabelle André. 2019. 'Harnessing Glycoenzyme Engineering for Synthesis of Bioactive Oligosaccharides'. *Interface Focus* 9 (2). https://doi.org/10.1098/rsfs.2018.0069.

Berg, Stefan, Devinder Kaur, Mary Jackson, and Patrick J Brennan. 2007. 'The Glycosyltransferases of Mycobacterium Tuberculosis—Roles in the Synthesis of Arabinogalactan, Lipoarabinomannan, and Other Glycoconjugates'. *Glycobiology* 17 (6): 35R-56R. https://doi.org/10.1093/glycob/cwm010.

Bissaro, Bastien. 2014. 'On Hydrolysis / Transglycosylation Modulation in Glycoside Hydrolases : Lessons Learnt from the Molecular Design of the First Non-Leloir Transarabinofuranosylases'. Université de Toulouse. http://www.theses.fr/2014ISAT0023.

Bissaro, Bastien, Julien Durand, Antoni Planas, Pierre Monsan, Xevi Biarnés, Antoni Planas, Pierre Monsan, et al. 2015. 'Molecular Design of Non-Leloir Furanose-Transferring Enzymes from an α-L-Arabinofuranosidase: A Rationale for the Engineering of Evolved Transglycosylases'. *ACS Catalysis* 5 (8): 4598–4611. https://doi.org/10.1021/acscatal.5b00949.

Bissaro, Bastien, Pierre Monsan, Régis Fauré, and Michael J. O'Donohue. 2015. 'Glycosynthesis in a Waterworld: New Insight into the Molecular Basis of Transglycosylation in Retaining Glycoside Hydrolases'. *Biochemical Journal* 467 (1):

17–35. https://doi.org/10.1042/BJ20141412.

Bissaro, Bastien, Olivier Saurel, Faten Arab-jaziri, Luc Saulnier, Alain Milon, Maija Tenkanen, Pierre Monsan, Michael J. O'Donohue, and Régis Fauré. 2014. 'Mutation of a PH-Modulating Residue in a GH51 α-L-Arabinofuranosidase Leads to a Severe Reduction of the Secondary Hydrolysis of Transfuranosylation Products'. *Biochimica et Biophysica Acta - General Subjects* 1840 (1): 626–36. https://doi.org/10.1016/j.bbagen.2013.10.013.

Blanco, Noelia, Ana B. Sanz, Jose M. Rodríguez-Peña, César Nombela, Vladimír Farkaš, Ramón Hurtado-Guerrero, and Javier Arroyo. 2015. 'Structural and Functional Analysis of Yeast Crh1 and Crh2 Transglycosylases'. *The FEBS Journal* 282 (4): 715–31. https://doi.org/10.1111/febs.13176.

Borsenberger, Vinciane, Emmie Dornez, Marie-Laure Desrousseaux, Stéphane Massou, Maija Tenkanen, Christophe M. Courtin, Claire Dumon, Michael J. O'Donohue, and Régis Fauré. 2014. 'A 1H NMR Study of the Specificity of α-L-Arabinofuranosidases on Natural and Unnatural Substrates'. *Biochimica et Biophysica Acta (BBA) - General Subjects* 1840 (10): 3106–14. https://doi.org/10.1016/j.bbagen.2014.07.001.

Borsenberger, Vinciane, Fernando Ferreira, Annick Pollet, Emmie Dornez, Marie-Laure Desrousseaux, Stéphane Massou, Christophe M. Courtin, Michael J. O'Donohue, and Régis Fauré. 2012. 'A Versatile and Colorful Screening Tool for the Identification of Arabinofuranose-Acting Enzymes'. *ChemBioChem* 13 (13): 1885–88. https://doi.org/10.1002/cbic.201200394.

Bruggen-Van der Lugt, A. W. Van, H. J. Kamphuis, G. A. De Ruiter, P. Mischnick, J. H. Van Boom, and F. M. Rombouts. 1992. 'New Structural Features of the Antigenic Extracellular Polysaccharides of *Penicillium* and *Aspergillus* Species Revealed with Exo-??-D- Galactofuranosidase'. *Journal of Bacteriology* 174 (19): 6096–6102.

Cardini, C. E., A. C. Paladini, R. Caputto, and L. F. Leloir. 1950. 'Uridine Diphosphate Glucose: The Coenzyme of the Galactose–Glucose Phosphate Isomerization'. *Nature* 165 (4188): 191–92. https://doi.org/10.1038/165191a0.

Champion, Elise, Frédéric Guérin, Claire Moulis, Sophie Barbe, Thu Hoai Tran, Sandrine Morel, Karine Descroix, et al. 2012. 'Applying Pairwise Combinations of Amino Acid Mutations for Sorting Out Highly Efficient Glucosylation Tools for Chemo-Enzymatic Synthesis of Bacterial Oligosaccharides'. *Journal of the American Chemical Society* 134 (45): 18677–88. https://doi.org/10.1021/ja306845b.

Chang, Aram, Shanteri Singh, George N. Phillips, and Jon S. Thorson. 2011. 'Glycosyltransferase Structural Biology and Its Role in the Design of Catalysts for Glycosylation'. *Current Opinion in Biotechnology*. https://doi.org/10.1016/j.copbio.2011.04.013.

Chlubnová, Ilona, Dominik Filipp, Vojtech Spiwok, Hana Dvoráková, Richard Daniellou, Caroline Nugier-Chauvin, Blanka Králová, et al. 2010. 'Enzymatic Synthesis of Oligo-D-Galactofuranosides and L-Arabinofuranosides: From Molecular Dynamics to Immunological Assays.' *Organic & Biomolecular Chemistry* 8 (9): 2092–2102. https://doi.org/10.1039/b926988f.

Chlubnová, Ilona, Blanka Králová, Hana Dvořáková, Petr Hošek, Vojtěch Spiwok, Dominik Filipp, Caroline Nugier-Chauvin, Richard Daniellou, and Vincent Ferrières. 2014. 'The Versatile Enzyme Ara51 Allowed Efficient Synthesis of Rare Pathogen-Related β-d-Galactofuranosyl-Pyranoside Disaccharides'. *Org. Biomol. Chem.* 12 (19): 3080–89.

https://doi.org/10.1039/c3ob42519c.

Chlubnova, Ilona, Laurent Legentil, Rémy Dureau, Alizé Pennec, Mélanie Almendros, Richard Daniellou, Caroline Nugier-Chauvin, and Vincent Ferrières. 2012. 'Specific and Non-Specific Enzymes for Furanosyl-Containing Conjugates: Biosynthesis, Metabolism, and Chemo-Enzymatic Synthesis'. *Carbohydrate Research* 356: 44–61. https://doi.org/10.1016/j.carres.2012.04.002.

Choi, Yun Hee, Jong Hoon Kim, Bum Seok Park, and Byung-Gee Kim. 2016. 'Solubilization and Iterative Saturation Mutagenesis of A1,3-Fucosyltransferase from Helicobacter Pylori to Enhance Its Catalytic Efficiency'. *Biotechnology and Bioengineering* 113 (8): 1666–75. https://doi.org/10.1002/bit.25944.

Classen, Birgit, Alexander Baumann, and Jon Utermoehlen. 2019. 'Arabinogalactan-Proteins in Spore-Producing Land Plants'. *Carbohydrate Polymers* 210 (April): 215–24. https://doi.org/10.1016/j.carbpol.2019.01.077.

Cousin, M.A., S. Notermans, P. Hoogerhout, and J.H. Van Boom. 1989. 'Detection of β-Galactofuranosidase Production by *Penicillium* and *Aspergillus* Species Using 4-Nitrophenyl β-D-Galactofuranoside'. *Journal of Applied Bacteriology* 66 (4): 311–17. https://doi.org/10.1111/j.1365-2672.1989.tb02484.x.

Danby, Phillip M., and Stephen G. Withers. 2016. 'Advances in Enzymatic Glycoside Synthesis'. *ACS Chemical Biology* 11 (7): 1784–94. https://doi.org/10.1021/acschembio.6b00340.

Daudé, David, Elise Champion, Sandrine Morel, David Guieysse, Magali Remaud-Siméon, and Isabelle André. 2013. 'Probing Substrate Promiscuity of Amylosucrase from *Neisseria Polysaccharea*'. *ChemCatChem* 5 (8): 2288–95. https://doi.org/10.1002/cctc.201300012.

David, Benoit, Philippe Arnaud, Charles Tellier, and Yves-Henri Sanejouand. 2019. 'Toward the Design of Efficient Transglycosidases: The Case of the GH1 of Thermus Thermophilus'. Edited by Valerie Daggett. *Protein Engineering, Design and Selection*, no. Md (October): 1–8. https://doi.org/10.1093/protein/gzz032.

David, Benoit, Romain Irague, Diane Jouanneau, Franck Daligault, Mirjam Czjzek, Yves-Henri Sanejouand, and Charles Tellier. 2017. 'Internal Water Dynamics Control the Transglycosylation/Hydrolysis Balance in the Agarase (AgaD) of *Zobellia Galactanivorans*'. *ACS Catalysis* 7 (5): 3357–67. https://doi.org/10.1021/acscatal.7b00348.

Davies, G. J., K. S. Wilson, and B. Henrissat. 1997. 'Nomenclature for Sugar-Binding Subsites in Glycosyl Hydrolases [1]'. *Biochemical Journal*. https://doi.org/10.1042/bj3210557.

Davies, G.J., V.M.-A. Ducros, A. Varrot, and D.L. Zechel. 2003. 'Mapping the Conformational Itinerary of β-Glycosidases by X-Ray Crystallography'. *Biochemical Society Transactions* 31 (3): 523–27. https://doi.org/10.1042/bst0310523.

Davies, Gideon, and Bernard Henrissat. 1995. 'Structures and Mechanisms of Glycosyl Hydrolases'. *Structure*. https://doi.org/10.1016/S0969-2126(01)00220-9.

Debeche, Takoua, Christophe Bliard, Philippe Debeire, and Michael J O'Donohue. 2002. 'Probing the Catalytically Essential Residues of the α-L-Arabinofuranosidase from *Thermobacillus Xylanilyticus*'. *Protein Engineering, Design and Selection* 15 (1): 21–28. https://doi.org/10.1093/protein/15.1.21.

Debeche, Takoua, Nicola Cummings, Ian Connerton, Philippe Debeire, and Michael J. O'Donohue. 2000. 'Genetic and Biochemical Characterization of a Highly Thermostable Alpha -L-Arabinofuranosidase from *Thermobacillus Xylanilyticus*'. *Applied and Environmental Microbiology* 66 (4): 1734–36. https://doi.org/10.1128/AEM.66.4.1734-1736.2000.

Demetrius, Lloyd. 1998. 'Role of Enzyme–Substrate Flexibility in Catalytic Activity: An Evolutionary Perspective'. *Journal of Theoretical Biology* 194 (2): 175–94. https://doi.org/10.1006/jtbi.1998.0748.

Demir, Özlem, and Adrian E. Roitberg. 2009. 'Modulation of Catalytic Function by Differential Plasticity of the Active Site: Case Study of Trypanosoma Cruzi Trans - Sialidase and Trypanosoma Rangeli Sialidase †'. *Biochemistry* 48 (15): 3398–3406. https://doi.org/10.1021/bi802230y.

Dilokpimol, Adiphol, Hiroyuki Nakai, Charlotte H. Gotfredsen, Martin J. Baumann, Natsuko Nakai, Maher Abou Hachem, and Birte Svensson. 2011. 'Recombinant Production and Characterisation of Two Related GH5 Endo-β-1,4-Mannanases from Aspergillus Nidulans FGSC A4 Showing Distinctly Different Transglycosylation Capacity'. *Biochimica et Biophysica Acta - Proteins and Proteomics* 1814 (12): 1720–29. https://doi.org/10.1016/j.bbapap.2011.08.003.

Dumon, Claire, Letian Song, Sophie Bozonnet, Régis Fauré, and Michael J O'Donohue. 2012. 'Progress and Future Prospects for Pentose-Specific Biocatalysts in Biorefining'. *Process Biochemistry* 47 (3): 346–57. https://doi.org/https://doi.org/10.1016/j.procbio.2011.06.017.

Dutoit, Raphaël, Maud Delsaute, Laetitia Collet, Corinne Vander Wauven, Dany Van Elder, Renaud Berlemont, Aurore Richel, Moreno Galleni, and Cédric Bauvois. 2019. 'Crystal Structure Determination of *Pseudomonas Stutzeri* A1501 Endoglucanase Cel5A: The Search for a Molecular Basis for Glycosynthesis in GH5_5 Enzymes'. *Acta Crystallographica Section D: Structural Biology* 75: 605–15. https://doi.org/10.1107/S2059798319007113.

Edelman, Jeffrey. 2006. 'The Formation of Oligosaccharides by Enzymic Transglycosylation'. In *Advances in Enzymology and Related Areas of Molecular Biology*, 189–232. https://doi.org/10.1002/9780470122624.ch5.

Ende, Wim Van den, Willem Lammens, André Van Laere, Lindsey Schroeven, and Katrien Le Roy. 2009. 'Donor and Acceptor Substrate Selectivity among Plant Glycoside Hydrolase Family 32 Enzymes'. *FEBS Journal* 276 (20): 5788–98. https://doi.org/10.1111/j.1742-4658.2009.07316.x.

Eppe, Guillaume, Sandy El Bkassiny, and Stéphane P. Vincent. 2015. 'Galactofuranose Biosynthesis: Discovery, Mechanisms and Therapeutic Relevance'. In *Carbohydrates in Drug Design and Discovery*, edited by Jesús Jiménez-Barbero, F. Javier Canada, and Sonsoles Martin-Santamaria, P209–45. Royal Society of Chemistry. https://doi.org/10.1039/9781849739993-FP001.

Euzen, Ronan, Gérald Lopez, Caroline Nugier-Chauvin, Vincent Ferrières, Daniel Plusquellec, Caroline Rémond, and Michael O'Donohue. 2005. 'A Chemoenzymatic Approach for the Synthesis of Unnatural Disaccharides ContainingD-Galacto- OrD-Fucofuranosides'. *European Journal of Organic Chemistry* 2005 (22): 4860–69. https://doi.org/10.1002/ejoc.200500525.

Fauré, Régis, Christophe M. Courtin, Jan A. Delcour, Claire Dumon, Craig B. Faulds,

Geoffrey B. Fincher, Sébastien Fort, et al. 2009. 'A Brief and Informationally Rich Naming System for Oligosaccharide Motifs of Heteroxylans Found in Plant Cell Walls'. *Australian Journal of Chemistry* 62 (6): 533–37. https://doi.org/10.1071/CH08458.

Feng, Hui Yong, Jullien Drone, Lionel Hoffmann, Vinh Tran, Charles Tellier, Claude Rabiller, and Michel Dion. 2005. 'Converting a β-Glycosidase into a β-Transglycosidase by Directed Evolution'. *Journal of Biological Chemistry* 280 (44): 37088–97. https://doi.org/10.1074/jbc.M502873200.

Ferré, Henrik, Anders Broberg, Jens Duus, and Karl K. Thomsen. 2000. 'A Novel Type of Arabinoxylan Arabinofuranohydrolase Isolated from Germinated Barley. Analysis of Substrate Preference and Specificity by Nano-Probe NMR'. *European Journal of Biochemistry* 267 (22): 6633–41. https://doi.org/10.1046/j.1432-1327.2000.01758.x.

Field, Robert A. 2011. 'Challenging Reaction Equilibria'. *Nature Chemical Biology* 7 (10): 658–59. https://doi.org/10.1038/nchembio.668.

Filice, Marco, and Marzia Marciello. 2013. 'Enzymatic Synthesis of Oligosaccharides: A Powerful Tool for a Sweet Challenge'. *Current Organic Chemistry* 17 (7): 701–18. https://doi.org/10.2174/1385272811317070006.

Florindo, Renata N., Valquiria P. Souza, Hemily S. Mutti, Cesar Camilo, Lívia Regina Manzine, Sandro R. Marana, Igor Polikarpov, and Alessandro S. Nascimento. 2018. 'Structural Insights into β-Glucosidase Transglycosylation Based on Biochemical, Structural and Computational Analysis of Two GH1 Enzymes from Trichoderma Harzianum'. *New Biotechnology* 40 (January): 218–27. https://doi.org/10.1016/j.nbt.2017.08.012.

Franceus, Jorick, Nikolas Capra, Tom Desmet, and Andy-Mark W.H. Thunnissen. 2019. 'Structural Comparison of a Promiscuous and a Highly Specific Sucrose 6F-Phosphate Phosphorylase'. *International Journal of Molecular Sciences* 20 (16): 3906. https://doi.org/10.3390/ijms20163906.

Frutuoso, M.A., and S.R. Marana. 2012. 'A Single Amino Acid Residue Determines the Ratio of Hydrolysis to Transglycosylation Catalyzed by β-Glucosidases'. *Protein & Peptide Letters* 20 (1): 102–6. https://doi.org/10.2174/09298665130113.

Garron, Marie-Line, and Bernard Henrissat. 2019. 'The Continuing Expansion of CAZymes and Their Families'. *Current Opinion in Chemical Biology* 53 (December): 82–87. https://doi.org/10.1016/j.cbpa.2019.08.004.

Geronimo, Inacrist, Christina M. Payne, and Mats Sandgren. 2018a. 'The Role of Catalytic Residue p*K*a on the Hydrolysis/Transglycosylation Partition in Family 3 β-Glucosidases'. *Organic and Biomolecular Chemistry* 16 (2): 316–24. https://doi.org/10.1039/c7ob02558k.

———. 2018b. 'Hydrolysis and Transglycosylation Transition States of Glycoside Hydrolase Family 3 β-Glucosidases Differ in Charge and Puckering Conformation'. Research-article. *The Journal of Physical Chemistry B* 122 (41): 9452–59. https://doi.org/10.1021/acs.jpcb.8b07118.

Grootaert, Charlotte, Willy Verstraete, and Tom Van de Wiele. 2007. 'Microbial Metabolism and Prebiotic Potency of Arabinoxylan Oligosaccharides in the Human Intestine'. *Trends in Food Science and Technology*. https://doi.org/10.1016/j.tifs.2006.08.004.

Gruppen, Harry, R A Hoffmann, Felix J.M. M Voragen A G J Kormelink, Alphons G.J. Voragen, J P Kamerling, Johannes F.G. G Vliegenthart, R A Hoffman, et al. 1992.

‘Characterisation by 1H NMR Spectroscopy of Enzymically Derived Oligosaccharides from Alkali-Extractable Wheat-Flour Arabinoxylan’. *Carbohydrate Research* 233 (C): 45–64. https://doi.org/10.1016/S0008-6215(00)90919-4.

Guberman, Mónica, and Peter H. Seeberger. 2019. ‘Automated Glycan Assembly: A Perspective’. *Journal of the American Chemical Society* 141 (14): 5581–92. https://doi.org/10.1021/jacs.9b00638.

Hanniffy, Orla M., Alexander S. Shashkov, Anthony P. Moran, Martina M. Prendergast, Sof’ya N. Senchenkova, Yuriy A. Knirel, and Angela V. Savage. 1999. ‘Chemical Structure of a Polysaccharide from *Campylobacter Jejuni* 176.83 (Serotype O:41) Containing Only Furanose Sugars’. *Carbohydrate Research* 319 (1–4): 124–32. https://doi.org/10.1016/S0008-6215(99)00129-9.

Harholt, Jesper, Jacob Krüger Jensen, Susanne Oxenbøll Sørensen, Caroline Orfila, Markus Pauly, and Henrik Vibe Scheller. 2006. ‘Arabinan deficient 1 Is a Putative Arabinosyltransferase Involved in Biosynthesis of Pectic Arabinan in Arabidopsis’. *Plant Physiology* 140 (1): 49–58. https://doi.org/10.1104/pp.105.072744.

Harvey, Hanjeong, Joseph Bondy-Denomy, Hélène Marquis, Kristina M. Sztanko, Alan R. Davidson, and Lori L. Burrows. 2018. ‘*Pseudomonas Aeruginosa* Defends against Phages through Type IV Pilus Glycosylation’. *Nature Microbiology* 3 (1): 47–52. https://doi.org/10.1038/s41564-017-0061-y.

Helbert, William, Laurent Poulet, Sophie Drouillard, Sophie Mathieu, Mélanie Loiodice, Marie Couturier, Vincent Lombard, et al. 2019. ‘Discovery of Novel Carbohydrate-Active Enzymes through the Rational Exploration of the Protein Sequences Space’. *Proceedings of the National Academy of Sciences* 116 (13): 201815791. https://doi.org/10.1073/pnas.1815791116.

Hijum, S. A. F. T. van, S. Kralj, L. K. Ozimek, L. Dijkhuizen, and I. G. H. van Geel-Schutten. 2006. ‘Structure-Function Relationships of Glucansucrase and Fructansucrase Enzymes from Lactic Acid Bacteria’. *Microbiology and Molecular Biology Reviews* 70 (1): 157–76. https://doi.org/10.1128/MMBR.70.1.157-176.2006.

Hoffmann, Rainer A., Bas R. Leeflang, Martina M.J. de Barse, Johannis P. Kamerling, and Johannes F.G. Vliegenthart. 1991. ‘Characterisation by 1H-n.m.r. Spectroscopy of Oligosaccharides, Derived from Arabinoxylans of White Endosperm of Wheat, That Contain the Elements →4)[α-L-Araf-(1-Ar3)]-β-D-Xylp-(1→ or →4)[α-l-Araf-(1→2)][α-L-Araf-(1→3)]-β-D-Xylp-(1→’. *Carbohydrate Research* 221 (1): 63–81. https://doi.org/10.1016/0008-6215(91)80049-S.

Honda, Yuji, Shinya Fushinobu, Masafumi Hidaka, Takayoshi Wakagi, Hirofumi Shoun, Hajime Taniguchi, and Motomitsu Kitaoka. 2008. ‘Alternative Strategy for Converting an Inverting Glycoside Hydrolase into a Glycosynthase’. *Glycobiology* 18 (4): 325–30. https://doi.org/10.1093/glycob/cwn011.

Honda, Yuji, and Motomitsu Kitaoka. 2006. ‘The First Glycosynthase Derived from an Inverting Glycoside Hydrolase’. *Journal of Biological Chemistry* 281 (3): 1426–31. https://doi.org/10.1074/jbc.M511202200.

Horton, Derek. 2008. ‘The Development of Carbohydrate Chemistry and Biology’. In *Carbohydrate Chemistry, Biology and Medical Applications*, 1–28. Elsevier. https://doi.org/10.1016/B978-0-08-054816-6.00001-X.

Houseknecht, Justin B., and Todd L. Lowary. 2001. ‘Chemistry and Biology of

Arabinofuranosyl- and Galactofuranosyl-Containing Polysaccharides'. *Current Opinion in Chemical Biology*. https://doi.org/10.1016/S1367-5931(01)00265-4.

Houseknecht, Justin B., Cornelis Altona, Christopher M. Hadad, and Todd L. Lowary. 2002. 'Oligofuranosides Containing Conformationally Restricted Residues: Synthesis and Conformational Analysis'. *Journal of Organic Chemistry* 67 (12): 4150–64. https://doi.org/10.1021/jo011127p.

Huang, Xuefei, Lijun Huang, Haisheng Wang, and Xin-Shan Ye. 2004. 'Iterative One-Pot Synthesis of Oligosaccharides'. *Angewandte Chemie International Edition* 43 (39): 5221–24. https://doi.org/10.1002/anie.200460176.

Imamura, Akihiro, and Todd Lowary. 2011. 'Chemical Synthesis of Furanose Glycosides'. *Trends in Glycoscience and Glycotechnology* 23 (131): 134–52. https://doi.org/10.4052/tigg.23.134.

Jahn, Michael, Hongming Chen, Johannes Müllegger, Jennifer Marles, R. Antony J. Warren, and Stephen G. Withers. 2004. 'Thioglycosynthases: Double Mutant Glycosidases That Serve as Scaffolds for Thioglycoside Synthesis'. *Chem. Commun.*, no. 3: 274–75. https://doi.org/10.1039/B313155F.

Jahn, Michael, Jennifer Marles, R. Antony J. Warren, and Stephen G. Withers. 2003. 'Thioglycoligases: Mutant Glycosidases for Thioglycoside Synthesis'. *Angewandte Chemie International Edition* 42 (3): 352–54. https://doi.org/10.1002/anie.200390114.

Jahn, Michael, and Stephen G. Withers. 2003. 'New Approaches to Enzymatic Oligosaccharide Synthesis: Glycosynthases and Thioglycoligases'. *Biocatalysis and Biotransformation* 21 (4–5): 159–66. https://doi.org/10.1080/1024220310001614351.

Jamek, Shariza B., Jan Muschiol, Jesper Holck, Birgitte Zeuner, Peter K. Busk, Jørn D. Mikkelsen, and Anne S. Meyer. 2018. 'Loop Protein Engineering for Improved Transglycosylation Activity of a β-N-Acetylhexosaminidase'. *ChemBioChem* 19 (17): 1858–65. https://doi.org/10.1002/cbic.201800181.

Jankute, Monika, Luke J. Alderwick, Alice R. Moorey, Maju Joe, Sudagar S. Gurcha, Lothar Eggeling, Todd L. Lowary, et al. 2018. 'The Singular Corynebacterium Glutamicum Emb Arabinofuranosyltransferase Polymerises the α(1 → 5) Arabinan Backbone in the Early Stages of Cell Wall Arabinan Biosynthesis'. *The Cell Surface* 2 (June): 38–53. https://doi.org/10.1016/j.tcsw.2018.06.003.

Johansson, Patrik, Harry Brumer, Martin J. Baumann, Åsa M. Kallas, Hongbin Henriksson, Stuart E. Denman, Tuula T. Teeri, and T. Alwyn Jones. 2004. 'Crystal Structures of a Poplar Xyloglucan Endotransglycosylase Reveal Details of Transglycosylation Acceptor Binding'. *The Plant Cell* 16 (4): 874–86. https://doi.org/10.1105/tpc.020065.

Jones, Christopher, Adriane R. Todeschini, Orlando A. Agrellos, José O. Previato, and Lucia Mendonça-Previato. 2004. 'Heterogeneity in the Biosynthesis of Mucin O -Glycans from *Trypanosoma Cruzi Tulahuen* Strain with the Expression of Novel Galactofuranosyl-Containing Oligosaccharides †'. *Biochemistry* 43 (37): 11889–97. https://doi.org/10.1021/bi048942u.

Joseleau, Jean Paul, Gérard Chambat, Michel Vignon, and Fernand Barnoud. 1977. 'Chemical and 13C N.M.R. Studies on Two Arabinans from the Inner Bark of Young Stems of *Rosa Glauca*'. *Carbohydrate Research* 58 (1): 165–75. https://doi.org/10.1016/S0008-6215(00)83412-6.

Kaji, Akira, and Tsuneo Saheki. 1975. 'Endo-Arabinanase from Bacillus Subtilis F-11'.

Biochimica et Biophysica Acta (BBA) - Enzymology 410 (2): 354–60. https://doi.org/10.1016/0005-2744(75)90237-5.

Kashiwagi, Gustavo A., Carmen R. Cori, Rosa M. de Lederkremer, and Carola Gallo-Rodriguez. 2019. 'Synthesis of the Hexasaccharide from *Trypanosoma Cruzi* Mucins with the Galp(1 → 2)Galf Unit Constructed with a Superarmed Thiogalactopyranosyl Donor'. *Carbohydrate Research* 482 (August): 107734. https://doi.org/10.1016/j.carres.2019.06.013.

Kawabata, Yasuyuki, Satoshi Kaneko, Isao Kusakabe, and Yasuo Gama. 1995. 'Synthesis of Regioisomeric Methyl α-L-Arabinofuranobiosides'. *Carbohydrate Research* 267 (1): 39–47. https://doi.org/10.1016/0008-6215(94)00290-V.

Kawamura, Shunsuke, Masashi Eto, Taiji Moto, Shinji Ikemizu, Tomohiro Araki, and Takao Torikata. 2004. 'Functional and Structural Effects of Mutagenic Replacement of Asn37 at Subsite F on the Lysozyme-Catalyzed Reaction'. *Bioscience, Biotechnology, and Biochemistry* 68 (3): 593–601. https://doi.org/10.1271/bbb.68.593.

Kelly, Ronan M., Hans Leemhuis, Henriëtte J. Rozeboom, Niels van Oosterwijk, Bauke W. Dijkstra, and Lubbert Dijkhuizen. 2008. 'Elimination of Competing Hydrolysis and Coupling Side Reactions of a Cyclodextrin Glucanotransferase by Directed Evolution'. *Biochemical Journal* 413 (3): 517–25. https://doi.org/10.1042/BJ20080353.

Kiessling, Laura L., and Rebecca A. Splain. 2010. 'Chemical Approaches to Glycobiology'. *Annual Review of Biochemistry* 79 (1): 619–53. https://doi.org/10.1146/annurev.biochem.77.070606.100917.

Kim, Myung Il, Hong Suk Kim, Jin Jung, and Sangkee Rhee. 2008. 'Crystal Structures and Mutagenesis of Sucrose Hydrolase from Xanthomonas Axonopodis Pv. Glycines: Insight into the Exclusively Hydrolytic Amylosucrase Fold'. *Journal of Molecular Biology* 380 (4): 636–47. https://doi.org/10.1016/j.jmb.2008.05.046.

Kinnaert, Christine, Mathilde Daugaard, Faranak Nami, and Mads H. Clausen. 2017. 'Chemical Synthesis of Oligosaccharides Related to the Cell Walls of Plants and Algae'. *Chemical Reviews* 117 (17): 11337–405. https://doi.org/10.1021/acs.chemrev.7b00162.

Koenigs, Wilhelm, and Eduard Knorr. 1901. 'Ueber Einige Derivate Des Traubenzuckers Und Der Galactose'. *Berichte Der Deutschen Chemischen Gesellschaft* 34 (1): 957–81. https://doi.org/10.1002/cber.190103401162.

Komachi, Yuji, Shintaro Hatakeyama, Haruka Motomatsu, Taiki Futagami, Karina Kizjakina, Pablo Sobrado, Keisuke Ekino, et al. 2013. 'GfsA Encodes a Novel Galactofuranosyltransferase Involved in Biosynthesis of Galactofuranose Antigen of O -Glycan in Aspergillus Nidulans and Aspergillus Fumigatus'. *Molecular Microbiology* 90 (5): 1054–73. https://doi.org/10.1111/mmi.12416.

Koné, Fankroma M.T. T, Mickaël Le Béchec, Jean-Pierre Pierre Sine, Michel Dion, and Charles Tellier. 2009. 'Digital Screening Methodology for the Directed Evolution of Transglycosidases.' *Protein Engineering, Design & Selection : PEDS* 22 (1): 37–44. https://doi.org/10.1093/protein/gzn065.

Konishi, Teruko, Hiroshi Ono, Mayumi Ohnishi-Kameyama, Satoshi Kaneko, and Tadashi Ishii. 2006. 'Identification of a Mung Bean Arabinofuranosyltransferase That Transfers Arabinofuranosyl Residues onto (1,5)-Linked α-L-Arabino-Oligosaccharides'. *Plant Physiology* 141 (3): 1098–1105. https://doi.org/10.1104/pp.106.080309.

Koshland, D. E. 1953. 'Stereochemistry and the Mechanism of Enzymatic Reactions'.

Biological Reviews 28 (4): 416–36. https://doi.org/10.1111/j.1469-185X.1953.tb01386.x.

Kotake, Toshihisa, Yukiko Yamanashi, Chiemi Imaizumi, and Yoichi Tsumuraya. 2016. 'Metabolism of L-Arabinose in Plants'. *Journal of Plant Research* 129 (5): 781–92. https://doi.org/10.1007/s10265-016-0834-z.

Krasnova, Larissa, and Chi-Huey Wong. 2016. *Understanding the Chemistry and Biology of Glycosylation with Glycan Synthesis*. *Annual Review of Biochemistry*. Vol. 85. https://doi.org/10.1146/annurev-biochem-060614-034420.

Kus, J. V., J. Kelly, L. Tessier, H. Harvey, D. G. Cvitkovitch, and L. L. Burrows. 2008. 'Modification of Pseudomonas Aeruginosa Pa5196 Type IV Pilins at Multiple Sites with D-Araf by a Novel GT-C Family Arabinosyltransferase, TfpW'. *Journal of Bacteriology* 190 (22): 7464–78. https://doi.org/10.1128/JB.01075-08.

Lafraya, Álvaro, Julia Sanz-Aparicio, Julio Polaina, and Julia Marín-Navarro. 2011. 'Fructo-Oligosaccharide Synthesis by Mutant Versions of *Saccharomyces Cerevisiae* Invertase'. *Applied and Environmental Microbiology* 77 (17): 6148–57. https://doi.org/10.1128/AEM.05032-11.

Laine, Roger A. 1994. 'Invited Commentary: A Calculation of All Possible Oligosaccharide Isomers Both Branched and Linear Yields 1.05×10^{12} Structures for a Reducing Hexasaccharide: The Isomer Barrier to Development of Single-Method Saccharide Sequencing or Synthesi'. *Glycobiology* 4 (6): 759–67. https://doi.org/10.1093/glycob/4.6.759.

Lairson, L.L., B. Henrissat, G.J. Davies, and S.G. Withers. 2008. 'Glycosyltransferases: Structures, Functions, and Mechanisms'. *Annual Review of Biochemistry* 77 (1): 521–55. https://doi.org/10.1146/annurev.biochem.76.061005.092322.

Lammens, Willem, Katrien Le Roy, Lindsey Schroeven, André Van Laere, Anja Rabijns, and Wim Van den Ende. 2009. 'Structural Insights into Glycoside Hydrolase Family 32 and 68 Enzymes: Functional Implications'. *Journal of Experimental Botany* 60 (3): 727–40. https://doi.org/10.1093/jxb/ern333.

Larsbrink, Johan, Atsushi Izumi, Glyn R. Hemsworth, Gideon J. Davies, and Harry Brumer. 2012. 'Structural Enzymology of Cellvibrio Japonicus Agd31B Protein Reveals α-Transglucosylase Activity in Glycoside Hydrolase Family 31'. *Journal of Biological Chemistry* 287 (52): 43288–99. https://doi.org/10.1074/jbc.M112.416511.

Latge, Jean-Paul. 2009. 'Galactofuranose Containing Molecules in *Aspergillus Fumigatus*'. *Medical Mycology* 47 (s1): S104–9. https://doi.org/10.1080/13693780802258832.

Lawson, Sherry L., Warren W. Wakarchuk, and Stephen G. Withers. 1996. 'Effects of Both Shortening and Lengthening the Active Site Nucleophile of *Bacillus Circulans* Xylanase on Catalytic Activity'. *Biochemistry* 35 (31): 10110–18. https://doi.org/10.1021/bi960586v.

Leal, Teresa Fontes, and Isabel de Sá-Nogueira. 2004. 'Purification, Characterization and Functional Analysis of an Endo-Arabinanase (AbnA) from *Bacillus Subtilis*'. *FEMS Microbiology Letters* 241 (1): 41–48. https://doi.org/10.1016/j.femsle.2004.10.003.

Lee, Richard E., Patrick J. Brennan, and Gurdyal S. Besra. 1997. 'Mycobacterial Arabinan Biosynthesis: The Use of Synthetic Arabinoside Acceptors in the Development of an Arabinosyl Transfer Assay'. *Glycobiology* 7 (8): 1121–28. https://doi.org/10.1093/glycob/7.8.1121.

Leemhuis, Hans, Bauke W. Dijkstra, and Lubbert Dijkhuizen. 2002. 'Mutations Converting

Cyclodextrin Glycosyltransferase from a Transglycosylase into a Starch Hydrolase'. *FEBS Letters* 514 (2–3): 189–92. https://doi.org/10.1016/S0014-5793(02)02362-1.

Leitão, Eduardo A., Vera C.B. Bittencourt, Rosa M.T. Haido, Ana P. Valente, Jasna Peter-Katalinic, Matthias Letzel, Lauro M. de Souza, and Eliana Barreto-Bergter. 2003. 'β-Galactofuranose-Containing O-Linked Oligosaccharides Present in the Cell Wall Peptidogalactomannan of *Aspergillus Fumigatus* Contain Immunodominant Epitopes'. *Glycobiology* 13 (10): 681–92. https://doi.org/10.1093/glycob/cwg089.

Levengood, Matthew R., Rebecca A. Splain, and Laura L. Kiessling. 2011. 'Monitoring Processivity and Length Control of a Carbohydrate Polymerase'. *Journal of the American Chemical Society* 133 (32): 12758–66. https://doi.org/10.1021/ja204448t.

Light, Samuel H., Laty A. Cahoon, Kiran V. Mahasenan, Mijoon Lee, Bill Boggess, Andrei S. Halavaty, Shahriar Mobashery, Nancy E. Freitag, and Wayne F. Anderson. 2017. 'Transferase Versus Hydrolase: The Role of Conformational Flexibility in Reaction Specificity'. *Structure* 25 (2): 1–10. https://doi.org/10.1016/j.str.2016.12.007.

Lipničanová, Sabina, Daniela Chmelová, Miroslav Ondrejovič, Vladimír Frecer, and Stanislav Miertuš. 2020. 'Diversity of Sialidases Found in the Human Body – A Review'. *International Journal of Biological Macromolecules* 148 (April): 857–68. https://doi.org/10.1016/j.ijbiomac.2020.01.123.

Lombard, Vincent, Hemalatha Golaconda Ramulu, Elodie Drula, Pedro M Coutinho, and Bernard Henrissat. 2014. 'The Carbohydrate-Active Enzymes Database (CAZy) in 2013'. *Nucleic Acids Research* 42 (D1): D490–95. https://doi.org/10.1093/nar/gkt1178.

Lowary, Todd L. 2003. 'Synthesis and Conformational Analysis of Arabinofuranosides, Galactofuranosides and Fructofuranosides'. *Current Opinion in Chemical Biology* 7 (6): 749–56. https://doi.org/10.1016/j.cbpa.2003.10.005.

Lü, Yang, Haitao Yang, Hongyan Hu, Ying Wang, Zihe Rao, and Cheng Jin. 2009. 'Mutation of Trp137 to Glutamate Completely Removes Transglycosyl Activity Associated with the *Aspergillus Fumigatus* AfChiB1'. *Glycoconjugate Journal* 26 (5): 525–34. https://doi.org/10.1007/s10719-008-9203-z.

Luang, Sukanya, Jung Il Cho, Bancha Mahong, Rodjana Opassiri, Takashi Akiyama, Kannika Phasai, Juthamath Komvongsa, et al. 2013. 'Rice Os9BGlu31 Is a Transglucosidase with the Capacity to Equilibrate Phenylpropanoid, Flavonoid, and Phytohormone Glycoconjugates'. *Journal of Biological Chemistry* 288 (14): 10111–23. https://doi.org/10.1074/jbc.M112.423533.

Lundemo, Pontus, Patrick Adlercreutz, and Eva Nordberg Karlsson. 2013. 'Improved Transferase/Hydrolase Ratio through Rational Design of a Family 1 β-Glucosidase from *Thermotoga Neapolitana*'. *Applied and Environmental Microbiology* 79 (11): 3400–3405. https://doi.org/10.1128/AEM.00359-13.

Maarel, Marc J.E.C. van der, and Hans Leemhuis. 2013. 'Starch Modification with Microbial Alpha-Glucanotransferase Enzymes'. *Carbohydrate Polymers* 93 (1): 116–21. https://doi.org/10.1016/j.carbpol.2012.01.065.

Macdonald, Spencer S., Zachary Armstrong, Connor Morgan-Lang, Magdalena Osowiecka, Kyle Robinson, Steven J. Hallam, and Stephen G. Withers. 2019. 'Development and Application of a High-Throughput Functional Metagenomic Screen for Glycoside Phosphorylases'. *Cell Chemical Biology* 26 (7): 1001-1012.e5. https://doi.org/10.1016/j.chembiol.2019.03.017.

Macdonald, Spencer S., Ankoor Patel, Veronica L. C. Larmour, Connor Morgan-Lang, Steven J. Hallam, Brian L. Mark, and Stephen G. Withers. 2018. 'Structural and Mechanistic Analysis of a β-Glycoside Phosphorylase Identified by Screening a Metagenomic Library'. *Journal of Biological Chemistry* 293 (9): 3451–67. https://doi.org/10.1074/jbc.RA117.000948.

Mackenzie, Lloyd F., Qingping Wang, R. Antony J. Warren, and Stephen G. Withers. 1998. 'Glycosynthases: Mutant Glycosidases for Oligosaccharide Synthesis'. *Journal of the American Chemical Society* 120 (22): 5583–84. https://doi.org/10.1021/ja980833d.

Madhuprakash, Jogi, Avinash Singh, Sanjit Kumar, Mau Sinha, Punit Kaur, Sujata Sharma, Appa R. Podile, and Tej P. Singh. 2013. 'Structure of Chitinase D from Serratia Proteamaculans Reveals the Structural Basis of Its Dual Action of Hydrolysis and Transglycosylation'. *International Journal of Biochemistry and Molecular Biology* 4 (4): 166–78.

Malet, Carles, and Antoni Planas. 1998. 'From β-Glucanase to β-Glucansynthase: Glycosyl Transfer to α-Glycosyl Fluorides Catalyzed by a Mutant Endoglucanase Lacking Its Catalytic Nucleophile'. *FEBS Letters* 440 (1–2): 208–12. https://doi.org/10.1016/S0014-5793(98)01448-3.

Marino, Carla, and Luciana Baldoni. 2014. 'Synthesis of D-Galactofuranose-Containing Molecules: Design of Galactofuranosyl Acceptors'. *ChemBioChem* 15 (2): 188–204. https://doi.org/10.1002/cbic.201300638.

Marino, Carla, Adriana Rinflerch, and Rosa M. De Lederkremer. 2017. 'Galactofuranose Antigens, a Target for Diagnosis of Fungal Infections in Humans'. *Future Science OA* 3 (3): FSO199. https://doi.org/10.4155/fsoa-2017-0030.

Mark, Pekka, Martin J. Baumann, Jens M. Eklöf, Fredrika Gullfot, Gurvan Michel, Åsa M. Kallas, Tuula T. Teeri, Harry Brumer, and Mirjam Czjzek. 2009. 'Analysis of Nasturtium Tm NXG1 Complexes by Crystallography and Molecular Dynamics Provides Detailed Insight into Substrate Recognition by Family GH16 Xyloglucan Endo -Transglycosylases and Endo -Hydrolases'. *Proteins: Structure, Function, and Bioinformatics* 75 (4): 820–36. https://doi.org/10.1002/prot.22291.

Mathew, Sindhu, Anna Aronsson, Eva Nordberg Karlsson, and Patrick Adlercreutz. 2018. 'Xylo- and Arabinoxylooligosaccharides from Wheat Bran by Endoxylanases, Utilisation by Probiotic Bacteria, and Structural Studies of the Enzymes'. *Applied Microbiology and Biotechnology* 102 (7): 3105–20. https://doi.org/10.1007/s00253-018-8823-x.

Matsuba, Yuki, Nobuhiro Sasaki, Masayuki Tera, Masachika Okamura, Yutaka Abe, Emi Okamoto, Haruka Nakamura, et al. 2010. 'A Novel Glucosylation Reaction on Anthocyanins Catalyzed by Acyl-Glucose–Dependent Glucosyltransferase in the Petals of Carnation and Delphinium'. *The Plant Cell* 22 (10): 3374–89. https://doi.org/10.1105/tpc.110.077487.

Matsunaga, Emiko, Yujiro Higuchi, Kazuki Mori, Kosuke Tashiro, and Kaoru Takegawa. 2017. 'Draft Genome Sequence of Streptomyces Sp. JHA26, a Strain That Harbors a PA14 Domain Containing β-D-Galactofuranosidase'. *Genome Announcements* 5 (15): 3–4. https://doi.org/10.1128/genomeA.00190-17.

Matsunaga, Emiko, Yujiro Higuchi, Kazuki Mori, Nao Yairo, Takuji Oka, Saki Shinozuka, Kosuke Tashiro, Minoru Izumi, Satoru Kuhara, and Kaoru Takegawa. 2015. 'Identification and Characterization of a Novel Galactofuranose-Specific β-D-Galactofuranosidase from *Streptomyces* Species'. Edited by Olaf Kniemeyer. *PLOS ONE*

10 (9): e0137230. https://doi.org/10.1371/journal.pone.0137230.

Matsunaga, Emiko, Yujiro Higuchi, Kazuki Mori, Nao Yairo, Saki Toyota, Takuji Oka, Kosuke Tashiro, and Kaoru Takegawa. 2017. 'Characterization of a PA14 Domain-Containing Galactofuranose-Specific β-D-Galactofuranosidase from *Streptomyces* Sp.' *Bioscience, Biotechnology, and Biochemistry* 81 (7): 1314–19. https://doi.org/10.1080/09168451.2017.1300518.

Mayer, Christoph, David L. Zechel, Stephen P. Reid, R. Antony J. Warren, and Stephen G. Withers. 2000. 'The E358S Mutant of Agrobacterium Sp. β-Glucosidase Is a Greatly Improved Glycosynthase'. *FEBS Letters* 466 (1): 40–44. https://doi.org/10.1016/S0014-5793(99)01751-2.

McCleary, Barry V., Vincent A. McKie, Anna Draga, Edward Rooney, David Mangan, and Jennifer Larkin. 2015. 'Hydrolysis of Wheat Flour Arabinoxylan, Acid-Debranched Wheat Flour Arabinoxylan and Arabino-Xylo-Oligosaccharides by β-Xylanase, α-l-Arabinofuranosidase and β-Xylosidase'. *Carbohydrate Research* 407 (April): 79–96. https://doi.org/10.1016/j.carres.2015.01.017.

Mestrom, Luuk, Marta Przypis, Daria Kowalczykiewicz, André Pollender, Antje Kumpf, Stefan R. Marsden, Isabel Bento, et al. 2019. *Leloir Glycosyltransferases in Applied Biocatalysis: A Multidisciplinary Approach*. *International Journal of Molecular Sciences*. Vol. 20. https://doi.org/10.3390/ijms20215263.

Miletti, Luiz Claudio, Karina Mariño, Carla Marino, Walter Colli, Maria Julia Manso Alves, and Rosa M.de Lederkremer. 2003. 'Evidence for Exo β-D-Galactofuranosidase in Trypanosoma Cruzi'. *Molecular and Biochemical Parasitology* 127 (1): 85–88. https://doi.org/10.1016/S0166-6851(02)00307-9.

Möhlig, M., C. Koebnick, M. O. Weickert, W. Lueder, B. Otto, J. Steiniger, M. Twilfert, F. Meuser, A. F. Pfeiffer, and H. J. Zunft. 2005. 'Arabinoxylan-Enriched Meal Increases Serum Ghrelin Levels in Healthy Humans'. *Hormone and Metabolic Research* 37 (5): 303–8. https://doi.org/10.1055/s-2005-861474.

Monsan, Pierre, Magali Remaud-Siméon, and Isabelle André. 2010. 'Transglucosidases as Efficient Tools for Oligosaccharide and Glucoconjugate Synthesis'. *Current Opinion in Microbiology*. https://doi.org/10.1016/j.mib.2010.03.002.

Moracci, Marco, Antonio Trincone, Giuseppe Perugino, Maria Ciaramella, and Mosé Rossi. 1998. 'Restoration of the Activity of Active-Site Mutants of the Hyperthermophilic β-Glycosidase from Sulfolobus Solfataricus: Dependence of the Mechanism on the Action of External Nucleophiles'. *Biochemistry* 37 (49): 17262–70. https://doi.org/10.1021/bi981855f.

Moulis, Claire, Isabelle André, and Magali Remaud-Simeon. 2016. 'GH13 Amylosucrases and GH70 Branching Sucrases, Atypical Enzymes in Their Respective Families'. *Cellular and Molecular Life Sciences* 73 (14): 2661–79. https://doi.org/10.1007/s00018-016-2244-8.

Muthana, Saddam, Hongzhi Cao, and Xi Chen. 2009. 'Recent Progress in Chemical and Chemoenzymatic Synthesis of Carbohydrates'. *Current Opinion in Chemical Biology*. https://doi.org/10.1016/j.cbpa.2009.09.013.

Nakai, Hiroyuki, Motomitsu Kitaoka, Birte Svensson, and Ken'ichi Ohtsubo. 2013. 'Recent Development of Phosphorylases Possessing Large Potential for Oligosaccharide Synthesis'. *Current Opinion in Chemical Biology*.

https://doi.org/10.1016/j.cbpa.2013.01.006.

Nidetzky, Bernd, Alexander Gutmann, and Chao Zhong. 2018. 'Leloir Glycosyltransferases as Biocatalysts for Chemical Production'. *ACS Catalysis* 8 (7): 6283–6300. https://doi.org/10.1021/acscatal.8b00710.

Nielsen, Michael Martin, and Christian Marcus Pedersen. 2018. 'Catalytic Glycosylations in Oligosaccharide Synthesis'. Review-article. *Chemical Reviews* 118 (17): 8285–8358. https://doi.org/10.1021/acs.chemrev.8b00144.

Nielsen, Ruby Illum. 2002. 'Microbial Xyloglucan Endotransglycosylase (XET)'. Google Patents.

Ogawa-Ohnishi, Mari, Wataru Matsushita, and Yoshikatsu Matsubayashi. 2013. 'Identification of Three Hydroxyproline O-Arabinosyltransferases in Arabidopsis Thaliana'. *Nature Chemical Biology* 9 (11): 726–30. https://doi.org/10.1038/nchembio.1351.

Olvera, Clarita, Sara Centeno-Leija, Paulina Ruiz-Leyva, and Agustín López-Munguía. 2012. 'Design of Chimeric Levansucrases with Improved Transglycosylation Activity'. *Applied and Environmental Microbiology* 78 (6): 1820–25. https://doi.org/10.1128/AEM.07222-11.

Ortiz-Soto, Maria Elena, and Jürgen Seibel. 2016. 'Expression of Functional Human Sialyltransferases ST3Gal1 and ST6Gal1 in *Escherichia Coli*'. Edited by Fernando Rodrigues-Lima. *PLOS ONE* 11 (5): e0155410. https://doi.org/10.1371/journal.pone.0155410.

Paës, Gabriel, Lars K. Skov, Michael J. O'Donohue, Caroline Rémond, Jette S. Kastrup, Michael Gajhede, and Osman Mirza. 2008. 'The Structure of the Complex between a Branched Pentasaccharide and Thermobacillus Xylanilyticus GH-51 Arabinofuranosidase Reveals Xylan-Binding Determinants and Induced Fit.' *Biochemistry* 47 (28): 7441–51. https://doi.org/10.1021/bi800424e.

Pan, F., M. Jackson, Y. Ma, and M. McNeil. 2001. 'Cell Wall Core Galactofuran Synthesis Is Essential for Growth of Mycobacteria'. *Journal of Bacteriology* 183 (13): 3991–98. https://doi.org/10.1128/JB.183.13.3991-3998.2001.

Panza, Matteo, Salvatore G. Pistorio, Keith J. Stine, Alexei V. Demchenko, Salvatore G. Pistorio, Keith J. Stine, and Alexei V. Demchenko. 2018. 'Automated Chemical Oligosaccharide Synthesis: Novel Approach to Traditional Challenges'. Review-article. *Chemical Reviews* 118 (17): 8105–50. https://doi.org/10.1021/acs.chemrev.8b00051.

Paris, Gastón, Laura Ratier, María Fernanda Amaya, Tong Nguyen, Pedro M. Alzari, and Alberto Carlos C. Frasch. 2005. 'A Sialidase Mutant Displaying Trans-Sialidase Activity'. *Journal of Molecular Biology* 345 (4): 923–34. https://doi.org/10.1016/j.jmb.2004.09.031.

Park, Tae Joon, Michel Weïwer, Xuejun Yuan, Sultan N. Baytas, Eva M. Munoz, Saravanababu Murugesan, and Robert J. Linhardt. 2007. 'Glycosylation in Room Temperature Ionic Liquid Using Unprotected and Unactivated Donors'. *Carbohydrate Research*. https://doi.org/10.1016/j.carres.2006.11.022.

Pastell, Helena, Päivi Tuomainen, Liisa Virkki, and Maija Tenkanen. 2008. 'Step-Wise Enzymatic Preparation and Structural Characterization of Singly and Doubly Substituted Arabinoxylo-Oligosaccharides with Non-Reducing End Terminal Branches.' *Carbohydrate Research* 343 (18): 3049–57. https://doi.org/10.1016/j.carres.2008.09.013.

Pastell, Helena, Peter Westermann, Anne S. Meyer, Päivi Tuomainen, and Maija Tenkanen. 2009. 'In Vitro Fermentation of Arabinoxylan-Derived Carbohydrates by Bifidobacteria and Mixed Fecal Microbiota'. *Journal of Agricultural and Food Chemistry* 57 (18): 8598–8606. https://doi.org/10.1021/jf901397b.

Pavic, Quentin, Aline Pillot, Olivier Tasseau, Laurent Legentil, and Sylvain Tranchimand. 2019. 'Improvement of the Versatility of an Arabinofuranosidase against Galactofuranose for the Synthesis of Galactofuranoconjugates'. *Organic & Biomolecular Chemistry* 17 (28): 6799–6808. https://doi.org/10.1039/C9OB01162E.

Pedersen, L L, and S J Turco. 2003. 'Galactofuranose Metabolism: A Potential Target for Antimicrobial Chemotherapy'. *Cellular and Molecular Life Sciences (CMLS)* 60 (2): 259–66. https://doi.org/10.1007/s000180300021.

Peltier, Pauline, Ronan Euzen, Richard Daniellou, Caroline Nugier-Chauvin, and Vincent Ferrières. 2008. 'Recent Knowledge and Innovations Related to Hexofuranosides: Structure, Synthesis and Applications'. *Carbohydrate Research* 343 (12): 1897–1923. https://doi.org/10.1016/j.carres.2008.02.010.

Pierdominici-Sottile, Gustavo, Nicole A. Horenstein, and Adrian E. Roitberg. 2011. 'Free Energy Study of the Catalytic Mechanism of Trypanosoma Cruzi Trans -Sialidase. From the Michaelis Complex to the Covalent Intermediate'. *Biochemistry* 50 (46): 10150–58. https://doi.org/10.1021/bi2009618.

Pierdominici-Sottile, Gustavo, Juliana Palma, and Adrian E. Roitberg. 2014. 'Free-Energy Computations Identify the Mutations Required to Confer Trans-Sialidase Activity into *Trypanosoma Rangeli* Sialidase'. *Proteins: Structure, Function, and Bioinformatics* 82 (3): 424–35. https://doi.org/10.1002/prot.24408.

Poria, Vikram, Jitendra Kumar Saini, Surender Singh, Lata Nain, and Ramesh Chander Kuhad. 2020. 'Arabinofuranosidases: Characteristics, Microbial Production, and Potential in Waste Valorization and Industrial Applications'. *Bioresource Technology*, 123019. https://doi.org/10.1016/j.biortech.2020.123019.

Potocki De Montalk, Gabrielle, Magali Remaud-Simeon, René Marc Willemot, Patricia Sarçabal, Véronique Planchot, and Pierre Monsan. 2000. 'Amylosucrase from *Neisseria Polysaccharea*: Novel Catalytic Properties'. *FEBS Letters* 471 (2–3): 219–23. https://doi.org/10.1016/S0014-5793(00)01406-X.

Poulin, Myles B., and Todd L. Lowary. 2010. 'Methods to Study the Biosynthesis of Bacterial Furanosides'. In *Methods in Enzymology*, 1st ed., 478:389–411. Elsevier Inc. https://doi.org/10.1016/S0076-6879(10)78019-8.

Poulin, Myles B., Ruokun Zhou, and Todd L. Lowary. 2012. 'Synthetic UDP-Galactofuranose Analogs Reveal Critical Enzyme-Substrate Interactions in GlfT2-Catalyzed Mycobacterial Galactan Assembly'. *Organic and Biomolecular Chemistry* 10 (20): 4074–87. https://doi.org/10.1039/c2ob25159k.

Pravda, Lukáš, Karel Berka, Radka Svobodová Vařeková, David Sehnal, Pavel Banáš, Roman A Laskowski, Jaroslav Koča, and Michal Otyepka. 2014. 'Anatomy of Enzyme Channels'. *BMC Bioinformatics* 15 (1): 379. https://doi.org/10.1186/s12859-014-0379-x.

Prieto, A., M. Bernabé, and J. A. Leal. 1995. 'Isolation, Purification and Chemical Characterization of Alkali-Extractable Polysaccharides from the Cell Walls of Talaromyces Species'. *Mycological Research*. https://doi.org/10.1016/S0953-7562(09)80318-3.

Puchart, Vladimír. 2015. ‘Glycoside Phosphorylases: Structure, Catalytic Properties and Biotechnological Potential’. *Biotechnology Advances*. https://doi.org/10.1016/j.biotechadv.2015.02.002.

Ramli, Nahrowi, Michiko Fujinaga, Mitsuaki Tabuchi, Kaoru Takegawa, and Shojiro Iwahara. 1995. ‘Isolation and Characterization of a Novel Endo-β-Galactofuranosidase from *Bacillus* Sp.’ *Bioscience, Biotechnology, and Biochemistry* 59 (10): 1856–60. https://doi.org/10.1271/bbb.59.1856.

Rebuffet, Etienne, Agnès Groisillier, Andrew Thompson, Alexandra Jeudy, Tristan Barbeyron, Mirjam Czjzek, and Gurvan Michel. 2011. ‘Discovery and Structural Characterization of a Novel Glycosidase Family of Marine Origin’. *Environmental Microbiology* 13 (5): 1253–70. https://doi.org/10.1111/j.1462-2920.2011.02426.x.

Rémond, C., I. Boukari, G. Chambat, and M. O’Donohue. 2008. ‘Action of a GH 51 α-L-Arabinofuranosidase on Wheat-Derived Arabinoxylans and Arabino-Xylooligosaccharides’. *Carbohydrate Polymers* 72 (3): 424–30. https://doi.org/10.1016/j.carbpol.2007.09.008.

Rémond, Caroline, Richard Plantier-Royon, Nathalie Aubry, Emmanuel Maes, Christophe Bliard, and Michael J. O’Donohue. 2004. ‘Synthesis of Pentose-Containing Disaccharides Using a Thermostable α-L-Arabinofuranosidase’. *Carbohydrate Research* 339 (11): 2019–25. https://doi.org/10.1016/j.carres.2004.04.017.

Rémond, Caroline, Richard Plantier-Royon, Nathalie Aubrya and Michael J. O’Donohue. 2005. ‘An Original Chemoenzymatic Route for the Synthesis of β-D-Galactofuranosides Using an α-l-Arabinofuranosidase’. *Carbohydrate Research* 340 (4): 637–44. https://doi.org/10.1016/j.carres.2005.01.016.

Ribeirão, Marcelo, Vera Lucia Pereira-Chioccola, Daniel Eichinger, Mauricio M. Rodrigues, and Sergio Schenkman. 1997. ‘Temperature Differences for Trans -Glycosylation and Hydrolysis Reaction Reveal an Acceptor Binding Site in the Catalytic Mechanism of Trypanosoma Cruzi Trans -Sialidase’. *Glycobiology* 7 (8): 1237–46. https://doi.org/10.1093/glycob/7.8.1237.

Richards, Michele R., and Todd L. Lowary. 2009. ‘Chemistry and Biology of Galactofuranose-Containing Polysaccharides’. *ChemBioChem* 10 (12): 1920–38. https://doi.org/10.1002/cbic.200900208.

Rietschel-Berst, M., N.H. Jentoft, P.D. Rick, C. Pletcher, F Fang, and J.E. Gander. 1977. ‘Extracellular Exo-β-Galactofuranosidase from *Penicillium Charlesii*: Isolation, Purification, and Properties’. *Journal of Biological Chemistry* 252 (10): 3219–26.

Rivera, Manuel Heriberto, Agustín López-Munguía, Xavier Soberón, and Gloria Saab-Rincón. 2003. ‘α-Amylase from *Bacillus Licheniformis* Mutants near to the Catalytic Site: Effects on Hydrolytic and Transglycosylation Activity’. *Protein Engineering* 16 (7): 505–14. https://doi.org/10.1093/protein/gzg060.

Rose, Natisha L., Gladys C. Completo, Shuang Jun Lin, Michael McNeil, Monica M. Palcic, and Todd L. Lowary. 2006. ‘Expression, Purification, and Characterization of a Galactofuranosyltransferase Involved in *Mycobacterium Tuberculosis* Arabinogalactan Biosynthesis’. *Journal of the American Chemical Society* 128 (20): 6721–29. https://doi.org/10.1021/ja058254d.

Rosengren, Anna, Per Hägglund, Lars Anderson, Patricia Pavon-Orozco, Ragna Peterson-Wulff, Wim Nerinckx, and Henrik Stålbrand. 2012. ‘The Role of Subsite +2 of the

Trichoderma Reesei β-Mannanase TrMan5A in Hydrolysis and Transglycosylation'. *Biocatalysis and Biotransformation* 30 (3): 338–52. https://doi.org/10.3109/10242422.2012.674726.

Saburi, Wataru, Momoko Kobayashi, Haruhide Mori, Masayuki Okuyama, and Atsuo Kimura. 2013. 'Replacement of the Catalytic Nucleophile Aspartyl Residue of Dextran Glucosidase by Cysteine Sulfinate Enhances Transglycosylation Activity'. *Journal of Biological Chemistry* 288 (44): 31670–77. https://doi.org/10.1074/jbc.M113.491449.

Sainz-Polo, M. Angela, Mercedes Ramírez-Escudero, Alvaro Lafraya, Beatriz González, Julia Marín-Navarro, Julio Polaina, and Julia Sanz-Aparicio. 2013. 'Three-Dimensional Structure of *Saccharomyces* Invertase: Role of a Non-Catalytic Domain in Oligomerization and Substrate Specificity'. *Journal of Biological Chemistry* 288 (14): 9755–66. https://doi.org/10.1074/jbc.M112.446435.

Sanchez, J. I., M. Marzorati, C. Grootaert, M. Baran, V. Van Craeyveld, C. M. Courtin, W. F. Broekaert, J. A. Delcour, W. Verstraete, and T. Van de Wiele. 2009. 'Arabinoxylan-Oligosaccharides (AXOS) Affect the Protein/Carbohydrate Fermentation Balance and Microbial Population Dynamics of the Simulator of Human Intestinal Microbial Ecosystem'. *Microbial Biotechnology* 2 (1): 101–13. https://doi.org/10.1111/j.1751-7915.2008.00064.x.

Saura-Valls, Marc, Régis Fauré, Harry Brumer, Tuula T. Teeri, Sylvain Cottaz, Hugues Driguez, and Antoni Planas. 2008. 'Active-Site Mapping of a Populus Xyloglucan Endo-Transglycosylase with a Library of Xylogluco-Oligosaccharides'. *Journal of Biological Chemistry* 283 (32): 21853–63. https://doi.org/10.1074/jbc.M803058200.

Schmölzer, Katharina, Martin Lemmerer, and Bernd Nidetzky. 2018. 'Glycosyltransferase Cascades Made Fit for Chemical Production: Integrated Biocatalytic Process for the Natural Polyphenol C-Glucoside Nothofagin'. *Biotechnology and Bioengineering* 115 (3): 545–56. https://doi.org/10.1002/bit.26491.

Schröder, Roswitha, TeresaF. Wegrzyn, KarenM. Bolitho, and RobertJ. Redgwell. 2004. 'Mannan Transglycosylase: A Novel Enzyme Activity in Cell Walls of Higher Plants'. *Planta* 219 (4): 1091–1102. https://doi.org/10.1007/s00425-004-1274-x.

Schroeven, Lindsey, Willem Lammens, André Van Laere, and Wim Van den Ende. 2008. 'Transforming Wheat Vacuolar Invertase into a High Affinity Sucrose:Sucrose 1-Fructosyltransferase'. *New Phytologist* 180 (4): 822–31. https://doi.org/10.1111/j.1469-8137.2008.02603.x.

Schultink, Alex, Kun Cheng, Yong Bum Park, Daniel J. Cosgrove, and Markus Pauly. 2013. 'The Identification of Two Arabinosyltransferases from Tomato Reveals Functional Equivalency of Xyloglucan Side Chain Substituents'. *Plant Physiology* 163 (1): 86–94. https://doi.org/10.1104/pp.113.221788.

Seeberger, Peter H. 2015. 'The Logic of Automated Glycan Assembly'. *Accounts of Chemical Research* 48 (5): 1450–63. https://doi.org/10.1021/ar5004362.

Seeberger, Peter H., and Daniel B. Werz. 2007. 'Synthesis and Medical Applications of Oligosaccharides'. *Nature* 446 (7139): 1046–51. https://doi.org/10.1038/nature05819.

Seiboth, Bernhard, and Benjamin Metz. 2011. 'Fungal Arabinan and L-Arabinose Metabolism'. *Applied Microbiology and Biotechnology* 89 (6): 1665–73. https://doi.org/10.1007/s00253-010-3071-8.

Shi, Hao, Ying Zhang, Baiyun Xu, Maobing Tu, and Fei Wang. 2014. 'Characterization of a

Novel GH2 Family α-L-Arabinofuranosidase from *Hyperthermophilic Bacterium Thermotoga Thermarum*'. *Biotechnology Letters* 36 (6): 1321–28. https://doi.org/10.1007/s10529-014-1493-6.

Sinnott, Michael L. 1990. 'Catalytic Mechanisms of Enzymic Glycosyl Transfer'. *Chemical Reviews* 90 (7): 1171–1202. https://doi.org/10.1021/cr00105a006.

Splain, Rebecca A., and Laura L. Kiessling. 2010. 'Synthesis of Galactofuranose-Based Acceptor Substrates for the Study of the Carbohydrate Polymerase GlfT2'. *Bioorganic & Medicinal Chemistry* 18 (11): 3753–59. https://doi.org/10.1016/j.bmc.2010.04.068.

Stam, M. R., E. G.J. Danchin, C. Rancurel, P. M. Coutinho, and B. Henrissat. 2006. 'Dividing the Large Glycoside Hydrolase Family 13 into Subfamilies: Towards Improved Functional Annotations of -Amylase-Related Proteins'. *Protein Engineering Design and Selection* 19 (12): 555–62. https://doi.org/10.1093/protein/gzl044.

Stevenson, Gordon, Brian Neal, D a N Liu, Matthew Hobbs, Nicolle H Packer, Michael Batley, John W Redmond, Lennart Lindquist, and Peter Reeves. 1994. 'K-12 and the Sequence of Its Rfb Gene Cluster . Structure of the 0 Antigen of *Escherichia Coli* K-12 and the Sequence of Its RIb Gene Cluster'. *J. Bacteriol.* 176 (13): 4144–56. https://doi.org/10.1128/jb.176.13.4144-4156.1994.

Stoco, Patrícia Hermes, Cassandra Aresi, Débora Denardin Lückemeyer, Maísa Michels Sperandio, Thaís Cristine Marques Sincero, Mário Steindel, Luiz Claudio Miletti, and Edmundo Carlos Grisard. 2012. '*Trypanosoma Rangeli* Expresses a β-Galactofuranosyl Transferase'. *Experimental Parasitology* 130 (3): 246–52. https://doi.org/10.1016/j.exppara.2011.12.005.

Sun, Yecheng, Xuguo Duan, Lei Wang, and Jing Wu. 2016. 'Enhanced Maltose Production through Mutagenesis of Acceptor Binding Subsite +2 in *Bacillus Stearothermophilus* Maltogenic Amylase'. *Journal of Biotechnology* 217: 53–61. https://doi.org/10.1016/j.jbiotec.2015.11.007.

Suzuki, Masayuki, Kyoko Kaneda, Yukiko Nakai, Motomitsu Kitaoka, and Hajime Taniguchi. 2009. 'Synthesis of Cellobiose from Starch by the Successive Actions of Two Phosphorylases'. *New Biotechnology*. https://doi.org/10.1016/j.nbt.2009.07.004.

Szczepina, Monica G., Ruixiang B. Zheng, Gladys C. Completo, Todd L. Lowary, and B. Mario Pinto. 2010. 'STD-NMR Studies of Two Acceptor Substrates of GlfT2, a Galactofuranosyltransferase from *Mycobacterium Tuberculosis*: Epitope Mapping Studies'. *Bioorganic and Medicinal Chemistry* 18 (14): 5123–28. https://doi.org/10.1016/j.bmc.2010.05.069.

Taha, Hashem A., Michele R. Richards, and Todd L. Lowary. 2013. 'Conformational Analysis of Furanoside-Containing Mono- and Oligosaccharides'. *Chemical Reviews* 113 (3): 1851–76. https://doi.org/10.1021/cr300249c.

Taira, Toki, Maho Fujiwara, Nicole Dennhart, Hiroko Hayashi, Shoko Onaga, Takayuki Ohnuma, Thomas Letzel, Shohei Sakuda, and Tamo Fukamizo. 2010. 'Transglycosylation Reaction Catalyzed by a Class V Chitinase from Cycad, Cycas Revoluta: A Study Involving Site-Directed Mutagenesis, HPLC, and Real-Time ESI-MS'. *Biochimica et Biophysica Acta - Proteins and Proteomics*. https://doi.org/10.1016/j.bbapap.2009.10.015.

Tam, Pui Hang, and Todd L. Lowary. 2009. 'Recent Advances in Mycobacterial Cell Wall Glycan Biosynthesis'. *Current Opinion in Chemical Biology*.

https://doi.org/10.1016/j.cbpa.2009.09.012.

Tefsen, Boris, Arthur F.J. Ram, Irma Van Die, and Françoise H. Routier. 2012. 'Galactofuranose in Eukaryotes: Aspects of Biosynthesis and Functional Impact'. *Glycobiology* 22 (4): 456–69. https://doi.org/10.1093/glycob/cwr144.

Teze, David, Franck Daligault, Vincent Ferrières, Yves Henri Sanejouand, and Charles Tellier. 2015. 'Semi-Rational Approach for Converting a GH36 α-Glycosidase into an α-Transglycosidase'. *Glycobiology* 25 (4): 420–27. https://doi.org/10.1093/glycob/cwu124.

Teze, David, Johann Hendrickx, Mirjam Czjzek, David Ropartz, Yves-Henri Sanejouand, Vinh Tran, Charles Tellier, and Michel Dion. 2014. 'Semi-Rational Approach for Converting a GH1 β-Glycosidase into a β-Transglycosidase'. *Prot. Eng. Des. Sel.* 27 (1): 13–19. https://doi.org/10.1093/protein/gzt057.

Teze, David, Johann Hendrickx, Michel Dion, Charles Tellier, Virgil L. Woods, Vinh Tran, and Yves-Henri Sanejouand. 2013. 'Conserved Water Molecules in Family 1 Glycosidases: A DXMS and Molecular Dynamics Study'. *Biochemistry* 52 (34): 5900–5910. https://doi.org/10.1021/bi400260b.

Thakur, Abhijeet, Kedar Sharma, and Arun Goyal. 2019. 'Chapter 12: α-l-Arabinofuranosidase: A Potential Enzyme for the Food Industry'. In *Green Bio-Processes*, edited by Binod Parameswaran, Sunita Varjani, and Sindhu Raveendran, 229–44. Energy, Environment, and Sustainability. Singapore: Springer Singapore. https://doi.org/10.1007/978-981-13-3263-0_12.

Tonozuka, Takashi, Akiko Tamaki, Gaku Yokoi, Takatsugu Miyazaki, Megumi Ichikawa, Atsushi Nishikawa, Yukari Ohta, et al. 2012. 'Crystal Structure of a Lactosucrose-Producing Enzyme, Arthrobacter Sp. K-1 β-Fructofuranosidase'. *Enzyme and Microbial Technology* 51 (6–7): 359–65. https://doi.org/10.1016/j.enzmictec.2012.08.004.

Toshima, Gen, Shunsuke Kawamura, Tomohiro Araki, and Takao Torikata. 2003. 'Histidine-114 at Subsites E and F Can Explain the Characteristic Enzymatic Activity of Guinea Hen Egg-White Lysozyme'. *Bioscience, Biotechnology, and Biochemistry* 67 (3): 540–46. https://doi.org/10.1271/bbb.67.540.

Touzel, Jean Pierre, Michael O'Donohue, Philippe Debeire, Eric Samain, and Christelle Breton. 2000. 'Thermobacillus Xylanilyticus Gen. Nov., Sp. Nov., a New Aerobic Thermophilic Xylan-Degrading Bacterium Isolated from Farm Soil'. *International Journal of Systematic and Evolutionary Microbiology* 50 (1): 315–20. https://doi.org/10.1099/00207713-50-1-315.

Tran, Vinh, Lionel Hoffmann, Claude Rabiller, Charles Tellier, and Michel Dion. 2010. 'Rational Design of a GH1 β-Glycosidase to Prevent Self-Condensation during the Transglycosylation Reaction'. *Protein Engineering, Design and Selection* 23 (1): 43–49. https://doi.org/10.1093/protein/gzp068.

Trincone, Antonio. 2015. 'Uncommon Glycosidases for the Enzymatic Preparation of Glycosides'. *Biomolecules*. https://doi.org/10.3390/biom5042160.

Umekawa, Midori, Wei Huang, Bing Li, Kiyotaka Fujita, Hisashi Ashida, Lai-Xi Wang, and Kenji Yamamoto. 2008. 'Mutants of Mucor Hiemalis Endo-β- N -Acetylglucosaminidase Show Enhanced Transglycosylation and Glycosynthase-like Activities'. *Journal of Biological Chemistry* 283 (8): 4469–79. https://doi.org/10.1074/jbc.M707137200.

Umesiri, Francis E., Aditya K. Sanki, Julie Boucau, Donald R. Ronning, and Steven J.

Sucheck. 2010. ‘Recent Advances toward the Inhibition of MAG and LAM Synthesis in *Mycobacterium Tuberculosis*’. *Medicinal Research Reviews*, n/a-n/a. https://doi.org/10.1002/med.20190.

Vembaiyan, Kannan, Jean A. Pearcey, Milan Bhasin, Todd L. Lowary, and Wei Zou. 2011. ‘Synthesis of Sugar-Amino Acid-Nucleosides as Potential Glycosyltransferase Inhibitors’. *Bioorganic and Medicinal Chemistry*. https://doi.org/10.1016/j.bmc.2010.11.044.

Viëtor, Remco J., Rainer A. Hoffmann, Seven A.G.F. Angelino, Alphons G.J. Voragen, Johannis P Kamerling, and Johannes F.G. Vliegenthart. 1994. ‘Structures of Small Oligomers Liberated from Barley Arabinocylans by Endoxylanase from *Aspergillus Awamori*’. *Carbohydrate Research* 254 (C): 245–55. https://doi.org/10.1016/0008-6215(94)84257-4.

Voragen, Alphons G.J., Gerd Jan Coenen, René P. Verhoef, and Henk A. Schols. 2009. ‘Pectin, a Versatile Polysaccharide Present in Plant Cell Walls’. *Structural Chemistry* 20 (2): 263–75. https://doi.org/10.1007/s11224-009-9442-z.

Wada, Jun, Yuji Honda, Masamichi Nagae, Ryuichi Kato, Soichi Wakatsuki, Takane Katayama, Hajime Taniguchi, Hidehiko Kumagai, Motomitsu Kitaoka, and Kenji Yamamoto. 2008. ‘1,2-Alpha-L-Fucosynthase: A Glycosynthase Derived from an Inverting α-Glycosidase with an Unusual Reaction Mechanism’. *FEBS Letters* 582 (27): 3739–43. https://doi.org/10.1016/j.febslet.2008.09.054.

Wakarchuk, W. W., A. Cunningham, D. C. Watson, and N. M. Young. 1998. ‘Role of Paired Basic Residues in the Expression of Active Recombinant Galactosyltransferases from the Bacterial Pathogen Neisseria Meningitidis’. *Protein Engineering Design and Selection* 11 (4): 295–302. https://doi.org/10.1093/protein/11.4.295.

Wen, Liuqing, Garrett Edmunds, Christopher Gibbons, Jiabin Zhang, Madhusudhan Reddy Gadi, Hailiang Zhu, Junqiang Fang, Xianwei Liu, Yun Kong, and Peng George Wang. 2018. ‘Toward Automated Enzymatic Synthesis of Oligosaccharides’. *Chemical Reviews* 118 (17): 8151–87. https://doi.org/10.1021/acs.chemrev.8b00066.

Wesener, Darryl A., Matthew R. Levengood, and Laura L. Kiessling. 2017. ‘Comparing Galactan Biosynthesis in *Mycobacterium Tuberculosis* and *Corynebacterium Diphtheriae*’. *Journal of Biological Chemistry* 292 (7): 2944–55. https://doi.org/10.1074/jbc.M116.759340.

Wilkens, Casper, Susan Andersen, Claire Dumon, Jean Guy Berrin, and Birte Svensson. 2017. ‘GH62 Arabinofuranosidases: Structure, Function and Applications’. *Biotechnology Advances* 35 (6): 792–804. https://doi.org/10.1016/j.biotechadv.2017.06.005.

Williams, Spencer. 2015. ‘Transglycosylases’. CAZypedia. 2015. https://www.cazypedia.org/index.php/Transglycosylases.

Wing, Corin, James C. Errey, Balaram Mukhopadhyay, John S. Blanchard, and Robert A. Field. 2006. ‘Expression and Initial Characterization of WbbI, a Putative D-Galf:α-D-Glc β-1,6-Galactofuranosyltransferase from Escherichia Coli K-12’. *Org. Biomol. Chem.* 4 (21): 3945–50. https://doi.org/10.1039/B609455D.

Winkler, Margit, Martina Geier, Steven P. Hanlon, Bernd Nidetzky, and Anton Glieder. 2018. ‘Human Enzymes for Organic Synthesis’. *Angewandte Chemie - International Edition* 57 (41): 13406–23. https://doi.org/10.1002/anie.201800678.

Withers, S. 2001. ‘Mechanisms of Glycosyl Transferases and Hydrolases’. *Carbohydrate*

Polymers 44 (4): 325–37. https://doi.org/10.1016/S0144-8617(00)00249-6.

Withers, Stephen G. 1999. '1998 Hoffmann La Roche Award Lecture Understanding and Exploiting Glycosidases'. *Canadian Journal of Chemistry* 77 (1): 1–11. https://doi.org/10.1139/v98-235.

Wittrup Larsen, Marianne, Dorota F. Zielinska, Mats Martinelle, Aurelio Hidalgo, Lars Juhl Jensen, Uwe T. Bornscheuer, and Karl Hult. 2010. 'Suppression of Water as a Nucleophile in Candida Antarctica Lipase B Catalysis'. *ChemBioChem* 11 (6): 796–801. https://doi.org/10.1002/cbic.200900743.

Xu, Cao, Katie L Liberatore, Cora A MacAlister, Zejun Huang, Yi-Hsuan Chu, Ke Jiang, Christopher Brooks, et al. 2015. 'A Cascade of Arabinosyltransferases Controls Shoot Meristem Size in Tomato'. *Nature Genetics* 47 (7): 784–92. https://doi.org/10.1038/ng.3309.

Yang, Jingwen, Qi Wang, Ye Zhou, Jingbo Li, Renjun Gao, and Zheng Guo. 2017. 'Engineering T. Naphthophila β-Glucosidase for Enhanced Synthesis of Galactooligosaccharides by Site-Directed Mutagenesis'. *Biochemical Engineering Journal* 127 (November): 1–8. https://doi.org/10.1016/j.bej.2017.07.008.

Yin, Jie, Lei Li, Neil Shaw, Yang Li, Jing Katherine Song, Wenpeng Zhang, Chengfeng Xia, et al. 2009. 'Structural Basis and Catalytic Mechanism for the Dual Functional Endo-β-N-Acetylglucosaminidase A'. Edited by Bostjan Kobe. *PLoS ONE* 4 (3): e4658. https://doi.org/10.1371/journal.pone.0004658.

Zakariassen, Henrik, Mona Cecilie Hansen, Maje Jøranli, Vincent G H Eijsink, and Morten Sørlie. 2011. 'Mutational Effects on Transglycosylating Activity of Family 18 Chitinases and Construction of a Hypertransglycosylating Mutant'. *Biochemistry* 50 (25): 5693–5703. https://doi.org/10.1021/bi2002532.

Zechel, David L., and Stephen G. Withers. 2000. 'Glycosidase Mechanisms: Anatomy of a Finely Tuned Catalyst'. *Accounts of Chemical Research* 33 (1): 11–18. https://doi.org/10.1021/ar970172+.

Zeuner, Birgitte, David Teze, Jan Muschiol, and Anne S. Meyer. 2019. 'Synthesis of Human Milk Oligosaccharides: Protein Engineering Strategies for Improved Enzymatic Transglycosylation'. *Molecules* 24 (11): 2033. https://doi.org/10.3390/molecules24112033.

Zhao, Jiao, Tobias Tandrup, Bastien Bissaro, Sophie Barbe, Isabelle André, Claire Dumon, Leila Lo Leggio, Michael J. O'Donohue, and Régis Fauré. 2020. 'Probing the Determinants of the Transglycosylation/Hydrolysis Partition in a Retaining α-L-Arabinofuranosidase'. *Unpublished*.

Chapter II.

Probing the determinants of the transglycosylation/hydrolysis partition in a retaining α-L-arabinofuranosidase

A key aim of this thesis is to pursue work on *Tx*Abf, refining understanding of the T/H partition of *Tx*Abf and incrementing our already rich knowledge base. To this end, this chapter focuses on a structure function study performed specifically on two first shell residues (F26 and L352) that are located in the vicinity of subsite -1. The aim of the study was to compare the functional roles of these two residues, using a variety of experimental and computational techniques to extract a maximum number of details. Advantageously, working in collaboration with the University of Copenhagen, this study provided us with the opportunity to establish some new 3D structures and probe molecular dynamics.

One of the most significant results in this work is the identification of an alternative acceptor binding site, which is the result of the mutation N216W. Moreover, this work allowed us to better appreciate the role played by R69 and reveal the determinant role played by loop flexibility in both donor and acceptor binding.

Probing the determinants of the transglycosylation/hydrolysis partition in a retaining α-L-arabinofuranosidase

Jiao Zhao[1], Tobias Tandrup[2], Bastien Bissaro[1], Sophie Barbe[1], Isabelle André[1], Claire Dumon[1], Leila Lo Leggio[2], Michael J. O'Donohue[1] and Régis Fauré[1]

[1] TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France

[2] University of Copenhagen, Department of Chemistry, Copenhagen, Denmark

Correspondence

M.J. O'Donohue and R. Fauré, Toulouse Biotechnology Institute, Bio & Chemical Engineering, Université de Toulouse - CNRS 5504 - INRA 792 - INSA de Toulouse, 135 avenue de Rangueil, 31077 Toulouse CEDEX 04, France

Tel: +33 5 6155 9410

E-mail: michael.odonohue@inra.fr (M.J.O.); regis.faure@insa-toulouse.fr (R.F.)

Abstract

The use of retaining glycoside hydrolases (rGHs) as synthetic tools for glycochemistry is highly topical and the focus of considerable R&D work. However, due to the incomplete identification of the molecular determinants of the transglycosylation/hydrolysis partition (T/H), rational engineering of rGHs to create transglycosylases (TGs) remains challenging. Therefore, to better understand the factors that underpin transglycosylation in a GH51 retaining α-L-arabinofuranosidase from *Thermobacillus xylanilyticus* (*Tx*Abf) we have pursued our detailed investigation of this enzyme's active site. Specifically, two mutations, F26L and L352M, located in the vicinity of the active site are studied, measuring the kinetic parameters related to transglycosylation, determining 3D structures and performing molecular dynamic simulations. The results reveal that the presence of L352M in the context of a triple mutant (also containing R69H and N216W) generates changes both in the donor and acceptor subsites, the latter being the result of a domino-like effect. Overall, the mutant R69H-N216W-L352M displays excellent transglycosylation activity (70% yield, 78% transfer rate and reduced secondary hydrolysis of the product). In the course of this study, the key role played by the conserved R69 residue is also underlined. The mutation of R69H affects both the

catalytic nucleophile and the acid/base and has a determinant effect on the T/H partition. Finally, our results reveal that increased loop flexibility in the the positive subsites probably creates new interactions with the acceptor, in particular with a hydrophobic binding platform composed of N216W, W248 and W302. Besides, acid/base flexibility can also contribute to enhancement of transglycosylation.

Keywords: retaining glycoside hydrolase; engineered transglycosylases; transglycosylation ability; molecular interactions; flexibility

Abbreviations

Abf, α-L-arabinofuranosidase; L-Ara*f*, α-L-arabinofuranosyl unit; AXOS, arabinoxylo-oligosaccharides; A^2XX, α-L-Ara*f*-(1,2)-β-D-Xyl*p*-(1,4)-β-D-Xyl*p*-(1,4)-D-Xyl*p*; A^3XX, α-L-Ara*f*-(1,3)-β-D-Xyl*p*-(1,4)-β-D-Xyl*p*-(1,4)-D-Xyl*p*; DP, degree of polymerization; α-L-Ara*f*O*p*NP, 4-nitrophenyl α-L-arabinofuranoside; GH, glycoside hydrolase; GT, glycosyltransferase; LG, leaving group; *p*NP, 4-nitrophenol; rGH, retaining glycoside hydrolase; R_T, transfer rate; *Tx*Abf, α-L-arabinofuranosidase from *Thermobacillus xylanilyticus*; T/H, transglycosylation/hydrolysis ratio; TG, non-Leloir transglycosylase; TS, transition state; vdW, van der Waals; XOS, xylo-oligosaccharides; D-Xyl*p*, D-xylopyranosyl unit; X, D-xylose; X_2, (1,4)-β-D-xylobiose; X_3, (1,4)-β-D-xylotriose; X_4, (1,4)-β-D-xylotetraose; X_5, (1,4)-β-D-xylopentaose; X_6, (1,4)-β-D-xylohexaose; XA^3X, β-D-Xyl*p*-(1,4)-[α-L-Araf-(1,3)]-β-D-Xyl*p*-(1,4)-D-Xyl*p*; XA^3XX, β-D-Xyl*p*-(1,4)-[α-L-Araf-(1,3)]-β-D-Xyl*p*-(1,4)-β-D-Xyl*p*-(1,4)-D-Xyl*p*.

1. Introduction

The development of efficient *in vitro* strategies to synthesize carbohydrates remains a major challenge for synthetic chemistry. Despite considerable progress, the complexity and diversity of carbohydrates means that the quest for generic approaches remains highly topical and relevant for understanding the role of carbohydrates in biological systems [1,2]. For decades, enzymes have been recognized as useful tools to tackle the complexity of carbohydrate chemistry and surmount the lack of selectivity of chemical catalysts. Among enzyme candidates, Leloir and non-Leloir carbohydrate-active enzymes are widely studied for their ability to synthesize various target glycoconjugates [3–5].

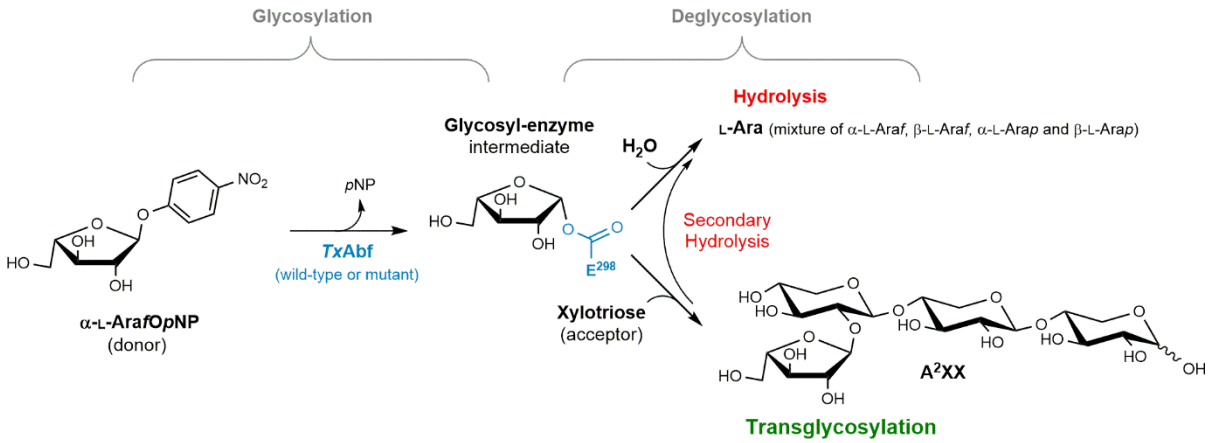


Fig. 1. Two-step displacement mechanism (also known as Ping-Pong Bi Bi mechanism in transglycosylation) of retaining *Tx*Abf. The first step (glycosylation) leads to release of the *p*NP leaving group from the donor, α-L-Ara*f*O*p*NP in this study, and concomitant formation of the covalent glycosyl-enzyme intermediate. Regarding the second step (deglycosylation), the covalent glycosyl-enzyme intermediate can be attacked by either a water molecule (hydrolysis) or an external acceptor (transglycosylation, xylotriose in this study). In the case of secondary hydrolysis, the transglycosylation product becomes a donor substrate with a subsequent deglycosylation step involving water.

Leloir glycosyltransferases (GTs), *which use* sugar nucleotides as donors, can be considered as Nature's solution for carbohydrate synthesis, since *in vivo* these enzymes are responsible for the formation of vast majority of carbohydrates. However, harnessing GTs for the purposes of synthetic chemistry is fraught with challenges (e.g., expression of GTs in heterologous systems, the cost of sugar nucleotides, etc.) [6–8]. Therefore, despite recent progress in the area [9–11], the development of alternative strategies employing other enzyme

classes, particularly retaining glycoside hydrolases (hereafter designated rGHs), is the subject of intense research. rGHs operate through a two-step displacement mechanism (Fig. 1) [12] that provides the basis for the occurrence of non-Leloir transglycosylases (TGs), which are actually variants of rGHs that perform synthetic roles in biological systems [13,14]. The occurrence of TGs, the fact that rGHs in general are extremely abundant (67% of all 755 496 classified GH modules among 83 of the 167 GH families in the CAZY database, http://www.cazy.org/ [15], February 24th, 2020) and that rGHs cover a wide range of substrate specificities makes this group of carbohydrate-active enzymes extremely interesting targets for synthetic chemists [16].

In the 2-step mechanism used by rGHs, a covalent glycosyl-enzyme intermediate is formed that can either be deglycosylated by a water molecule (i.e., hydrolysis) or by a carbohydrate acceptor, which leads to transglycosylation [12,17]. In most rGHs operating in aqueous medium, hydrolysis is the principal outcome of the reaction, because of omnipresent water. However, naturally-occurring TGs reveal that specific protein modifications can diminish water-mediated deglycosylation and thus favour transglycosylation. This evidence has prompted the creation of glycosynthetic rGHs using protein engineering, one of the best examples being the glycosynthase strategy [18–21]. In this strategy, catalytically-impotent mutants are fed with fluoride-activated substrates. However, despite the relative simplicity of this elegant strategy, it is rather difficult to apply when using furanoside donors, because the synthesis of fluoride-activated sugars has so far proven impossible. Alternative methods to engineer artificial TGs rely on the induction of shifts in the transglycosylation/hydrolysis (T/H) partition. This can be achieved either by engineering the reaction conditions (e.g., reducing water activity), or by modifying specific amino acids that influence the T/H partition [22–24]. Three main approaches have been identified to create TGs starting from rGHs [16]. The first involves the modification of substrate interactions in the donor subsite(s), leading to increases in the transition state (TS) energy barriers [25,26]. The aim is to reduce the efficiency of water-mediated deglycosylation, thus favouring transglycosylation, although overall catalytic activity is also affected. The second approach relies on modification of acceptor subsites, the objective being to improve enzyme-acceptor interactions and thus favour sugar-mediated deglycosylation (*i.e.*, transglycosylation). This is often achieved by introducing hydrophobic amino acid sidechains at the acceptor site(s), which form favourable interactions with incoming glycoside acceptors and repel water molecules [25,27]. When combined with donor subsite engineering, the introduction of favourable acceptor site

interactions can partially compensate for the overall reduction in efficiency due to the poor formation of the glycosyl-enzyme intermediate [28]. However, when acceptor subsite(s) engineering is used as a standalone strategy, it rarely affords radical diminution of hydrolysis [17]. The third method to modulate the T/H partition targets the ability of the enzyme to fix, convey and/or activate catalytic water molecules [29–32]. This generally requires knowledge related to the location of water channels and/or the identification of specific amino acids responsible for positioning water for catalysis.

Recent reports describe a semi-rational method to engineer TGs that involves targeting highly conserved residues that are located within, or in the vicinity of subsite -1 [26,33]. Interestingly, this approach is described as a potentially generic strategy [34]. However, despite this prospect and the previous use of a variety of techniques, including *in vitro* mutagenesis and *in silico* design (e.g., molecular dynamics simulations [32,35] and quantum mechanical/molecular mechanical studies [36,37]), an accurate and reliable method to predictably optimize the T/H partition has yet to be found.

In recent work performed on the GH51 α-L-arabinofuranosidase from *Thermobacillus xylanyliticus* (*Tx*Abf), we used a combinatory approach to create the first non-Leloir transarabinofuranosylases [17]. These enzymes are actually derivatives of *Tx*Abf bearing three mutations, R69H-N216W-L352M and F26L-R69H-N216W (Fig. 2), which display remarkable transglycosylation capability (almost quantitative overall yield, with A^2XX as main product [38]) [17]. Each of the point mutations in these enzymes plays a distinct role in defining the novel activity. In subsite -1, R69H lies within a H-bond distance (1.9 Å) of the catalytic nucleophile E298. It has been postulated that this mutation impairs the nucleophilic strength of E298, leading to destabilization of pK_a cycling, an intrinsic feature of the double-displacement mechanism [39]. N216W introduces hydrophobicity into subsite +2 and influences the positioning of the acceptor, thus increasing transglycosylation and regioselectivity for A^2XX. Additionally, forming part of subsite -1, L352M is located close (3.5 Å) to *O*-5 of the L-Ara*f* moiety. This mutation decreases the catalytic efficiency of *Tx*Abf [17]. Furthermore, in previous work F26L was pinpointed as a potential target to improve transglycosylation [40] and more recently we have shown that this mutation procures increased ability to transfer the L-Ara*f* donor onto a xylotriose acceptor [34]. Significantly, F26 is also located in subsite -1, is 3.8 Å distant from E298 and lies in the vicinity (3.8 Å) of the *O*-5 position of the L-Ara*f* moiety. Finally, previous work has also identified the beneficial

G179F mutant (Fig. 2). This mutation was first predicted using an *in silico* method and then tested experimentally, revealing that it leads to better binding of the xylotriose acceptor [17].

In the present study, we constructed a novel triple mutant F26L-R69H-N216W, providing the basis to compare this mutant with R69H-N216W-L352M, and thus appraise the relative impacts of F26L and L352M on the *Tx*Abf T/H partition. Moreover, we designed two quadruple mutants F26L-R69H-N216W-L352M and R69H-G179F-N216W-L352M to investigate to what extent synergistic effects can be generated. To acquire atomic level insight we solved crystal structures of mutated enzymes and performed molecular dynamic simulations.

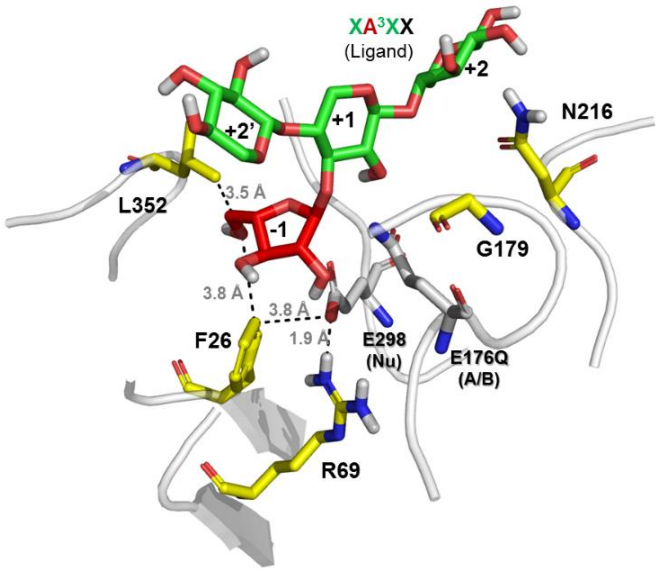


Fig. 2. View of the *Tx*Abf-E176Q:XA^{3}XX complex active site showing the positioning of F26, R69, G179, N216 and L352 residues as well as the mutated acid/base (A/B) E176Q and the catalytic nucleophile (Nu) E298 (PDB code: 2VRQ) [41]. The reducing D-Xyl*p* unit, X in black, of pentasaccharide XA^{3}XX (i.e., β-D-Xyl*p*-(1,4)-[α-L-Araf-(1,3)]-β-D-Xyl*p*-(1,4)-β-D-Xyl*p*-(1,4)-D-Xyl*p*) is not observed.

2. Results

2.1. Specific activity of wild-type *Tx*Abf and mutants thereof

To measure specific activities (SA) in both hydrolysis and transglycosylation modes, the release of *p*-nitrophenol (*p*NP) from α-L-Ara*f*O*p*NP was monitored (Table 1). Release of *p*NP relates to the first step in the double-displacement mechanism (glycosylation), thus the SA measurement is an aggregate value that captures global activity. This includes

transglycosylation (SA_T), hydrolysis and self-condensation (SA_H). In this assay, all mutants displayed decreased SA_H (0.1 to 40% compared to *Tx*Abf). However, unlike glycosynthases [42], catalytic activity remained measurable and significant in both modes. It is noteworthy that some mutants displayed a SA_T/SA_H ratio >1, meaning that these preferentially perform transglycosylation in the presence of the acceptor. However, this enhanced capability correlated with decreased specific activity in both hydrolysis and transglycosylation modes (between 0.3 and 5.9% for SA_T compared to the wild-type).

Table 1. Specific activities in hydrolysis (SA_H) and transglycosylation (SA_T) modes.[a]

Enzyme	SA_H ($IU.mg^{-1}$)[b]	SA_T ($IU.mg^{-1}$)[c]	SA_T/SA_H[d]
wt	261.79	125.49	0.5
F26L	105.76	83.26	0.8
R69H	2.84	7.41	2.6
L352M[e]	86.91	47.00	0.5
R69H-N216W	1.60	2.89	1.8
R69H-N216W-L352M	0.97	1.82	1.9
F26L-R69H-N216W	0.32	0.60	1.9
R69H-G179F-N216W-L352M	0.29	0.42	1.4
F26L-R69H-N216W-L352M	0.27	0.93	3.4

[a]One unit (IU) of enzyme specific activity corresponds to the amount of enzyme releasing 1 µmol of *p*NP per minute. Experiments were performed in triplicate and relative errors were always less than 10%.

[b]Reactions operating in hydrolysis mode contain only 5 mM α-L-Ara*f*O*p*NP.

[c]Reactions operating in transglycosylation mode contain both donor (5 mM α-L-Ara*f*O*p*NP) and acceptor (10 mM xylotriose).

[d]The ratio between SA_T and SA_H reveals the extent of activation or inhibition by the acceptor.

[e]SA of L352M was calculated using previously reported data acquired in the same reaction conditions (i.e., 45 °C in 50 mM sodium phosphate buffer at pH 7.0) [17].

2.2. Kinetic parameters in both hydrolytic and transglycosylation modes

Kinetic parameters were determined in both hydrolysis and transglycosylation modes for donor and acceptor substrates individually. While the kinetic profile of reactions catalysed by *Tx*Abf followed the classic Michaelis-Menten model (Equation 1), mutants performing high

levels of self-condensation in the hydrolysis mode described a two-phase reaction trajectory that did not reach saturation (Fig. 3A). This result is reminiscent of the behaviour of *Tx*Abf R69-containing mutants and also that of a *Thermus thermophilus* β-glycosidase consuming a *o*-nitrophenyl β-D-galactopyranoside donor, which also displayed two-phase reaction profile [17,43]. To model this reaction, a modified Michaelis-Menten equation (2) was used that contains a nonspecific constant (N_s). This constant reflects activation by the donor substrate, which drives self-condensation. Unlike hydrolysis mode, plotting SA as a function of either donor or acceptor concentration in transglycosylation mode generated curves that can be fitted using the Michaelis-Menten model for all enzymes. Nevertheless, even in the presence of 100 mM xylotriose, reactions catalysed by mutants failed to reach maximal rate, implying that saturation was not attained (Fig. 3B). For *Tx*Abf and mutants R69H, L352M and R69H-L352M, a previous Brønsted-Hammett analysis has shown that when the donor substrate bears a good leaving group, deglycosylation constitutes the rate limiting step [17]. Accordingly, we postulate that this is also true for the mutants N216W and G179F, since neither is sufficiently close to the donor binding site to impact on the nucleophile and/or acid/base residues. This assumption is also made for R69H-N216W-L352M and R69H-G179F-N216W-L352M, meaning that when the reaction involves a *p*NP-bearing donor the ratio k_{cat}/K_M and the rate constant k_{cat} correlate with the glycosylation and deglycosylation steps, respectively.

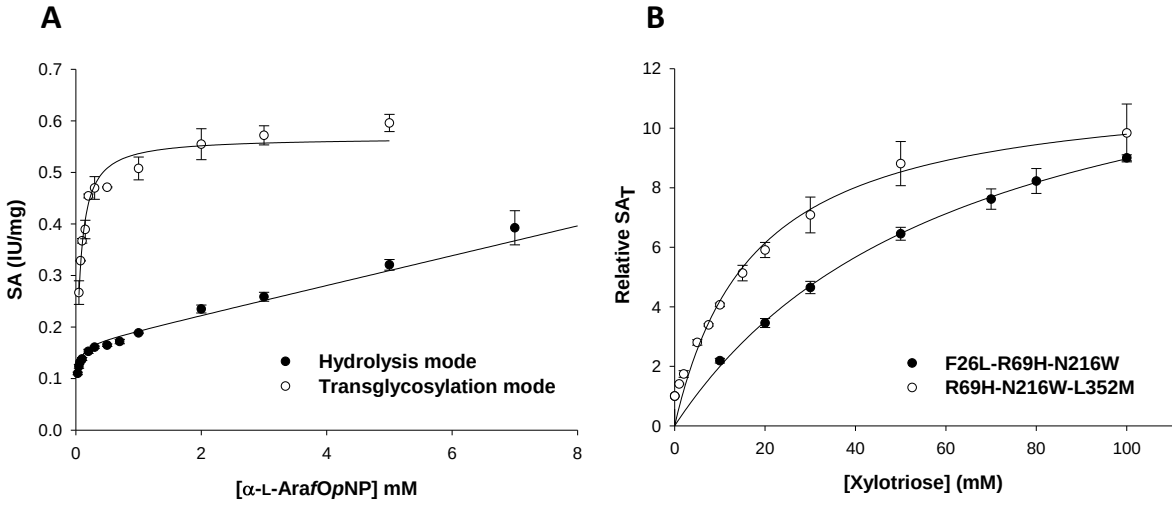


Fig. 3. Steady-state kinetics of F26L-R69H-N216W and R69H-N216W-L352M. (A) F26L-R69H-N216W in hydrolysis (●) and transglycosylation (○) modes; (B) Relative activity ($SA_T/SA_{[Xylotriose]\ =\ 0\ mM}$) as a function of xylotriose concentration indicating the acceptor-

mediated activation of F26L-R69H-N216W (●) and R69H-N216W-L352M (○) in transglycosylation mode with fixed 3 mM α-L-Ara*f*O*p*NP as donor.

When operating in hydrolysis mode (Table 2), the three mutants displayed reduced (nearly 1 000-fold) catalytic turnover compared to *Tx*Abf, which reflects impaired water-mediated deglycosylation. Consequently, since both glycosylation and deglycosylation TS display similar properties, we believe that glycosylation is also impacted, a postulate that is supported by much lower k_{cat}/K_M values compared to that of *Tx*Abf, these being in the range 0.2-1.9% of the wild-type value. Significantly, lowered catalytic efficiency has been previously described as an intrinsic feature of TGs [26,44]. Fortunately, in the case of F26L-R69H-N216W and R69H-G179F-N216W-L352M acting on α-L-Ara*f*O*p*NP, lower catalytic turnover is offset by lowered K_M values, meaning that the glycosyl-enzyme intermediate is formed and displays a sufficiently long half-life to allow glycoside acceptor-mediated deglycosylation.

In transglycosylation mode (*i.e.*, involving acceptor-mediated deglycosylation), equation (3) was used to fit the kinetic profiles of reactions catalysed by mutant enzymes. When using a constant concentration of 10 mM xylotriose acceptor (Table 2), the k_{cat} value of the triple mutants increased 13- to 32-fold when compared to the k_{cat} value measured in hydrolysis mode. Additionally, the presence of 10 mM acceptor barely altered (2-fold higher) the K_M (donor) value of the reaction catalysed by R69H-N216W-L352M, but significantly increased that of F26L-R69H-N216W. Regarding the k_{cat}/K_M, the addition of 10 mM xylotriose procured a 6-fold increase of the reaction catalysed by R69H-N216W-L352M, but had little effect (1.2-fold decrease) on the reaction containing F26L-R69H-N216W (Table 2). The fact that the K_M (acceptor) value of the reaction involving F26L-R69H-N216W is 75 mM (Table 3) implies that the interaction of xylotriose with the enzyme is weak and probably explains why even at 10 mM it does not further increase catalytic efficiency.

Table 2. Kinetic parameters on α-L-Ara*f*O*p*NP (donor)[a] in hydrolysis mode and transglycosylation mode[b] (data shown in bracket) at pH 7.0.[2]

Enzyme	SA_{th} ($IU.mg^{-1}$)	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($s^{-1}.mM^{-1}$)	Ns[c] ($s^{-1}.mM^{-1}$)
wt [c]	145	0.25	139	556	-
R69H-N216W-L352M[c]	0.60 (7.76)	0.48 (1.03)	0.58 (7.46)	1.21 (7.24)	0.03 (-)
F26L-R69H-N216W	0.16 (4.98)	0.01 (0.53)	0.15 (4.78)	10.81 (9.02)	0.03 (-)
R69H-G179F-N216W-L352M	0.14	0.02	0.14	6.45	0.03

[a]Experiments were performed in triplicate and relative errors were always less than 10%.

[b]In transglycosylation mode, the concentration of xylotriose acceptor was fixed at 10 mM.

[c]Ns is a nonspecific constant that is included in the modified Michaelis-Menten equation (2) to account for activation of the enzyme by the self-condensation product: $SA_{app} = SA_{th}\cdot[S]/(K_M + [S]) + N_S\cdot[S]$ where SA_{th} is the theoretical maximum activity achieved if the enzyme operates according to the Michaelis-Menten model.

[d]Data in hydrolysis mode from previous work acquired in the same conditions [17].

Table 3. Kinetic parameters on xylotriose (acceptor) in transglycosylation mode at pH 7.0.[a]

Enzyme	SA_{max} ($IU.mg^{-1}$)	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($s^{-1}.mM^{-1}$)
R69H-N216W-L352M-	8.28	23.32	7.96	0.34
F26L-R69H-N216W	4.49	75.43	4.31	0.06

[a]The concentration of α-L-Ara*f*O*p*NP donor was fixed at 3 mM. Experiments were performed in triplicate and relative errors were always less than 10%.

2.3. Transglycosylation profiles of mutants

H_2O to D_2O exchange did not affect overall catalytic efficiency, nor the transglycosylation profiles of *Tx*Abf and mutants thereof (Fig. S2). Moreover, the kinetic parameters (Tables 2-3 and S1-S2) and transglycosylation yields (Table S3) of reactions catalyzed by R69H-L352M-N216W and F26L-R69H-N216W were only slightly sensitive to pH. However,

[2] Kinetic parameters were measured for some hilighted mutants, not for all the mutants appeared in the article. Those data were sufficient to analyse the relationship with transglycosylation yield.

transglycosylation yields were highly influenced by the acceptor/donor substrate ratio, with maximum yields being obtained at a ratio of 3.33 (Table S4). On the other hand, for a given acceptor/donor substrate ratio, the global substrate concentrations (donor + acceptor) slightly affected yield (Table S5). Taking all of these results into account we defined the reaction conditions for time course ^{1}H NMR monitoring [17].

Table 4. A^2XX transglycosylation and self-condensation yields (determined by NMR) for reactions catalyzed by *Tx*Abf and mutants thereof.[a]

Enzyme	Yield (%)		
	A^2XX 5.20 ppm[c]	Self-condensation (1,2)[b] 5.88 ppm[c]	Self-condensation (1,3)[b] 5.81 ppm[c]
wt	4	5	1
R69H-N216W	57	2	18
R69H-N216W-L352M	70	2	9
F26L-R69H-N216W	59	2	15
R69H-G179F-N216W-L352M	59	3	9
F26L-R69H-N216W-L352M	66	2	10

[a]Kinetic assays were carried out using 5 mM α-L-Ara*f*O*p*NP and 10 mM xylotriose at 45 °C and pH 7.0 in buffered 10% D_2O.

[b]Self-condensation products are α-L-Ara*f*-(1,2)-α-L-Ara*f*O*p*NP and α-L-Ara*f*-(1,3)-α-L-Ara*f*O*p*NP.

[c]NMR chemical shift of the anomeric proton of transferred α-L-Ara*f* unit of products at 45 °C.

When transferring α-L-Ara*f* units onto xylotriose, wild-type *Tx*Abf produces a mixture of AXOS regioisomers at an overall 9% yield (Fig. S3), containing A^2XX (4% yield, Table 4). All mutants bearing the substitution N216W displayed significantly improved regioselectivity, which drove the reaction towards the formation of A^2XX, the yields of XA^3X and A^3XX being significantly reduced (Fig. S3) [17]. Since the distinctive anomeric signals of XA^3X and A^3XX (at 5.32 and 5.25 ppm respectively) are barely detectable during reactions catalyzed by N216W-containing mutants (Fig. S3), we used A^2XX yield as an indicator of global transglycosylation efficiency.

Compared to *Tx*Abf, all mutants strongly modulated the T/H partition in favour of transglycosylation by combining both high transglycosylation yields (57-70%) and high transfer rate (R_T, 59-78%, Table S6), with R69H being the principal cause of these

enhancements [17]. Nevertheless, other amino acid substitutions also contributed subtle improvements. The previously generated R69H-N216W-L352M proved to be the most potent enzyme (up to 70% yield and a transfer rate of 78%) for A^2XX synthesis (Fig. 4). Even though the transglycosylation yield of F26L-R69H-N216W is close to that of R69H-N216W (Table 4), F26L promotes higher transfer rate (R_T, 1.1-fold) and alleviates secondary hydrolysis when combined with R69H-N216W (Table S6). Adding F26L to the most potent triple mutant background (i.e., F26L + R69H-N216W-L352M) slightly lowers A^2XX synthesis, while maintaining the self-condensation level (12%). In this respect, while all mutants synthesize A^2XX as the major transglycosylation product, mutants R69H-N216W and F26L-R69H-N216W display increased self-condensation capability (17-20% yield) and the introduction of L352M mutation to any of the mutant restores self-condensation to an intermediate level (11-12% yield). This is indicative of enhanced competition between xylotriose (acceptor) and α-L-Ara*f*OpNP (both donor and acceptor) for occupation of the positive subsites. The addition of G179F to R69H-N216W-L352M was intended to reduce secondary hydrolysis of A^2XX. However, in addition to significant diminution of secondary hydrolysis (Table S6), this mutation actually led to a 1.2-fold reduction in A^2XX yield (Table 4 and S6, Figs. 4 and S4).

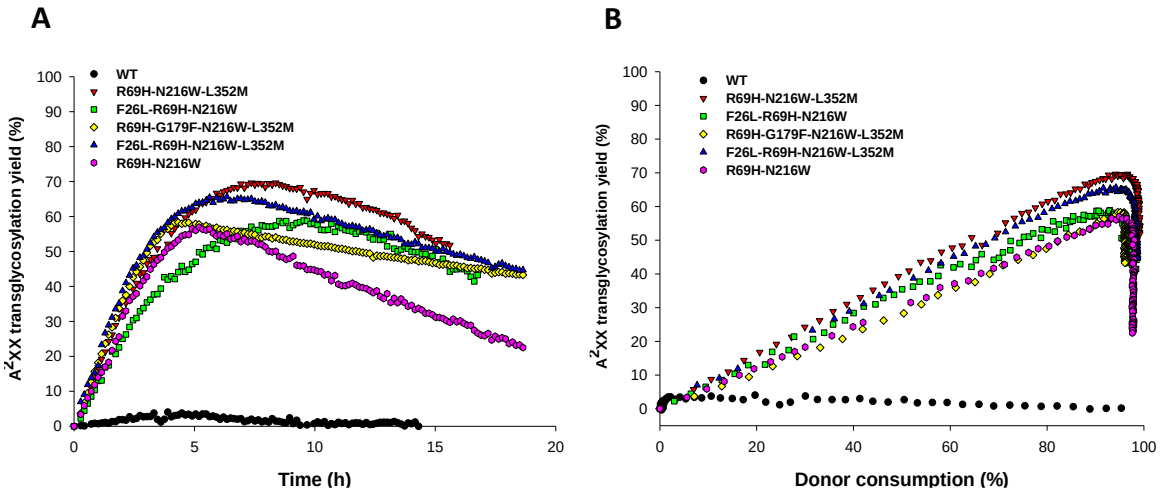


Fig. 4. NMR monitoring of A^2XX evolution as a function of (A) time and (B) donor conversion. All assays were carried out at 45 °C and pH 7.0 in buffered 10% D_2O, with 5 mM α-L-Ara*f*O*p*NP and 10 mM xylotriose as donor and acceptor respectively.

Secondary hydrolysis is exhibited when the linearity of donor concentration-dependent A^2XX production is lost (Fig. 4B). Therefore, plotting A^2XX yield/donor consumption (in %) *versus*

donor consumption provides a convenient way to observe the (tipping) point of loss of linearity (Fig. S4). For R69H-G179F-N216W-L352M, despite displaying a lower R_T, the tipping point occurs when the donor is almost completely (96%) consumed. However, for the mutants R69H-N216W-L352M, F26L-R69H-N216W-L352M and F26L-R69H-N216W, the tipping point is observed earlier at a mean value of 82% (Table S6).

2.4. Transglycosylation using XOS

To evaluate the impact of degree of polymerization (DP) of the acceptor on the ability of R69H-N216W-L352M to perform transglycosylation, D-xylose and a range of xylo-oligosaccharides (XOS), from xylobiose to xylohexaose, were used in reactions containing α-L-Ara*f*O*p*NP. When using D-xylose (X, DP = 1) as the acceptor, transglycosylation is barely detectable, with the yield being 2 and 5% respectively for the two regioisomer products (undeterminded linkages). It is noteworthy that R69H-N216W-L352M maintains its regioselectivity towards the transfer of α-L-Ara*f* moiety onto the *O*-2 position of the non-reducing end of XOS (Table S7). Using xylobiose (DP = 2) as acceptor procured two trisaccharides, which are very likely to be stereoisomers of A^2X (i.e., α and β forms at the reducing end of D-Xyl*p* moiety) [45], or regioisomers, XA^2/A^4X and A^2X, the latter being the least probable [46]. The use of longer XOS (DP ≥ 3) generated products that were all detected at 5.20 ppm meaning that each of the AXOS products share the same regioselectivity (i.e., the α-L-Ara*f* moiety is linked to the *O*-2 position of the non-reducing end of the XOS) as the product generated when xylotriose is used as acceptor. Moreover, the use of XOS displaying a DP ≥ 2 procured transglycosylation products at high yield (63-75%) and secondary hydrolysis was almost identical (Fig. S5), although SA_T decreased as a function of increasing acceptor DP (Table S7).

2.5. 3D structure of mutants R69H-L352M and R69H-N216W-L352M

The crystal structures of *apo*-R69H-L352M and *apo*-R69H-N216W-L352M revealed that L352M provokes a domino-like effect in its neighbouring environment. In particular, both W302 and W248, which were previously reported to interact with both the donor (W302) and acceptor within subsites +1 and +2 [41], are affected by L352M (Fig. 5A). Despite the fact that a structure bearing F26L is not available, we postulate that this mutation will have a much shorter range effect, affecting only subsite -1. Indeed, this is supported by the observation that the L352M (combined with R69H-N216W) mutation has greater influence on the T/H ratio in

favour of transglycosylation. Considering R69H, structural analysis shows that the shorter histidine sidechain induces a conformational change of N175 (Fig. 5B). In both the seleno-methionine substituted wild-type structure and that of the mutant E176Q (PDB codes: 2VRK and 2VRQ respectively), residue N175 interacts with both the acid/base (E176) and the nucleophile (E298). Therefore, R69H has an indirect, but significant effect on the catalytic apparatus. Finally, regarding the N216W mutation, the data clearly indicate that the presence of tryptophan in subsite +2 provides a productive substrate interaction that probably reinforces acceptor binding.

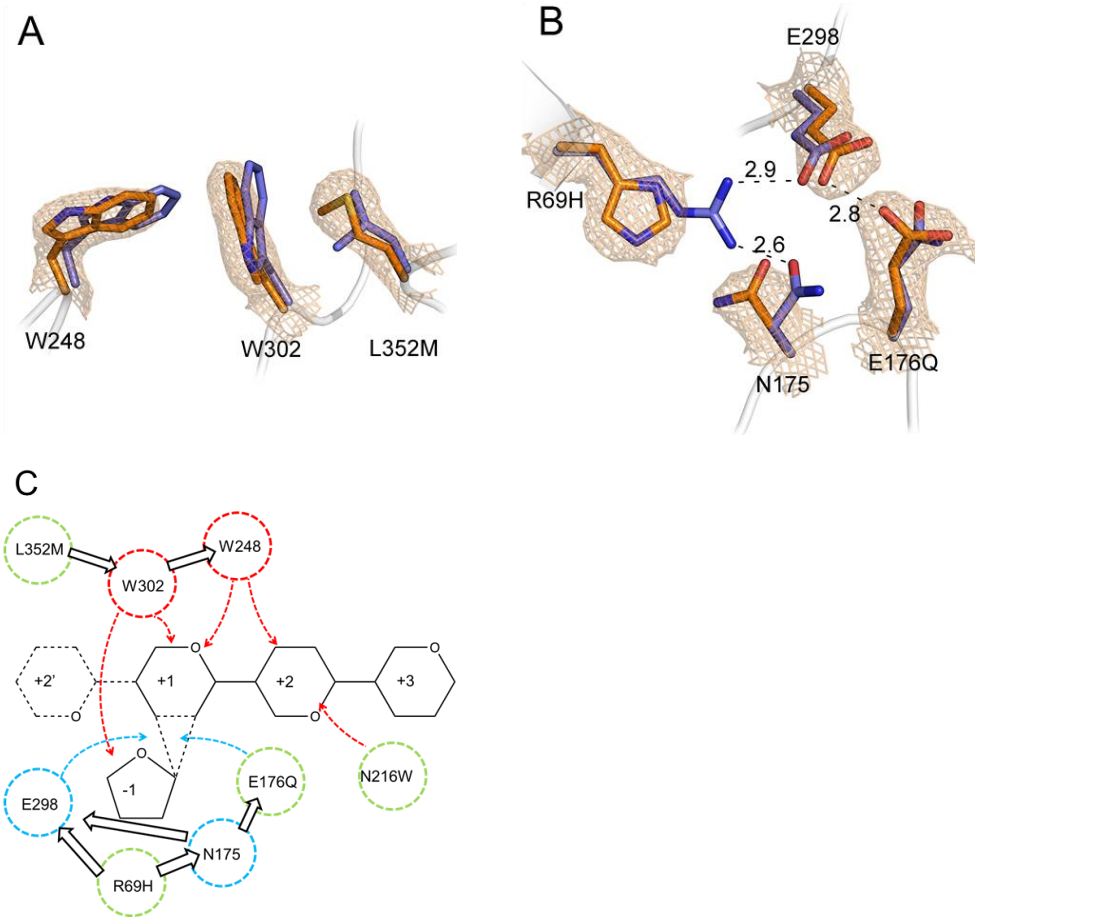


Fig. 5. (A) and (B) Superposition of *Tx*Abf (in blue, PDB code: 2VRQ) and R69H-L352M (in orange, $2F_o$-F_c electron density shown at 1.0 σ cut-off in orange mesh. H-bonds shown as black dashes with distances indicated in Å); (C) Schematic of the primary and secondary effects of the mutations. Residues indicated are: applied mutations (green); proposed to be involved in altered electrostatic interactions (blue); proposed to be involved in substrate binding (red). Coloured arrows indicate residues and locations of enzymatic activity (blue) and binding (red). Black arrows indicate secondary effects by changes in coordinates resulting from a mutation.

Crystal soaking experiments provided structures of R69H-L352M and R69H-N216W-L352M complexed with A^2XX. Analysis of these failed to reveal any substrate-induced alterations. However it is important to note that both enzymes are active and thus *in situ* cleavage of A^2XX is unavoidable. Consequently, we only detected the L-arabinofuranosyl subunit in the active site, present in two conformations (Fig. S6).

Unfortunately, the dataset of the mutant containing G179F displayed lower resolution and was thus more difficult to exploit. Nevertheless, the presence of phenylalanine at position 179 appears to create a new hydrophobic interaction platform (Fig. S7). However, the poor quality of the structure does not permit us to comment further on possible secondary structure rearrangements.

2.6. Molecular dynamics (MD) simulations on wild-type *Tx*Abf and its mutants F26L-R69H-N216W and R69H-N216W-L352M in complex with A^2XX product

Analysis of the MD simulation performed on the *Tx*Abf:A^2XX (i.e., with transglycosylation product) complex revealed that the oligosaccharide remained bound to the enzyme throughout the simulation. Binding of A^2XX was stabilized by a network of van der Waals (vdW) interactions involving F26 and Y242 in subsite -1, H98 and W99 in subsite +2', and W248 and W302 in subsite +1 (Fig. 6A). Generally, the loops surrounding the active site appear to be stabilized by the presence of the ligand, consistent with previous observations (Fig. S8) [47]. In comparison with wild-type enzyme, the MD simulation performed on F26L-R69H-N216W revealed only slight conformational differences in the vicinity of the active site. One notable exception concerns the apparently greater flexibility of the β4α4 loop that bears the catalytic acid/base E176 (Fig. 6B, Fig. S8 and Fig. S9). Regarding subsite +1, most of the ligand binding features are conserved, the main difference being the introduction of a tryptophan residue at position 216. This provides an additional stacking interaction with D-Xyl*p* moieties. Consequently, this mutation confers the means on F26L-R69H-N216W to bind xylotriose in a distinct way, different from that of wild-type *Tx*Abf (Fig. 6A-B).

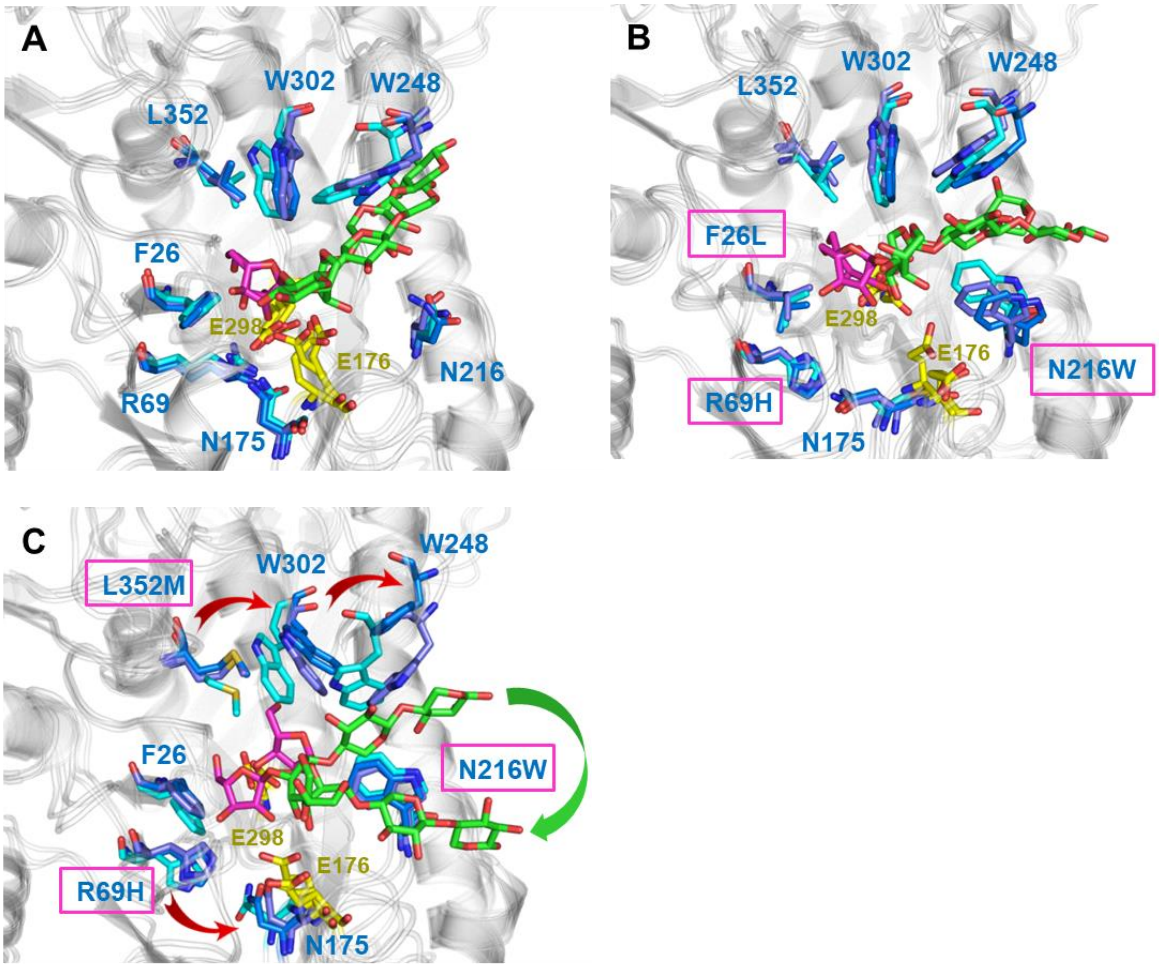


Fig. 6. Active site view of distinct conformations observed along MD simulations of (A) wild-type *Tx*Abf, (B) mutant F26L-R69H-N216W and (C) mutant R69H-N216W-L352M. Catalytic residues are colored in yellow. Selected amino acid residues are shown in stick and colored in graduated blue according to their occurrence along the MD simulation. A^2XX is shown as sticks, colored in magenta for L-Ara*f* and green for the xylotriosyl moiety, in the initial and final conformations. Mutated residues are highlighted by enframed labels.

Finally, MD simulation revealed that the combination of L352M with R69H-N216W led to greater conformational rearrangements of loops β6α6 and β7α7, compared to both wild-type *Tx*Abf and F26L-R69H-N216W (Fig. S8). On the other hand, the flexibility of loop β8α8, bearing the L352M, was identical in all the enzymes studied (Fig. S9). Together these dynamic behavioural differences increase the solvent exposure of the active site of R69H-N216W-L352M compared to wild-type *Tx*Abf. Additionally, the introduction of a methionyl moiety at position 352 appears to provoke a domino cascade that affects the conformation of residues (particularly W302 and W248) in subsites +1 and +2 (Fig. 6C). Interestingly, this

diminishes vdW interactions between the xylotriosyl moiety of A^2XX and W302 and W248, causing it to move away from loops β6α6 and β7α7. Compared to wild-type *Tx*Abf, these changes translate into a distinct binding mode involving a novel vdW interaction with the opposite face of the tryptophanyl moiety in position 216 (Fig. 6C).

3. Discussion

A previous study closely analysed the transglycosylation reaction performed by the retaining α-L-arabinofuranosidase, *Tx*Abf and provided clues on how to further alter the T/H ratio in favour of transglycosylation. Accordingly, the best performing mutant identified combines three point mutations, R69H, N216W, and L352M. In this work, we set out to further investigate the impact of L352M, comparing it with another subsite -1 mutation, F26L.

3.1. Two subsite -1 residues contribute differently to the active site

Mutations F26L and L352M display different effects on the T/H ratio of *Tx*Abf, consistent with the fact that F26, a highly conserved residue in family GH51 [34], participates exclusively to subsite -1, whereas residue L352 (less conserved) [48] is located in a more ambiguous position between subsite -1 and the acceptor subsites. Both R69H-N216W-L352M and F26L-R69H-N216W demonstrate outstanding transglycosylation ability, coupled to low catalytic efficiency (k_{cat}/K_M). This latter observation supports the postulate that both naturally occurring and engineered TGs are generally less efficient catalysts than their hydrolytic counterparts [16,26,49]. This impaired catalytic ability is probably related to more energy-demanding TS for both glycosylation and deglycosylation steps, and thus the prolonged lifetime of the covalent glycosyl-enzyme intermediate [50]. On the other hand, the extended lifetime of the glycosyl-enzyme intermediate is likely to favour statistically unfavourable deglycosylation by the glycoside acceptor (0.01 M), which is a minority species compared to water (55.6 M). In hydrolysis mode, F26L-R69H-N216W displayed a remarkable decrease in the values of both K_M and k_{cat}, suggesting that F26L severely disturbs the ability of the pocket-like subsite -1 to bind the donor substrate in a productive configuration for water-mediated deglycosylation. In contrast, the results for R69H-N216W-L352M (relatively unchanged K_M value and lower k_{cat}), clearly indicate that the impact of L352M-bearing mutant mainly affects the deglycosylation step [17]. It is noteworthy that the use of xylotriose as acceptor further demonstrates the different effects engendered by F26L and L352M. Even

though the k_{cat} values are enhanced in the presence of xylotriose, the F26L-bearing enzyme appears less able to bind xylotriose (higher K_M value for acceptor and lower R_T) than the L352M-bearing one, implying that the latter generates a positive effect with regard to xylotriose binding in positive subsites.

3.2. The conserved residue R69 plays a key role in T/H modulation

R69 is fully conserved throughout clan GH-A [51] and thus in GH51 [17,52]. Therefore, the fact that all R69H-N216W-derived mutants display similar transglycosylation profiles (including self-condensation) leads us to propose that R69H is the driving force that tips T/H ratio in favour of transglycosylation. To muster support for this postulate, it is first relevant to note that in transglycosylation mode all R69H-containing mutants display activation in the presence of acceptor ($SA_T/SA_H > 1$, Table 1). Moreover, when comparing hydrolysis and transglycosylation modes, the k_{cat} values are significantly improved for the latter. This suggests that deglycosylation by glycoside acceptors is enhanced. In this regard, structural analysis reveals that mutation of R69 indirectly influences (acting through a hydrogen bond network that involves the conserved residue N175) the catalytic acid/base E176, which is also observed in the MD simulation (Figs. B and 6C), almost certainly affecting the protonation capability of this residue. Therefore, it is highly likely that R69H impairs the ability of the enzyme to activate incoming acceptors (irrespective of whether it is a water molecule or a glycoside acceptor). In turn, we postulate that this partially diminishes the thermodynamic advantage of water (although water concentration remains a determining factor), and concomitantly increases the likelihood of deglycosylation by glycoside acceptors, which display lower deprotonation enthalpy than water [53]. Finally, it is also relevant to note that N175 is thought to be involved in transition state stabilization [37]. Therefore, the indirect modification of the interaction N175···OH-2 of α-L-Ara*f* unit by R69H might increase the TS energy barrier.

3.3. Active site flexibility modulates T/H

The results of previous work performed on a TG/hydrolytic rGH pair from family GH16 (clan GH-B) has led to the proposal that alterations in local active site flexibility (*i.e.*, plasticity and adaptation) can shift the T/H partition in favor of transglycosylation [54]. In this study, enhanced transglycosylation was attributed to increased flexibility in the donor subsite, coupled to more flexibility in the acceptor site. On the contrary, the study of GH31 TG/rGH

pair (clan GH-D) revealed that enhanced transglycosylation could be the result of conformational rigidity and the intervention of a 'hydrophobic shield' that prevents nucleophilic attack by catalytic water [55]. Finally, in a recent study increased transglycosylation was partly ascribed to higher local flexibility of the catalytic acid/base residue. According to the authors, this adds entropic cost to the height of the free energy barrier of the reaction, and thus slows down the hydrolysis step. Reciprocally, this favors (or increases the probability) of acceptor-mediated deglycosylation [56]. In our work, comparison of the root-mean-square-fluctuations (RMSF, Fig. S9) along the protein sequence clearly reveal that the best performing TG R69H-N216W-L352M displays higher flexibility in acceptor subsites, particularly in loops β6α6 and β7α7, bearing W248 and W302 respectively (Fig. S8). Considering the enhancement of the affinity (lower K_M value) for the acceptor, we postulate that this is due to greater flexibility. Additionally, R69H-N216W-L352M and F26L-R69H-N216W, which displays higher transfer rate than R69H-N216W (78 and 67 compared to 59% respectively), provides evidence that higher flexibility of the acid/base E176 might also enhance the T/H partition, which is consistent with the aforementioned findings of David et al. [56].

3.4. Positive subsite determinants

Our data reveal a significant difference in the ability of D-xylose and XOS to act as acceptors and show that XOS (DP > 2) are the best acceptors. These observations clearly point to the fact that the binding ability of the native subsite +1 alone is insufficient and that modifications, including the addition of a putative subsite +2, are essential for enhancement of transglycosylation. In this regard, the mutant R69H-N216W-L352M no doubt creates the conditions for better acceptor binding. Significantly, the presence of hydrophobic residues in aglycon-binding sites has already been identified as a key determinant of TG activity in other GHs [57,58]. In the case of the GH35 β-galactosidases from *Aspergillus niger*, it was reported that F264, Y304, and W806 constitute a dynamic hydrophobic platform that accommodates the sugar at subsite +1. In the case of *Tx*Abf and mutants thereof, it is likely that the hydrophobicity of subsite +1 is insufficient, whereas molecular modelling of the R69H-N216W-L352M:A^2XX complex indicates that both W216 and W248 are involved in stacking interactions with the subsite +2 D-Xyl*p* moiety. Finally, it is noteworthy that both the SA_T and yield decrease for XOS with DP ≥ 3 (Table S7), which implies that extra D-Xyl*p* moieties are less well-accommodated, although the presence of an undetected subsite +3 cannot be fully excluded.

3.5. Synergistic effects of mutations affect secondary hydrolysis

As a lone mutation, L352M eliminates secondary hydrolysis of A^2XX, and its combination with R69H does not alter this property [17]. However, to some extent the addition of N216W to R69H-L352M restores secondary hydrolysis, even though A^2XX synthesis is enhanced. In the present study, we confirm that while the combination of L352M with R69H-N216W decreases (2-fold) the synthesis rate, it nevertheless improves the transfer rate (R_T, 1.3-fold) and the A^2XX synthesis /secondary hydrolysis ratio (v_T/v_{HII}, 1.4-fold; Table S6). The logical conclusion of these observations is that L352M alone creates this effect, impacting the conformational flexibility of the neighbouring W302 and W248 residues and, in turn, diminishing A^2XX hydrolysis. However, this beneficial trait is probably partially countered by the effect of N216W, which apparently creates a hydrophobic platform (i.e., W302 + W248 + N216W, W178 also might be involved) that provides the basis for productive positioning for xylotriose-mediated deglycosylation toward A^2XX synthesis followed by a distinct binding mode with an altered positioning less favourable for its hydrolysis.

We assume that G179F binds α-L-Ara*f*O*p*NP more tightly thanks to the creation of a more hydrophobic environment (Fig. S7) and stronger interaction with the *p*NP moiety [17]. The kinetic data presented in the present work support this assumption, because the catalytic efficiency of hydrolysis catalysed by R69H-G179F-N216W-L352M is significantly increased (5.3-fold) compared to that of R69H-N216W-L352M, this change being driven by a decrease of the K_M value. Presumably, better donor binding by R69H-G179F-N216W-L352M leads to the acceleration of the glycosylation rate. Noticeably, even though transglycosylation yield, R_T and v_T are unfavourably affected by the introduction of G179F, secondary hydrolysis (v_{HII}) of the transglycosylation product is delayed and drastically reduced. This infers that binding of the *p*NP-bearing donor is more efficient than that of A^2XX, which only becomes significant once α-L-Ara*f*O*p*NP is almost depleted.

4. Conclusion

In the present work, we combined enzyme kinetics, 3D structure determination and *in silico* molecular dynamic simulations to further elucidate the molecular determinants that promote T/H partition in *Tx*Abf mutants. Our results provide further evidence supporting the postulate that R69H is a key T/H modulator. The introduction of histidine 69 disrupts the hydrogen bond network of both catalytic residues, thus leading to hydrolytic impotency. Additionally, we

show that L352M plays a dual effect, affecting both subsite -1 and, through a domino-like effect, the acceptor subsites. The F26L-containing mutant is a less efficient TG than the prototype R69H-N216W-L352M, because its impact is limited to subsite -1 only. Overall, our results are consistent with the idea that glycosyl-transferring GHs are sluggish enzymes (lower k_{cat}/K_M values) when compared to hydrolytic counterparts and that relaxed donor binding and acid/base flexibility can contribute to enhancement of transglycosylation. Finally, results herein are consistent with the postulate that strengthened interactions with incoming glycoside acceptors contribute to T/H shifts in favor of transglycoslation. Together these results further underline the fact that finding a generic strategy to alter the T/H partition in rGHs is not simple (but neither impossible), because the switch from hydrolysis to transglycosylation is the result of a significant number of subtle modifications to the active site.

5. Materials and Methods

5.1. Substrates and Chemicals

The substrate, 4-nitrophenyl α-L-arabinofuranoside (α-L-Ara*f*O*p*NP) was purchased from CarboSynth and xylotriose from Wako Chemicals Europe GmbH respectively. The other XOS (X_2 and X_4-X_6) were purchased from Megazyme. Molecular biology reagents were purchased from New England BioLabs.

5.2. Mutagenesis, protein expression and purification

In vitro site-directed mutagenesis was achieved using the QuikChange II Site-Directed Mutagenesis Kit (Agilent). A plasmid containing the complete coding sequence of *Tx*Abf (GenBank accession number CAA76421.2), pET24-*Tx*Abf (the original pET vector was from Novagen), was used as the DNA template for mutagenesis. Mutations were introduced using a series of oligonucleotide primers (Eurogentec). For recombination of mutations F26L, R69H, N216W G179F and L352M, the following primers (5'-3') were employed (underlined codon and mutated base in bold): F26L, 5'- CGGCCAT**TTA**TCGGAACATCTC -3'; R69H, 5'- TCCGGTCCT**CCA**CTGGCCGGGCG -3'; G179F, 5'- GGCAACGAGAACTGG**TT**CTGCG GCGGCAACAT -3'; N216W, 5'- TGCGTGCGGCGCG**TGG**ACGGCCGACTACCA -3'; L352M, 5' - CAGCTCGTCAACGTG**A**TGCAATCCGTCATCC -3'.

Double and multiple mutations were obtained by introducing stepwise the different point mutations. At each step, mutated DNA was used to transform XL1-Blue competent cells and the success of the mutagenesis protocol was systematically monitored using DNA sequencing of purified plasmid (GATC Biotech).

Expression and purification of enzymes were performed as previously described [59,60]. Briefly, a target plasmid was used to transform *E. coli* BL21(DE3) cells that were further cultured in LB medium (containing kanamycin, 50 $\mu g.mL^{-1}$) at 37 °C with shaking. For expression, IPTG (1 mM) was added to cultures of transformed *E. coli* that had reached an OD_{600} of 0.5-0.6. These were then grown for a further 4 h at 37 °C and then centrifugation (6 000×g, 15 min, 10 °C) was used to recover the cell pellets, which were suspended in 200 mM TALON buffer pH 8.0. Following a second centrifugation, pellets were resuspended in TALON buffer (20 mM Tris-HCl pH 8 with 300 mM NaCl). Enzyme crude extracts were obtained by sonication (Fisherbrand™ Model FB705 Dismembrator) of cell suspension on ice by 4 cycles of 15 s 'on', 45 s 'off', 15 s 'on' and 4 min 'off', with 40% of maximal power of the probe. Non-specific proteins were precipitated using heat treatment at 75 °C for 30 min and eliminated using centrifugation (11 000 rpm, 30 min, 10 °C) thus yielding a soluble cell extract. This extract was applied to cobalt resin (TALON® Metal Affinity Resin, Clontech) that binds *C*-terminal $(His)_6$-tagged *Tx*Abf and mutants thereof. The bound enzyme was eluted using 100 mM imidazole in TALON buffer, thus procuring fractions containing purified recombinant *Tx*Abfs that were checked using SDS-PAGE. Purified protein solutions were concentrated and desalinated using a 10 kDa cutoff Amicon® Ultra filter (Sigma-Aldrich) and 20 mM Tris-HCl buffer pH 8.0, and then stored at 4 °C. Protein concentrations were determined at 280 nm using a NanoDrop ND-1000 spectrophotometer. Theoretical molecular weight and extinction coefficients were calculated using the ExPASy ProtParam tool (https://web.expasy.org/protparam/).

5.3. Enzymatic assay

Enzyme activities were measured using a discontinuous assay [60]. Even though the optimal temperatures for activity for R69H-N216W-L352M, F26L-R69H-L352M and *Tx*Abf are 65, 70 and 75°C respectively (Fig. S1) [48,59], previous kinetic assays were arbitrarily carried out at 45°C [17], thus this temperature was used for all the enzymatic assays. Reactions operating in hydrolysis or transglycosylation mode were performed in triplicate at 45 °C, using 5 mM α-L-Ara*f*O*p*NP as donor and, when relevant, 10 mM xylotriose as acceptor (i.e., with a constant

[donor]/[acceptor] ratio of 2, thus avoiding excessive acceptor concentrations incompatible with scale-up of the reaction) prepared in 50 mM sodium phosphate buffer at pH 7.0 with 1 mg/mL BSA. Prior to enzyme addition, reaction mixtures were pre-incubated at 45 °C. Once launched, reactions were conducted for 10-20 min, removing 40 µL samples at regular intervals. These samples were immediately mixed with 200 µL of 1 M sodium carbonate and placed on ice. The release of *p*NP was monitored at 401 nm using a spectrophotometer (Infinite M200 Microplate reader, Tecan) and quantified using an appropriate standard curve that was prepared using pure *p*NP. Negative controls containing all of the reactants except the enzyme were used to correct for spontaneous hydrolysis of the donor substrate. Initial reaction rates (V_i) were determined from the linear regions of time-dependent plots by Excel, which correspond to less than 15% consumption of the donor substrate. One unit (IU) of enzyme specific activity (SA) corresponds to the amount of enzyme releasing one µmol of *p*NP per minute. SA_H and SA_T designate the specific activity measured when the reaction was operated in hydrolysis and tranglycosylation mode respectively.

The kinetic parameters K_M, k_{cat} and the efficiency coefficient k_{cat}/K_M were determined by measuring enzyme specific activity (SA) at various substrate concentrations. In hydrolysis mode, the α-L-Ara*f*O*p*NP concentration was varied from 0.05-10 mM, data obtained from the mutants were fitted to the modified Michaelis-Menten model (equation 2) by nonlinear regression plot of SA versus [S] (substrate concentration) using SigmaPlot 11.0 software. In transglycosylation mode, donor kinetic parameters were measured while maintaining the xylotriose at 10 mM. Likewise, maintaining the α-L-Ara*f*O*p*NP concentration at 3 mM while varying xylotriose concentration, from 0-100 mM, allowed us to measure the kinetic acceptor parameters. The two sets of data were fitted with Michaelis-Menten model in equation (1) to obtain the apparent SA'_{th} (SA''_{th}) and K'_M (K''_M). Concerning the two substrate reaction of Ping-Pong Bi-Bi mechanism, the modified equation (3) was applied [17,61,62]. The parameters for donor and acceptor were then calculated from the derived equation (4) and (5).

$$SA = \frac{SA_{th}[S]}{K_M + [S]} \quad (1)$$

$$SA = \frac{SA_{th}[S]}{K_M + [S]} + Ns \cdot [S] \quad (2)$$

$$SA = \frac{SA_{th}[Donor] \cdot [Acceptor]}{K_M^{Donor}[Acceptor] + K_M^{Acceptor}[Donor] + [Donor] \cdot [Acceptor]} \quad (3)$$

When fixing [Acceptor] at 10mM and varying [Donor], we obtained:

$$SA'_{th} = \frac{[Acceptor]}{K_M^{Acceptor} + [Acceptor]} \cdot SA_{th} \quad , \quad K'_M = K_M^{Donor} \cdot \frac{[Acceptor]}{K_M^{Acceptor} + [Acceptor]} \qquad (4)$$

Similarly, when fixing [Donor] at 3mM and varying [Acceptor], we obtained:

$$SA''_{th} = \frac{[Donor]}{K_M^{Donor} + [Donor]} \cdot SA_{th} \quad , \quad K''_M = K_M^{Acceptor} \cdot \frac{[Donor]}{K_M^{Donor} + [Donor]} \qquad (5)$$

SA'_{th}, K'_M, SA''_{th} and K''_M were apparent parameters in the transglycosylation mode obtained from equation (1). SA_{th} is the theoretical maximum activity achieved if the enzyme operates according to the Michaelis-Menten model.

5.4. NMR analysis

To monitor reactions ^{1}H NMR spectra were collected using a Bruker Avance II spectrometer equipped with a TCI probe and operating at 500MHz. Reactions were performed at 45 °C in the presence of α-L-Ara*f*O*p*NP and xylotriose at a ratio of 1-2 (5 and 10 mM respectively) and enzyme in a total volume of 600 µL (containing 10% D_2O in 10 mM sodium phosphate buffer at pH 7.0 with 1 mg·mL^{-1} BSA, v/v) in 5 mm NMR tubes. To initiate reactions, an aliquot of enzyme solution was added to the reaction, the concentration of enzyme (3 nM for *Tx*Abf, 0.13 µM for R69H-N216W, 0.22 µM for R69H-N216W-L352M, 0.60 µM for F26L-R69H-N216W, 1.79 µM for R69H-G179F-N216W-L352M and 0.57 µM for F26L-R69H-N216W-L352M) being adjusted to suit the 20 h reaction time frame. Time course NMR monitoring was achieved by performing pseudo-2D kinetics experiments based on a phase sensitive NOESY sequence with presaturation, with spectra being accumulated every 8.7 mins (2 × 32 scans). Using previously calculated specific activities it was possible to adjust enzyme quantities such that reactions occurred up to maximum of 18 h, although the actual duration of the reaction was dependent on the enzyme used. The pH in 10% D_2O was measured using a glass pH electrode and the modified equation $pH_{(10\%\ D_2O)} = pH_{electrode} + 0.04$ [63]. The chemical shift reference was based on the HOD signal calibrated at 4.55 ppm at 45 °C [64]. Experiments to investigate the transglycosylation ability depending on XOS length were performed on Bruker Ascend Advance III spectrometer at 800 MHz equipped with a TCI CryoProbe. Donor and acceptor concentrations were maintained at 5 and 10 mM respectively, but 50 µL reaction solution was used within 1.7 mm NMR tubes. Spectra were accumulated every 3.7 mins (32 scans).

Donor (α-L-Ara*f*O*p*NP) consumption and the apparition of transglycosylation products (AXOS) were quantified by integrating the relevant anomeric proton signals from the internal

anomeric proton from α-L-Ara*f* unit (5.78 ppm for donor, products were shown in Table 4), and normalized by the initial integral of α-L-Ara*f*O*p*NP. The AXOS yields (in %) were determined by product concentration against initial donor concentration. The transglycosylation rate (R_T) of the donor substrate onto xylotriose acceptor (μmol of formed A^2XX/μmol of consumed α-L-Ara*f*O*p*NP) was derived from the plot of A^2XX transglycosylation yield as a function of donor conversion. Conveniently, being independent of the duration of the reaction, R_T indicates the transglycosylation proportion in the consumed donor and allows to compare different enzymes that display both transglycosylation and hydrolysis.

5.5. Crystallographic structure determination

Crystals of *Tx*Abf mutants R69H-L352M, R69H-N216W-L352M and R69H-G179F-L352M were grown at RT using the sitting-drop vapor diffusion method. Crystallization drops were set up in MRC two-well plates using an Oryx-8 crystallization robot (Douglas Instruments). Crystallization conditions were 0.2 M ammonium acetate, 0.1 M bis-tris pH 5.5 and 45% MPD, or 0.2 M $MgCl_2$, 0.1 M HEPES pH 7.5 and 34% (w/v) PEG400 with protein concentrations of 14.2 mg/ml (R69H-L352M), 10.7 mg/ml (R69H-G179F-L352M) and 10.9 mg/ml (R69H-N216W-L352M) in 1 mM Tris-HCl pH 8.0, 150 mM NaCl (Table S8). Concentrations were estimated based on A_{280} using a NanoDrop-1000 (Thermofisher Scientific). Crystals were grown in 0.4 μl drops in a volume ratio of 3:1 (protein to reservoir). Crystals were reproducible in these conditions, and appeared within a few weeks. No cryoprotectant was added before cooling the crystals in liquid nitrogen.

In an effort to form crystallographic complex structures, soaking experiments were performed at both room temperature and 4 °C, yielding similar results in refinement. Crystals of mutated variants R69H-L352M and R69H-N216W-L352M were soaked by adding a small volume (2-5 μL) of mother-liquor containing 33.0 mM to 330.0 mM of either A^2XX or xylobiose.

Diffraction data were collected at MAX-lab, Sweden, or at the ESRF, France at the beamlines indicated in Table S8. The data were processed and scaled using either MOSFLM and AIMLESS [65,66], or XDS and XSCALE [67], all in the hexagonal space group $P6_522$. Refinement was performed in Refmac5 [68] of the CCP4 suite [69], with an initial rigid body refinement, using the structure and transferred R_{Free} data of the E176Q:XA^3XX structure, without substrate or solvent molecules. Rigid body refinement was followed by restrained refinement using NCS restraints. Each round of refinement was followed by manual structure

remodeling and validation using COOT [70]. Data collection and refinement statistics are presented in Table S8.

5.6. Molecular modelling procedures

Starting from the high resolution crystal structures of wild-type *Tx*Abf (PDB code: 2VRQ) [41] and herein disclosed mutant R69H-N216W-L352M (PDB code: to be defined), 3D models were constructed for WT, R69H-N216W-L352M and F26L-N216W-L352M enzymes in complex with A^2XX. Missing residues 86-107 of βα2 loop, in structures of mutants were added by comparison with crystallographic structure of wild-type enzyme. Mutations not observed in X-ray structures were added using Pymol Molecular Graphics System (Schrodinger, Portland, OR, USA) [71]. A^2XX was constructed using *tleap* program in AMBER software package (Case et al. 2016) and manually docked in active site of enzymes using as template co-crystallized E176Q:XA^3XX *Tx*Abf structure (PDB ID 2VRQ). In this conformation, L-Ara*f* moiety was placed in subsite -1 while D-Xyl*p* units were positioned in +2', +1, +2 subsites. MD simulations were subsequently performed with the AMBER ff14SB force-field [72] for enzymes and GLYCAM_06j [73] for carbohydrate ligands. To obtain a neutral charge of the simulated systems, counter-ions were included. Each enzyme or enzyme: ligand complex together with the counter-ions was solvated with TIP3P water molecules, using the rectangular parallelepiped box with a minimum distance of 0.12 nm between the solute and the simulation box edge. All molecular systems were then subjected to energy minimization using the AMBER16 suite of programs [74]. Following the minimization steps, MD simulations were carried out upon slow heating to 300 K under constant volume over a period of 100 ps. At the final required temperature (300 K), the system was equilibrated under constant volume conditions over 20 ps and then it was turned to constant pressure (1 bar) conditions over 80 ps. Harmonic constraints were first applied on the protein backbone and ligand and subsequently, they were gradually removed along the MD preparation schedule. The final production phase of simulations was then carried out for a total of 20 ns at constant temperature (300 K) and pressure (1 bar) conditions. The resulting trajectories were analysed using the cpptraj module of the AMBER16 package. The rmsf was calculated for the protein backbone atoms. Graphics were prepared using Pymol Molecular Graphics System, (Schrodinger, Portland, OR, USA) [71].

6. Acknowledgments

The PhD fellowship of J. Zhao was supported by CSC (China Scholarship Council). This work was performed in the framework of the French Danish research collaboration Program (IFD, N° 15/2015/CSU.8.2.1). The NMR analyses were performed using facilities at MetaToul (Metabolomics & Fluxomics Facilities, Toulouse, France, www.metatoul.fr), which is part of the national infrastructure MetaboHUB (The French National infrastructure for metabolomics and fluxomics, www.metabohub.fr) and is supported by grants from the Région Midi-Pyrénées, the European Regional Development Fund, SICOVAL, IBiSa-France, CNRS, and INRA. G. Lippens and N. Cox (TBI) are gratefully acknowledged for insightful discussions and technical development of the NMR pseudo-2D kinetics experiments. We also thank Jens-Christian N. Poulsen for the contribution in the crystallization work.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

R.F. and M.J.O designed and supervised the study. J.Z. performed the biochemical experiment. B.B. and R.F. produced the enzymes for crystallographic studies. T.T., C.D. and L.L.L carried out crystallographic studies. S.B. and I.A. performed the molecular dynamics simulation study. All authors analyzed data and contributed to write the manuscript.

References

1 Danby PM & Withers SG (2016) Advances in Enzymatic Glycoside Synthesis. *ACS Chem. Biol.* **11**, 1784–1794.

2 Krasnova L & Wong CH (2019) Oligosaccharide Synthesis and Translational Innovation. *J. Am. Chem. Soc.* **141**, 3735–3754.

3 Trincone A & Giordano A (2006) Glycosyl Hydrolases and Glycosyltransferases in the Synthesis of Oligosaccharides. *Curr. Org. Chem.* **10**, 1163–1193.

4 Filice M & Marciello M (2013) Enzymatic Synthesis of Oligosaccharides: A Powerful Tool for a Sweet Challenge. *Curr. Org. Chem.* **17**, 701–718.

5 Benkoulouche M, Fauré R, Remaud-Siméon M, Moulis C & André I (2019) Harnessing glycoenzyme engineering for synthesis of bioactive oligosaccharides. *Interface Focus* **9**.

6 Field RA (2011) Challenging reaction equilibria. *Nat. Chem. Biol.* **7**, 658–659.

7 Chang A, Singh S, Phillips GN & Thorson JS (2011) Glycosyltransferase structural biology and its role in the design of catalysts for glycosylation. *Curr. Opin. Biotechnol.* **22**, 800–

808.

8 Nidetzky B, Gutmann A & Zhong C (2018) Leloir Glycosyltransferases as Biocatalysts for Chemical Production. *ACS Catal.* **8**, 6283–6300.

9 Eisele A, Zaun H, Kuballa J & Elling L (2018) In Vitro One-Pot Enzymatic synthesis of hyaluronic acid from sucrose and N-Acetylglucosamine: optimization of the enzyme module system and nucleotide sugar regeneration. *ChemCatChem* **10**, 2969–2981.

10 Mestrom L, Przypis M, Kowalczykiewicz D, Pollender A, Kumpf A, Marsden SR, Bento I, Jarzębski AB, Szymańska K, Chruściel A, Tischler D, Schoevaart R, Hanefeld U & Hagedoorn PL (2019) *Leloir glycosyltransferases in applied biocatalysis: A multidisciplinary approach.*

11 Kallolimath S, Gruber C, Steinkellner H & Castilho A (2018) Promoter Choice Impacts the Efficiency of Plant Glyco-Engineering. *Biotechnol. J.* **13**, 1–7.

12 Koshland DE (1953) Stereochemistry and the mechanism of enzymatic reactions. *Biol. Rev.* **28**, 416–436.

13 Baumann MJ, Eklöf JM, Michel G, Kallas ÅM, Teeri TT, Czjzek M & Brumer H (2007) Structural evidence for the evolution of xyloglucanase activity from xyloglucan endo - transglycosylases: biological implications for cell wall metabolism. *Plant Cell* **19**, 1947–1963.

14 Passerini D, Vuillemin M, Ufarté L, Morel S, Loux V, Fontagné-Faucher C, Monsan P, Remaud-Siméon M & Moulis C (2015) Inventory of the GH70 enzymes encoded by Leuconostoc citreum NRRL B-1299 - Identification of three novel α-transglucosylases. *FEBS J.* **282**, 2115–2130.

15 Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM & Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, D490–D495.

16 Bissaro B, Monsan P, Fauré R & O'Donohue MJ (2015) Glycosynthesis in a waterworld: new insight into the molecular basis of transglycosylation in retaining glycoside hydrolases. *Biochem. J.* **467**, 17–35.

17 Bissaro B, Durand J, Planas A, Monsan P, Biarnés X, Planas A, Monsan P, O'Donohue MJ, Fauré R, Donohue MJO, Fauré R, O'Donohue MJ, Fauré R, O'Donohue MJ & Fauré R (2015) Molecular design of non-Leloir furanose-transferring enzymes from an α-L-arabinofuranosidase: a rationale for the engineering of evolved transglycosylases. *ACS Catal.* **5**, 4598–4611.

18 Moracci M, Trincone A, Perugino G, Ciaramella M & Rossi M (1998) Restoration of the activity of active-site mutants of the hyperthermophilic β-glycosidase from Sulfolobus solfataricus: Dependence of the mechanism on the action of external nucleophiles. *Biochemistry* **37**, 17262–17270.

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

20 Mackenzie LF, Wang Q, Warren RAJ & Withers SG (1998) Glycosynthases: mutant glycosidases for oligosaccharide synthesis. *J. Am. Chem. Soc.* **120**, 5583–5584.

21 Mayer C, Zechel DL, Reid SP, Warren RAJ & Withers SG (2000) The E358S mutant of *Agrobacterium sp*. β-glucosidase is a greatly improved glycosynthase. *FEBS Lett.* **466**,

40–44.

22 Huber RE, Gaunt MT, Sept RL & Babiak MJ (1983) Differences in the effects of pH on the hydrolytic and transgalactosylic reactions of β-galactosidase (*Escherichia coli*). *Can. J. Biochem. Cell Biol.* **61**, 198–206.

23 Abdul Manas NH, Md. Illias R & Mahadi NM (2018) Strategy in manipulating transglycosylation activity of glycosyl hydrolase for oligosaccharide production. *Crit. Rev. Biotechnol.* **38**, 272–293.

24 Wang LX & Huang W (2009) Enzymatic transglycosylation for glycoconjugate synthesis. *Curr. Opin. Chem. Biol.* **13**, 592–600.

25 Madhuprakash J, Dalhus B, Rani TS, Podile AR, Eijsink VGH & Sørlie M (2018) Key residues affecting transglycosylation activity in family 18 chitinases: Insights into Donor and Acceptor Subsites. *Biochemistry* **57**, 4325–4337.

26 Teze D, Hendrickx J, Czjzek M, Ropartz D, Sanejouand Y-H, Tran V, Tellier C & Dion M (2014) Semi-rational approach for converting a GH1 β-glycosidase into a β-transglycosidase. *Prot. Eng. Des. Sel.* **27**, 13–19.

27 Rivera MH, López-Munguía A, Soberón X & Saab-Rincón G (2003) α-amylase from Bacillus licheniformis mutants near to the catalytic site: Effects on hydrolytic and transglycosylation activity. *Protein Eng.* **16**, 505–514.

28 Zakariassen H, Hansen MC, Jøranli M, Eijsink VGH & Sørlie M (2011) Mutational Effects on Transglycosylating Activity of Family 18 Chitinases and Construction of a Hypertransglycosylating Mutant. *Biochemistry* **50**, 5693–5703.

29 Kuriki T, Kaneko H, Yanase M, Takata H, Shimada J, Handa S, Takada T, Umeyama H & Okada S (1996) Controlling substrate preference and transglycosylation activity of neopullulanase by manipulating steric constraint and hydrophobicity in active center. *J. Biol. Chem.* **271**, 17321–17329.

30 Teze D, Hendrickx J, Dion M, Tellier C, Woods VL, Tran V & Sanejouand Y-H (2013) Conserved Water Molecules in Family 1 Glycosidases: A DXMS and Molecular Dynamics Study. *Biochemistry* **52**, 5900–5910.

31 David B, Irague R, Jouanneau D, Daligault F, Czjzek M, Sanejouand Y-H & Tellier C (2017) Internal Water Dynamics Control the Transglycosylation/Hydrolysis Balance in the Agarase (AgaD) of Zobellia galactanivorans. *ACS Catal.* **7**, 3357–3367.

32 Florindo RN, Souza VP, Mutti HS, Camilo C, Manzine LR, Marana SR, Polikarpov I & Nascimento AS (2018) Structural insights into β-glucosidase transglycosylation based on biochemical, structural and computational analysis of two GH1 enzymes from *Trichoderma harzianum*. *N. Biotechnol.* **40**, 218–227.

33 Teze D, Daligault F, Ferrières V, Sanejouand YH & Tellier C (2015) Semi-rational approach for converting a GH36 α-glycosidase into an α-transglycosidase. *Glycobiology* **25**, 420–427.

34 Teze D, Zhao J, Wiemann M, Kazi ZG, Lupo R, Rønne ME, Carlström G, Duus JØ, Sanejouand Y-H, O'Donohue MJ, Karlsson EN, Fauré R, Stålbrand H & Svensson B (2020) Rational enzyme design without structural knowledge: a sequence-based approach for efficient generation of glycosylation catalysts. *ChemRxiv*.

35 Jamek SB, Muschiol J, Holck J, Zeuner B, Busk PK, Mikkelsen JD & Meyer AS (2018) Loop protein engineering for improved transglycosylation activity of a β-N-

Acetylhexosaminidase. *ChemBioChem* **19**, 1858–1865.

36 Romero-Téllez S, Lluch JM, González-Lafont À & Masgrau L (2019) Comparing hydrolysis and transglycosylation reactions catalyzed by thermus thermophilus β-glycosidase. a combined md and QM/MM study. *Front. Chem.* **7**, 1–9.

37 Raich L, Borodkin V, Fang W, Castro-López J, van Aalten DMF, Hurtado-Guerrero R & Rovira C (2016) A Trapped Covalent Intermediate of a Glycoside Hydrolase on the Pathway to Transglycosylation. Insights from Experiments and Quantum Mechanics/Molecular Mechanics Simulations. *J. Am. Chem. Soc.* **138**, 3325–3332.

38 Fauré R, Courtin CM, Delcour JA, Dumon C, Faulds CB, Fincher GB, Fort S, Fry SC, Halila S, Kabel MA, Pouvreau L, Quemener B, Rivet A, Saulnier L, Schols HA, Driguez H & O'Donohue MJ (2009) A brief and informationally rich naming system for Oligosaccharide motifs of heteroxylans found in plant cell walls. *Aust. J. Chem.* **62**, 533–537.

39 McIntosh LP, Hand G, Johnson PE, Joshi MD, Körner M, Plesniak LA, Ziser L, Wakarchuk WW & Withers SG (1996) The p*K*a of the general acid/base carboxyl group of a glycosidase cycles during catalysis: a 13C-NMR study of bacillus circulans xylanase. *Biochemistry* **35**, 9958–66.

40 Arab-Jaziri F, Bissaro B, Dion M, Saurel O, Harrison D, Ferreira F, Milon A, Tellier C, Fauré R & O'Donohue MJ (2013) Engineering transglycosidase activity into a GH51 α-L-arabinofuranosidase. *New. Biotechnol.* **30**, 536–544.

41 Paës G, Skov LK, O'Donohue MJ, Rémond C, Kastrup JS, Gajhede M & Mirza O (2008) The structure of the complex between a branched pentasaccharide and Thermobacillus xylanilyticus GH-51 arabinofuranosidase reveals xylan-binding determinants and induced fit. *Biochemistry* **47**, 7441–51.

42 Perugino G, Trincone A, Rossi M & Moracci M (2004) Oligosaccharide synthesis by glycosynthases. *Trends Biotechnol.* **22**, 31–37.

43 Feng HY, Drone J, Hoffmann L, Tran V, Tellier C, Rabiller C & Dion M (2005) Converting a β-glycosidase into a β-transglycosidase by directed evolution. *J. Biol. Chem.* **280**, 37088–37097.

44 Durand J, Biarnés X, Watterlot L, Bonzom C, Borsenberger V, Planas A, Bozonnet S, O'Donohue MJ & Fauré R (2016) A single point mutation alters the transglycosylation/hydrolysis partition, significantly enhancing the synthetic capability of an endo -glycoceramidase. *ACS Catal.* **6**, 8264–8275.

45 Viëtor RJ, Hoffmann RA, Angelino SAGF, Voragen AGJ, Kamerling JP & Vliegenthart JFG (1994) Structures of small oligomers liberated from barley arabinoxylans by endoxylanase from *Aspergillus-Awamori*. *Carbohydr. Res.* **254**, 245–255.

46 Gruppen H, Hoffmann R., Kormelink FJ., Voragen AGJ, Kamerling J. & Vliegenthart JFG (1992) Characterisation by 1H NMR spectroscopy of enzymically derived oligosaccharides from alkali-extractable wheat-flour arabinoxylan. *Carbohydr. Res.* **233**, 45–64.

47 Arab-Jaziri F, Bissaro B, Barbe S, Saurel O, Débat H, Dumon C, Gervais V, Milon A, André I, Fauré R & O'Donohue MJ (2012) Functional roles of H98 and W99 and β2α2 loop dynamics in the α-L-arabinofuranosidase from Thermobacillus xylanilyticus. *FEBS J.* **279**, 3598–3611.

48 Arab-Jaziri F, Bissaro B, Tellier C, Dion M, Fauré R & O'Donohue MJ (2015) Enhancing

the chemoenzymatic synthesis of arabinosylated xylo-oligosaccharides by GH51 α-L-arabinofuranosidase. *Carbohydr. Res.* **401**, 64–72.

49 Luang S, Cho J Il, Mahong B, Opassiri R, Akiyama T, Phasai K, Komvongsa J, Sasaki N, Hua YL, Matsuba Y, Ozeki Y, Jeon JS & Ketudat Cairns JR (2013) Rice Os9BGlu31 is a transglucosidase with the capacity to equilibrate phenylpropanoid, flavonoid, and phytohormone glycoconjugates. *J. Biol. Chem.* **288**, 10111–10123.

50 Piens K, Fauré R, Sundqvist G, Baumann MJ, Saura-Valls M, Teeri TT, Cottaz S, Planas A, Driguez H & Brumer H (2008) Mechanism-based labeling defines the free energy change for formation of the covalent glycosyl-enzyme intermediate in a xyloglucan endo-transglycosylase. *J. Biol. Chem.* **283**, 21864–21872.

51 Zhang Y, Ju J, Peng H, Gao F, Zhou C, Zeng Y, Xue Y, Li Y, Henrissat B, Gao GF & Ma Y (2008) Biochemical and structural characterization of the intracellular mannanase AaManA of *Alicyclobacillus acidocaldarius* reveals a novel glycoside hydrolase family belonging to clan GH-A. *J. Biol. Chem.* **283**, 31551–31558.

52 Borsenberger V, Dornez E, Desrousseaux M-L, Massou S, Tenkanen M, Courtin CM, Dumon C, O'Donohue MJ & Fauré R (2014) A 1H NMR study of the specificity of α-L-arabinofuranosidases on natural and unnatural substrates. *Biochim. Biophys. Acta - Gen. Subj.* **1840**, 3106–3114.

53 Geronimo I, Payne CM & Sandgren M (2018) The role of catalytic residue p*K*a on the hydrolysis/transglycosylation partition in family 3 β-glucosidases. *Org. Biomol. Chem.* **16**, 316–324.

54 Mark P, Baumann MJ, Eklöf JM, Gullfot F, Michel G, Kallas ÅM, Teeri TT, Brumer H & Czjzek M (2009) Analysis of nasturtium Tm NXG1 complexes by crystallography and molecular dynamics provides detailed insight into substrate recognition by family GH16 xyloglucan endo -transglycosylases and endo -hydrolases. *Proteins Struct. Funct. Bioinforma.* **75**, 820–836.

55 Light SH, Cahoon LA, Mahasenan K V., Lee M, Boggess B, Halavaty AS, Mobashery S, Freitag NE & Anderson WF (2017) Transferase versus hydrolase: The role of conformational flexibility in reaction specificity. *Structure* **25**, 1–10.

56 David B, Arnaud P, Tellier C & Sanejouand Y-H (2019) Toward the design of efficient transglycosidases: the case of the GH1 of *Thermus thermophilus*. *Protein Eng. Des. Sel.*, 1–8.

57 Umemoto N, Ohnuma T, Osawa T, Numata T & Fukamizo T (2015) Modulation of the transglycosylation activity of plant family GH18 chitinase by removing or introducing a tryptophan side chain. *FEBS Lett.* **589**, 2327–2333.

58 Rico-Díaz A, Ramírez-Escudero M, Vizoso-Vázquez Á, Cerdán ME, Becerra M & Sanz-Aparicio J (2017) Structural features of *Aspergillus niger* β-galactosidase define its activity against glycoside linkages. *FEBS J.* **284**, 1815–1829.

59 Debeche T, Cummings N, Connerton I, Debeire P & O'Donohue MJ (2000) Genetic and Biochemical Characterization of a Highly Thermostable alpha-L-Arabinofuranosidase from Thermobacillus xylanilyticus. *Appl. Environ. Microbiol.* **66**, 1734–1736.

60 Bissaro B, Saurel O, Arab-jaziri F, Saulnier L, Milon A, Tenkanen M, Monsan P, O'Donohue MJ & Fauré R (2014) Mutation of a pH-modulating residue in a GH51 α-L-arabinofuranosidase leads to a severe reduction of the secondary hydrolysis of transfuranosylation products. *Biochim. Biophys. Acta - Gen. Subj.* **1840**, 626–636.

61 Mannervik B (1973) A branching reaction mechanism of glutathione reductase. *Biochem. Biophys. Res. Commun.* **53**, 1151–1158.

62 Cleland WW (1967) Enzyme Kinetics. *Annu. Rev. Biochem.* **36**, 77–112.

63 Glasoe PK & Long FA (1960) Use of glass electrodes to measure acidities in deuterium oxide 1,2. *J. Phys. Chem.* **64**, 188–190.

64 Gottlieb HE, Kotlyar V & Nudelman A (1997) NMR chemical shifts of common laboratory solvents as trace impurities. *J. Org. Chem.* **62**, 7512–7515.

65 Battye TGG, Kontogiannis L, Johnson O, Powell HR & Leslie AGW (2011) iMOSFLM : a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 271–281.

66 Evans PR & Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69**, 1204–1214.

67 Kabsch W (2010) XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 125–132.

68 Murshudov GN, Vagin AA & Dodson EJ (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53**, 240–255.

69 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A & Wilson KS (2011) Overview of the CCP 4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 235–242.

70 Emsley P & Cowtan K (2004) Coot : model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2126–2132.

71 DeLano, W.L. (2002) The PyMOL Molecular Graphics System. Delano Scientific, San Carlos.

72 Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE & Simmerling C (2015) ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713.

73 Kirschner KN, Yongye AB, Tschampel SM, González-Outeiriño J, Daniels CR, Foley BL & Woods RJ (2008) GLYCAM06: A generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.* **29**, 622–655.

74 Case DA, Betz RM, Cerutti DS, Cheatham Iii TE, Darden TA, Duke RE, Giese TJ, Gohlke H, Goetz AW, Homeyer N, Izadi S, Janowski P, Kaus J, Kovalenko A, Lee TS, LeGrand.S, Li, P.C., Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, H.T. Nguyen I, Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith JS & R.C. Walker, J. Wang, R.M. Wolf, X. Wu LX and PAK (2016) AMBER 2016. *Univ. California, San Fr.* **810**.

Supporting Information

Probing the determinants of the transglycosylation/hydrolysis partition in a retaining α-L-arabinofuranosidase

Jiao Zhao[1], Tobias Tandrup[2], Bastien Bissaro[1], Sophie Barbe[1], Isabelle André[1], Claire Dumon[1], Leila Lo Leggio[2], Michael J. O'Donohue[1] and Régis Fauré[1]

[1] TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France

[2] University of Copenhagen, Department of Chemistry, Copenhagen, Denmark

Table S1. Kinetic parameters on α-L-Ara*f*O*p*NP (donor) in hydrolysis mode and transglycosylation mode (data shown in bracket) at pH 5.8.[a]

Enzyme	SA_{th} ($IU.mg^{-1}$)	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($s^{-1}.mM^{-1}$)	Ns ($s^{-1}.mM^{-1}$)
R69H-L352M-N216W	(6.37)	(1.4)	(6.12)	(4.37)	(-)
F26L-R69H-N216W	0.19 (3.43)	0.02 (0.6)	0.18 (3.29)	7.42 (5.49)	0.036 (-)

[a]Enzymatic assays were performed at 45 °C in 50 mM sodium acetate buffer pH 5.8. Experiments were performed in triplicate and relative errors were always less than 10%. In hydrolysis mode, the modified Michaelis-Menten model represented by equation (2) was applied. In transglycosylation mode, the concentration of xylotriose acceptor was fixed at 10 mM and the two-substrate equation (3) was applied.

Table S2. Kinetic parameters on xylotriose (acceptor) in transglycosylation mode at pH 5.8.[a]

Enzyme	SA_{th} ($IU.mg^{-1}$)	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($s^{-1}.mM^{-1}$)
R69H-L352M-N216W	6.38	15.98	6.13	0.39
F26L-R69H-N216W	4.36	44.77	4.19	0.09

[a]Enzymatic assays were performed at 45 °C in 50 mM sodium acetate buffer pH 5.8. Experiments were performed in triplicate and relative errors were always less than 10%. The concentration of α-L-Ara*f*O*p*NP (donor) was fixed at 3 mM and the two-substrate equation (3) was applied.

Table S3. A^2XX transglycosylation yield as function of operating pH for reactions catalyzed by *Tx*Abf triple mutants.[a]

Enzyme	Maximal A^2XX yield (%)	
	pH 5.8	pH 7.0
	5.20 ppm[b]	
R69H-N216W-L352M	68	70
F26L-R69H-N216W	58	59

[a]The enzymatic reactions were performed using 5 mM α-L-Ara*f*O*p*NP as donor and 10 mM of xylotriose as acceptor at 45 °C and pH 7.0 (10 mM sodium phosphate buffer) or pH 5.8 (10 mM sodium acetate buffer) in 10% D_2O with 0.22 μM of R69H-N216W-L352M and 0.3 μM of F26L-R69H-N216W.

[b]NMR chemical shift of the α-L-Ara*f* unit anomeric proton of A^2XX was determined at 45°C.

Table S4. Influence of the acceptor/donor substrate ratio on the A2XX transglycosylation yields.a

Substrate		[Acceptor]/[Donor] ratio	A^2XX yield (%)
[α-L-Ara*f*O*p*NP donor] (mM)	[Xylotriose acceptor] (mM)		
3	1.5	0.5	23
3	3	1	37
5	10	2	59
3	10	3.33	85
1	10	10	84

[a]Assays were carried out at 45 °C and pH 7.0 in buffered 10% D_2O and using 0.3 μM of F26L-R69H-N216W.

Table S5. Influence of the substrate concentrations on the A^2XX transglycosylation yields.[a]

Substrate		[Acceptor]/[Donor] ratio[b]	A^2XX yield (%)
[α-L-Ara*f*O*p*NP donor] (mM)	[Xylotriose acceptor] (mM)		
5	10	2	59
15	30		64

[a]Assays were carried out at 45 °C and pH 7.0 in buffered 10% D_2O, using 0.3 µM of F26L-R69H-N216W.

[b]Althought their concentrations varied, the ratio of [Acceptor]/[Donor] substrates was kept constant at 2.

Table S6. Apparent transglycosylation and secondary hydrolysis parameters for *Tx*Abf mutants.

Enzyme	R_T (%)[a]	Donor consumption at tipping point (%)[b]	R^2 value for linearity phase[c]	Synthesis rate v_T (mM.h^{-1}.µM^{-1})	2nd Hydrolysis rate v_{HII} (mM.h^{-1}.µM^{-1})	v_T/v_{HII}[d]
R69H-N216W	59	82	0.9974	7.30	1.11	6.6
R69H-N216W-L352M	78	80	0.9980	3.69	0.41	9.0
F26L-R69H-N216W	67	84	0.9939	1.00	0.20	5.1
R69H-G179F-N216W-L352M	59	98	0.9952	0.50	0.03	15.5
F26L-R69H-N216W-L352M	74	82	0.9994	1.70	0.16	10.9

[a]L-Ara*f* transfer rate (R_T) corresponds to the initial slope of the donor substrate consumption-dependent plot of the transglycosylation product yield (Fig. 4B).

[b]The tipping point matches to the donor consumption related with the linearity loss of the plot transglycosylation A^2XX yield/consumed donor (in %) *versus* donor consumption (Fig. 4B).

[c]The root-mean-square deviations (R^2) ascertains the accuracy of the linearity phase of the regression curve before the tipping point in Fig. 4B.

[d]Apparent transglycosylation rate (v_T, initial A^2XX synthesis rate per µM of enzyme) and secondary hydrolysis rate (v_{HII}, A^2XX hydrolysis rate after the maximum yield per µM of enzyme) were derived from the time-course monitoring of A^2XX transglycosylation yield in Fig. 4A. The ratio v_T/v_{HII} depicts the ability of the enzymes to perform the synthesis of A^2XX while considering their secondary hydrolysis capability.

Table S7. Transglycosylation yields for reactions catalyzed by R69H-N216W-L352M *Tx*Abf using different XOS as acceptor.[a]

XOS	SA_T[b] ($IU.mg^{-1}$)	Main product	Maximal yield (%)
X	0.73	A^n	5
X_2	2.33	A^2X	75
X_3	1.52	A^2XX	70
X_4	1.40	A^2XXX	70
X_5	1.34	A^2XXXX	60
X_6	1.34	A^2XXXXX	63

[a]Reactions were performed at 45 °C and pH 7.0 in buffered 10% D_2O containing 0.49 µM R69H-N216W-L352M (except for xylotriose where 0.23 µM enzyme was used), 5 mM α-L-Ara*f*O*p*NP as donor and 10 mM XOS as acceptors. The experiment containing xylotriose was monitored by NMR at 500 MHz, whereas all the other experiments were performed at 800 MHz.

[b]SA_T (transglycosylation mode) values were measured by following the donor (α-L-Ara*f*O*p*NP) consumption, monitoring the disappearance of the anomeric proton signal using ^1H NMR. The SA measurments acquired using NMR and UV spectroscopy (i.e., *p*NP release) provided consitent data (data not shown).

Table S8. Crystal structure data collection and refinement statistics.

Crystallization conditions, data collection and refinement statistics	R69H-L352M	R69H-L352M:A2XX	R69H-N216W-L352M	R69H-N216W-L352M:A2XX	R69H-G179F-L352M
Crystallization conditions	0.2 M ammonium acetate, 0.1 M Bis-tris pH 5.5, 45% (v/v) MPD	0.2 M ammonium acetate, 0.1 M Bis-tris pH 5.5, 45% (v/v) MPD	0.2 M $MgCl_2$, 0.1 M HEPES pH 7.5, 34% (v/v) PEG400	0.2 M ammonium acetate, 0.1 M Bis-tris pH 5.5, 45% (v/v) MPD	0.2 M sodium thiocyanate, 20% (v/v) PEG3350
Soaking substrate	-	66 mM	-	220 mM	-
Data collection					
Beamline	I911-3	ID30B	I911-3	ID30B	ID23-1
Wavelength [Å]	0.999	0.969	0.999	0.969	0.976
Space group	$P6_522$	$P6_522$	$P6_522$	$P6_522$	$P6_522$
Unit cell dimensions	156.13	156.69	156.57	157.06	156.49
(a, b, c)[Å]	156.13	156.69	156.57	157.06	156.49
α=90.0, β=90.0, γ=120.0	377.12	378.66	376.2	379.59	376.57
Resolution [Å]	76.46-2.3 (2.34-2.30)	50.0-1.85 (1.90-1.85)	92.11-2.8 (2.87-2.80)	50.0-2.00 (2.00-2.05)	50.0-3.10 (3.18-3.10)
Completeness [%]	99.9 (99.4)	99.7 (97.5)	99.9 (99.9)	99.9 (100.0)	99.0 (100.0)
R_{MEAS} [%]	21.6 (112.0)	15.4 (167.4)	44.9 (336.5)	25.7 (286.7)	86.4 (310.3)
I/σ(I)	14.5 (2.7)	14.17 (1.59)	7.27 (0.82)	11.19 (1.25)	4.81 (1.54)
$CC_{1/2}$ [%]	98.8 (72.7)	99.8 (57.6)	97.9 (25.5)	99.8 (50.3)	98.5 (50.1)
Unique reflections	120665 (5811)	231584 (16556)	67727 (4883)	184518 (13516)	50293 (3668)
Observed reflections	1894792 (91107)	4406879 (228093)	728372 (53399)	4026095 (297436)	1335524 (76051)
Redundancy	15.7 (15.6)	19.03 (13.77)	10.75 (10.93)	21.71 (22.0)	26.55 (20.73)
Refinement					
No. mol. Asu*a*.	3	3	3	3	3
Bond length RMSD [Å]	0.018	0.0237	0.0127	0.0189	0.0103
Bond angles RMSD [°]	1.8671	2.1615	1.6661	1.9473	1.408
R_{Work} [%]	16.2	14.41	19.96	14.70	33.82
R_{Free} [%]	19.2	16.61	26.61	17.32	37.31

*a*Number of molecules per asymmetric unit.

Numbers in parenthesis indicate the highest resolution shell.

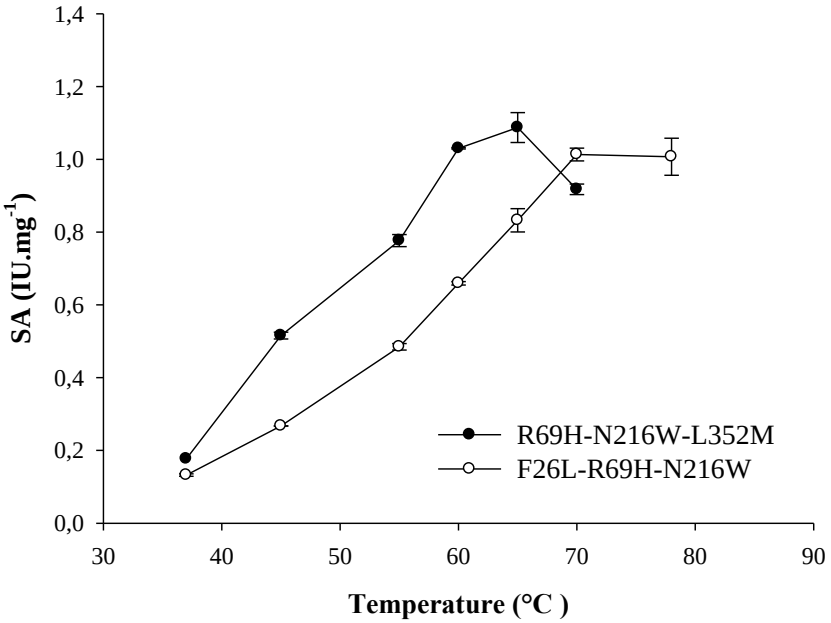


Fig. S1. Temperature dependence of specific activity for R69H-N216W-L352M and F26L-R69H-N216W in hydrolysis mode.

Assays were carried out in triplicates using 4.5 mM α-L-Ara*f*O*p*NP in 50 mM sodium phosphate buffer pH 7.0.

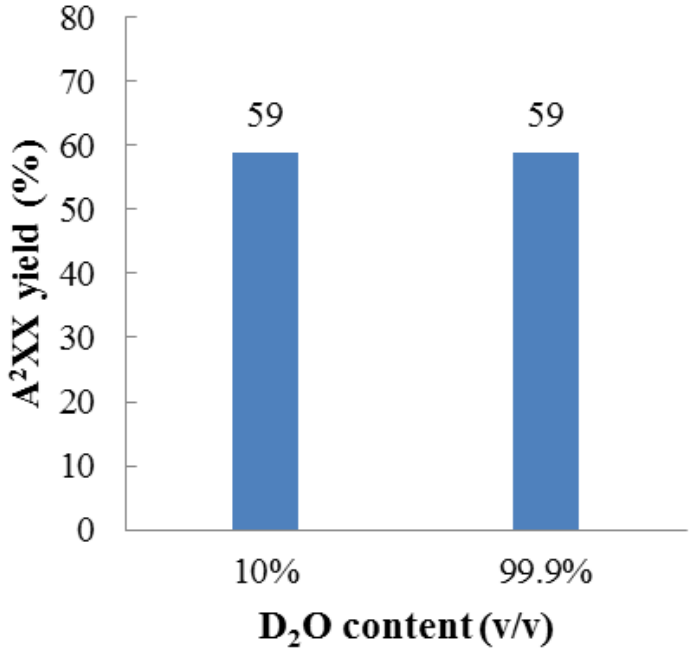


Fig. S2. Impact of D_2O content on transglycosylation performance.

Assays were carried out using 5 mM α-L-Ara*f*O*p*NP and 10 mM xylotriose in either 10% D_2O or 99.9% D_2O (both containing 10 mM sodium phosphate buffer at pH or pD 7.0, respectively), with 0.3 µM of F26L-R69H-L352M. Value of pD was measured by determining pH using a glass pH electrode and then applying the equation $pD = pH_{electrode} + 0.41$ [63].

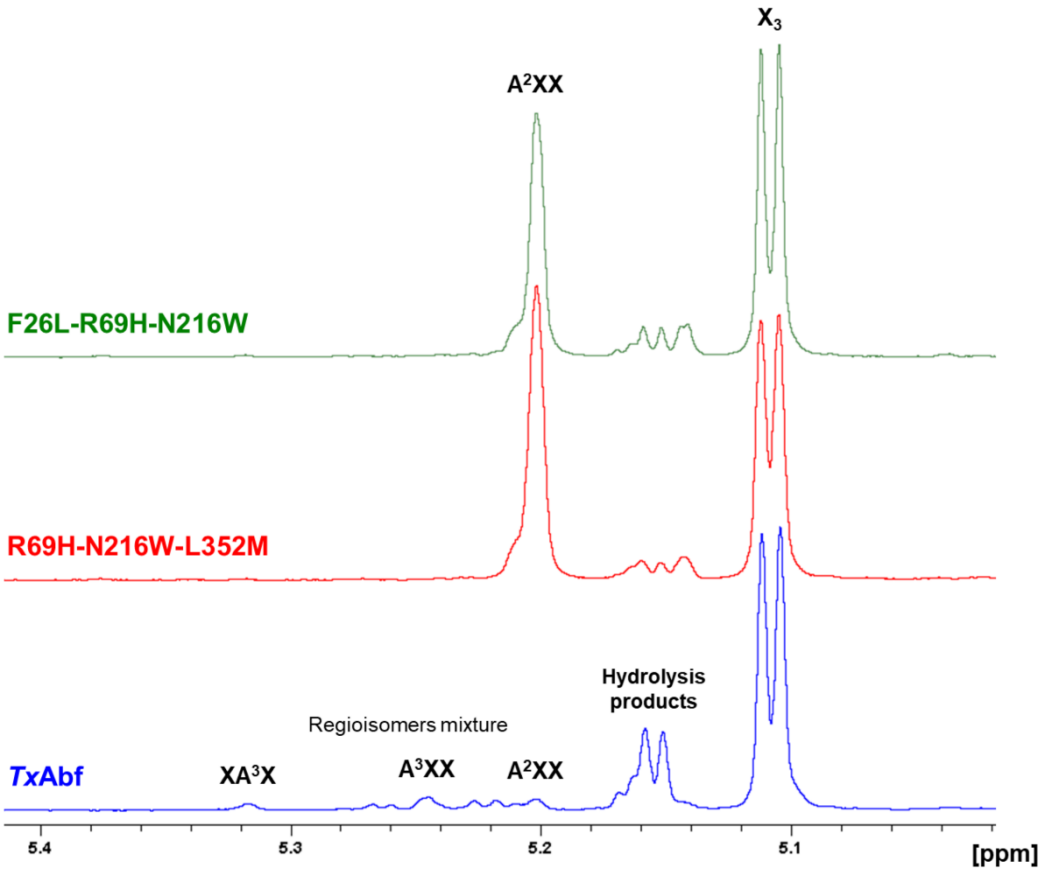


Fig. S3. NMR monitoring of transarabinofuranosylation by *Tx*Abf and triple mutants thereof: ^{1}H NMR anomeric signals of α-L-Ara*f*-containing regioisomers for wild-type and regioselective N216W-containing mutants. NMR chemical shift of the anomeric proton of α-L-Ara*f* unit of AXOS were reported in the literature around 25 °C [17]. Displacement of ^{1}H chemical shifts for α-L-Ara*f* anomeric signal of each AXOS towards blinded region (Δδ = -0.08 ppm) is observed at 45 °C. Besides the anomeric proton of the α-L-Ara*f* unit of AXOS at 5.20 ppm, the corresponding signals in XA^{3}X and A^{3}XX can be observed at 5.32 and 5.25 ppm for the wild-type at 45 °C.

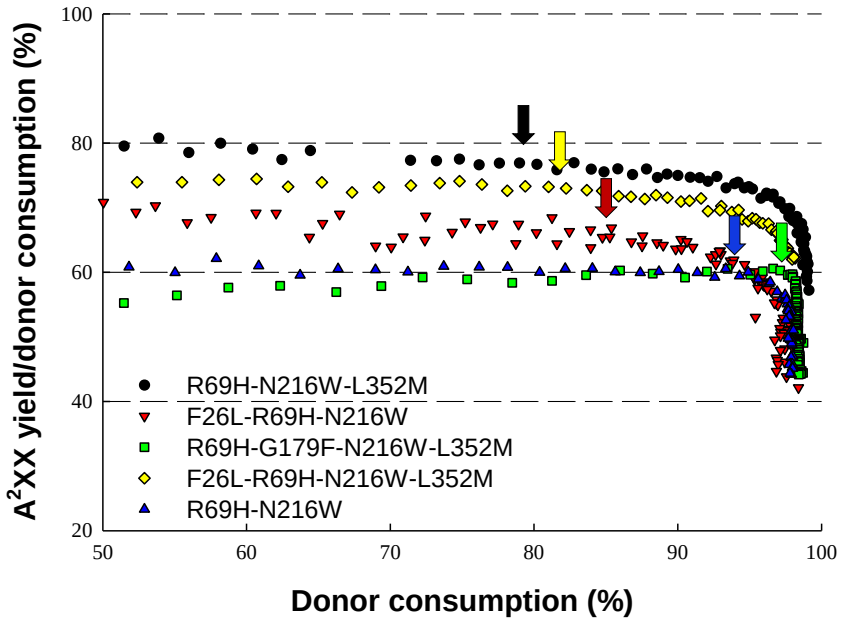


Fig. S4. A^2XX yield/donor consumption (in %) as function of donor consumption revealing the secondary hydrolysis tipping point (Table S6). Tipping points of each enzyme are indicated by arrows.

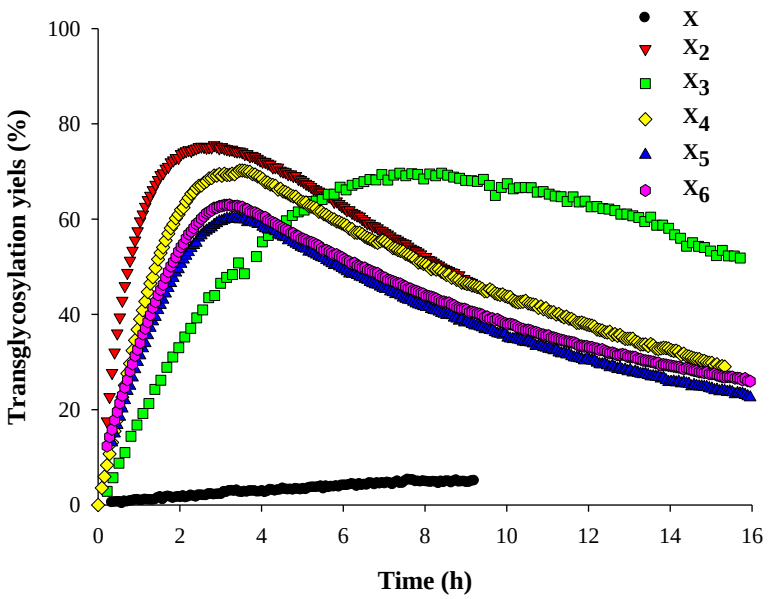


Fig. S5. Major transglycosylation product (A^2X_n) evolution as a function of time.

All assays were carried out using 5 mM *p*NP-α-L-Ara*f* as donor and 10 mM of different XOS as acceptor at 45 °C and pH 7.0 in buffered 10% D_2O, with 0.49 μM R69H-N216W-L352M (except for xylotriose, 0.23 μM enzyme was employed).

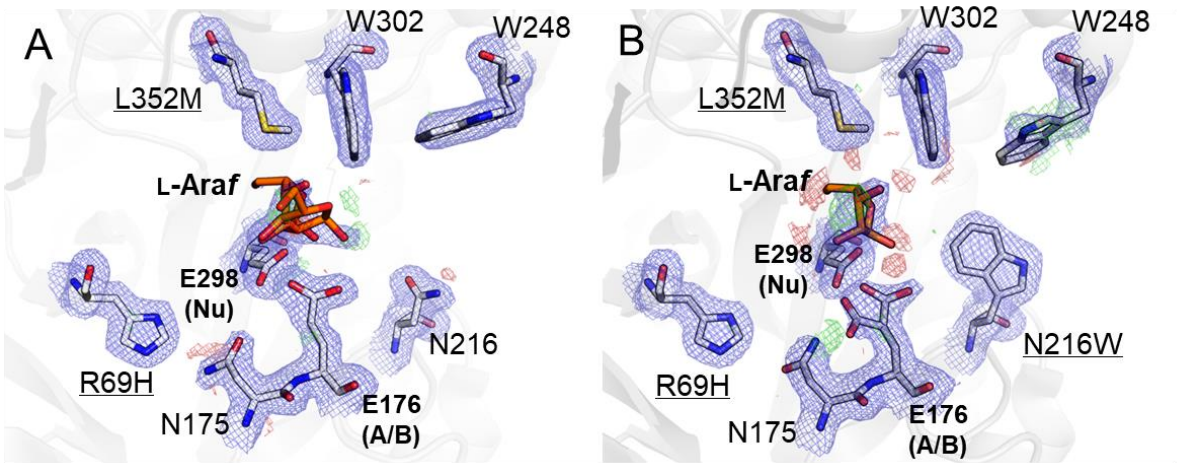


Fig. S6. Active sites of complex structures. R69H-L352M (left) and R69H-N216W-L352M (right) mutants soaked in A^2XX. Only the L-arabinofuranosyl (L-Ara*f*) subunit could be modelled into the available electron density. $2F_O$-F_C map shown at 1.0 σ cut-off in blue mesh. Difference density shown at 3.0 σ cut-off in green (positive) and red (negative). Only chain A of both structures are shown.

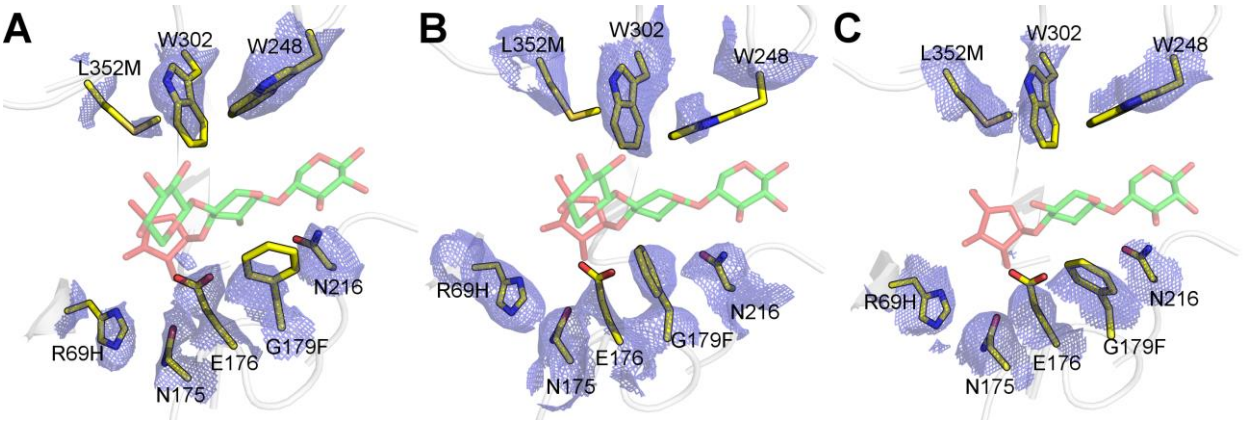


Fig. S7. Active site of R69H-G179F-L352M. The 3.1 Å electron density reveals the G179F mutation to act as a hydrophobic interaction platform. Due to the available resolution, no secondary effects could confidently be revealed. Substrate XA^{3}XX (observed as A^{3}XX in molecule C) from the E176Q mutant (PDB code: 2VRQ) shown in green, with the L-Ara*f* subunit shown in red . $2F_O$-F_C map shown at 1.0 σ cut-off in blue mesh. Molecules A, B and C of the asymmetric unit correspond to panel naming.

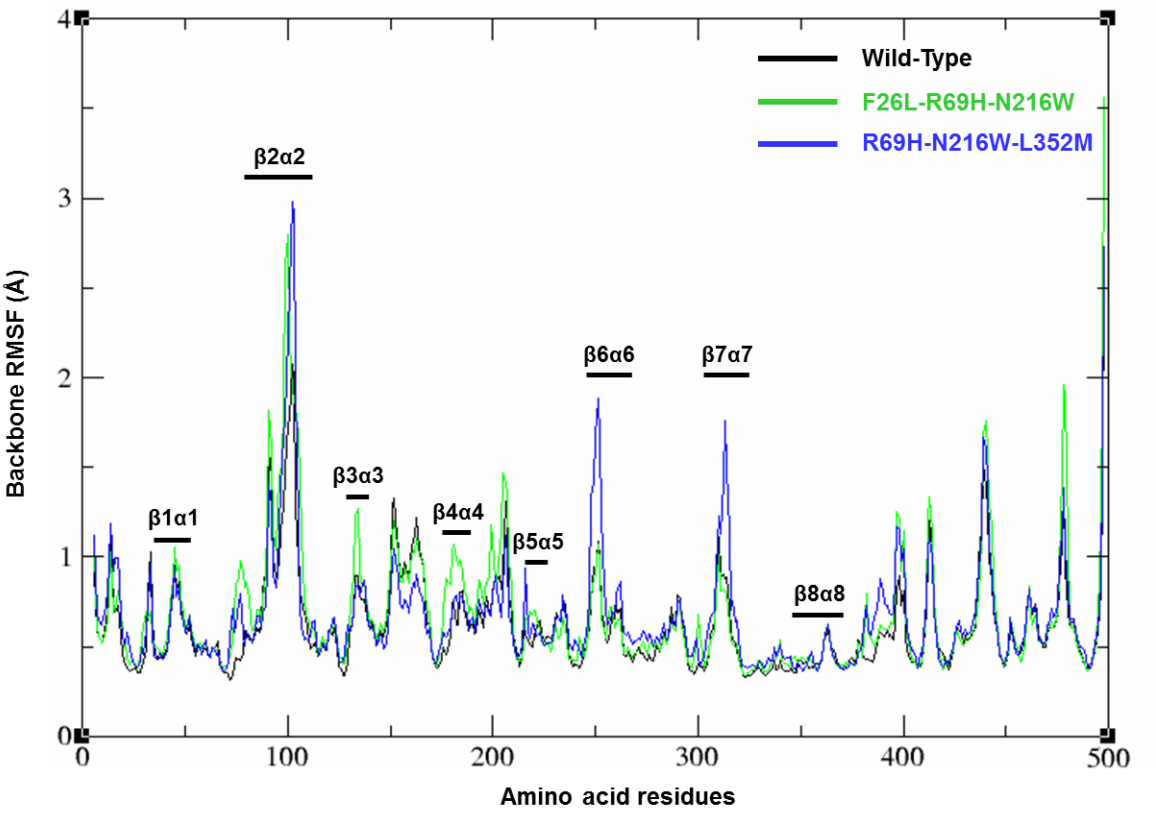


Fig. S8. Backbone root-mean-square-fluctuations (RMSF) per residue for wild-type *Tx*Ab*f* (black color), mutant F26L-R69H-N216W (green color) and mutant R69H-N216W-L352M (blue color) in the presence of ligand A^2XX during the MD simulations.

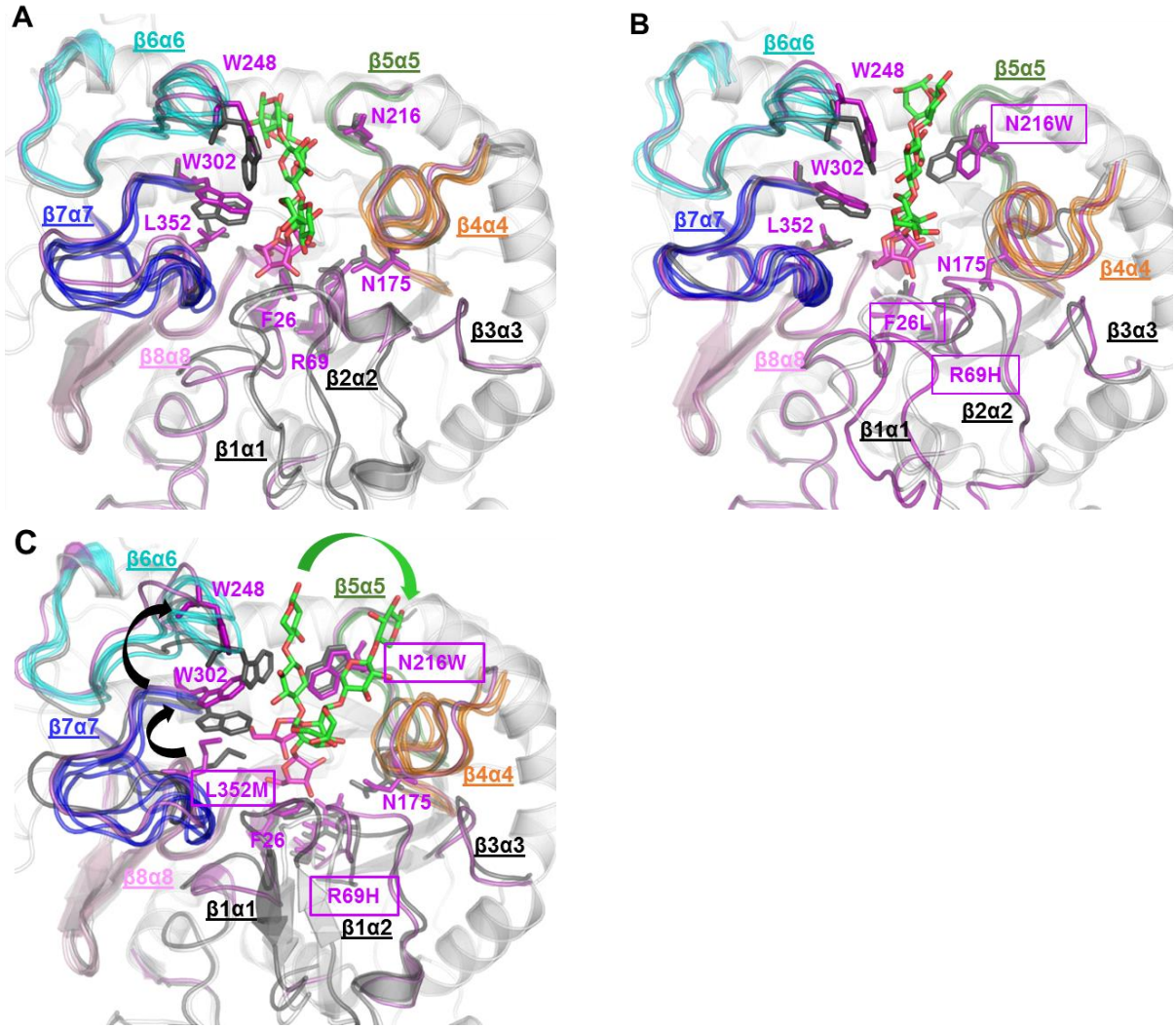


Fig. S9. Conformational changes observed along MD simulations (20 ns) performed on (A) wild-type *Tx*Ab*f*, (B) mutant F26L-R69H-N216W and (C) mutant R69H-N216W-L352M. Snapshots taken along the simulation are shown. For each represented loop, the initial state is shown in gray color and the final conformation in magenta. A^2XX is shown as sticks, colored in magenta for L-Ara*f* and green for the xylotriosyl moiety, in initial and final conformations. Mutated residues are highlighted by enframed labels.

Chapter III.

Synthesis of α-L-Ara*f* and β-D-Gal*f* series difuranosides using mutants of a GH51 α-L-arabinofuranosidase

Still focusing on our desire to refine our understanding of the T/H partition and increment the now extensive knowledge base related to *Tx*Abf, in this chapter we address a question that has previously been touched upon, but not fully studied. Indeed, self-condensation, which is a simplified form of transglycosylation in which a single glycoside substrate acts as both donor and acceptor, was studied to some extent at the beginning of the 21st century. However, the issue of self-condensation has not been revisited using all of the knowledge that has been accumulated over the last 10 years.

In this paper different mutant enzymes were studied. This revealed that self-condensation is radically improved with the double mutant R69H-N216W procuring yields of homo-D-galactofuranosides in excess of 70%. In addition to revealing how mutants increase transglycosylation, this work also provides insight into how specific mutations alter regioselectivity. This information will be precious for future engineering work aimed at finer sculpting of *Tx*Abf for specific synthetic purposes.

Synthesis of α-L-Ara*f* and β-D-Gal*f* series difuranosides using mutants of a GH51 α-L-arabinofuranosidase

Jiao Zhao, Jérémy Esque, Isabelle André, Michael J. O'Donohue*, Régis Fauré*

TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France

* Corresponding authors.

E-mail addresses: michael.odonohue@inra.fr (M.J. O'Donohue), regis.faure@insa-toulouse.fr (R. Fauré).

Abstract

The GH-51 α-L-arabinofuranosidase from *Thermobacillus xylanilyticus* (*Tx*Abf) possesses versatile catalytic properties, displaying not only the ability to hydrolyze glycosidic linkages but also to synthesize difuranosides in α-L-Ara*f* and β-D-Gal*f* series. Herein, mutants are investigated to evaluate their ability to perform self-condensation, assessing both yield improvements and changes in regioselectivity. Overall yields of oligo-L-arabino- and oligo-D-galactofuranosides were increased up to 4.7-fold compared to the wild-type enzyme. In depth characterization revealed that the mutants exhibit increased transfer rates and thus a hydrolysis/self-condensation ratio in favor of synthesis. The consequence of the substitution N216W is the creation of an additional binding subsite that provides the basis for an alternative acceptor substrate binding mode. As a result, mutants bearing N216W synthesize not only (1,2)-linked difuranosides, (1,3)- but also (1,5)-linked difuranosides. Since the self-condensation is under kinetic control, the yield of homo-disaccharides could be maximized using higher substrate concentrations. In this way, using the mutant R69H-N216W it was possible to procure >70% yield of oligo-galactofuranosides. Overall, this study further demonstrates the potential usefulness of *Tx*Abf mutants for glycosynthesis and shows how these might be used to synthesize biologically-relevant glycosconjugate.

Keywords: biocatalysis, self-condensation, retaining glycoside hydrolase, difuranosides, D-galactofuranosides, regioselectivity

Abbreviations

Abfs, α-L-arabinofuranosidases; α-L-Ara*f*, α-L-arabinofuranosyl unit; α-L-Ara*f*O*p*NP, 4-nitrophenyl α-L-arabinofuranoside; β-D-Gal*f*, β-D-galactofuranosyl unit; β-D-Gal*f*O*p*NP, 4-nitrophenyl β-D-galactofuranoside; LG, leaving group; rGH, retaining glycoside hydrolase; *p*NP, 4-nitrophenol; rGH, retaining glycoside hydrolase; SA, specific activity; *Tx*Abf, α-L-arabinofuranosidase from *Thermobacillus xylanilyticus*; T/H, transglycosylation/hydrolysis ratio.

1. Introduction

Furanoses, are carbohydrates displaying a five-membered ring structure and are ubiquitous in Nature. The main example is D-(deoxy)ribofuranose, which forms an integral part of nucleic acids. Other furanoses are also widespread, with both L-arabinofuranose (L-Ara*f*) and its 5-hydroxymethyl analog D-galactofuranose (D-Gal*f*) frequently found in naturally occurring oligo- and polysaccharides, but absent in mammalian glyco-motifs [1–4]. L-Ara*f* moieties are mostly associated with cell wall, intercellular-matrix and extracellular structures (e.g. arabinoxylans [5,6], arabinogalactans [7] and arabinans [8,9]) in higher plant polysaccharides [6,8,9], while D-Gal*f* is widespread in glycoconjugates produced by pathogenic microorganisms [1,3,10]. Notably, the galactan portion of mycobacterial arabinogalactan is formed by a linear chain of alternating β-(1,5)- and β-(1,6)-linked D-Gal*f* residues [11]. Other glycomotifs, such as β-D-Gal*f*-(1,2)-β-D-Gal*f* are found in the mucins of the protozoan *Trypanosoma cruzi* [12] and β-(1,3)-linked D-Gal*f* moieties are present in oligosaccharidic structures of fungi [13–15] and bacteria [16].

Unsurprisingly, the widespread nature L-Ara*f* and D-Gal*f* makes these two furanose forms interesting targets for synthetic chemistry. In the case of L-Ara*f*-containing compounds, these are of considerable interest for research purposes, particularly as substrates for plant cell wall acting enzymes [17]. Similarly, since D-Gal*f* forms part of antigenic epitopes in glycoconjugates of several pathogenic organisms, it constitutes a target for synthesis aimed at the development of diagnostic tools [18], or chemotherapeutic strategies [1].

In biological systems, furanose-containing carbohydrates are synthesized by arabino- and galactofuranosyltransferases, enzymes that use nucleotide sugar donors[18–21]. However, for *in vitro* synthesis, the use of such enzymes is challenging for a variety of reasons, even though

progress has recently been made in this field [22–25]. Similarly, the use of well-established synthetic chemistry strategies is fraught with difficulties, particularly because sugars in furanose configuration display less thermodynamic stability than their pyranose counterparts [26,27]. An alternative strategy to access furanose-bearing glycoconjugates is the use of transfuranosylases. Like glycosyltransferases, transfuranosylases are enzymes that possess the ability to synthesize furanose-containing compounds, using reactive furanose glycosyl sugars donor. However, unlike glycosyltransferases, transfuranosylases belong to the glycoside hydrolase (GH) class of enzymes and do not require the use of costly nucleotide sugars.

Transglycosylases are so-called retaining glycoside hydrolases (rGHs), meaning that these enzymes employ a double displacement mechanism to break and synthesize glycosidic bonds [28]. In fact, transglycosylases are a specific subtype of rGHs that, while operating the same catalytic mechanism as hydrolytic rGHs, display the ability to primarily perform transglycosylation, a reaction that involves the transfer of glycosyl donors onto glycoside acceptors. Regarding transfuranosylases, to our knowledge naturally-occurring examples are very rare, with transfructosylases being the most well-studied examples [29,30]. Nevertheless, hydrolytic α-L-arabinofuranosidases belonging to the GH51 family and possessing some ability to catalyze transarabinofuranosylation and transgalactofuranosylation have been described [31–33]. However, while α-L-arabinofuranosidases are relatively widespread and numerous in the GH classification (GH43 and 51 are archetypal arabinofuranosidase families) [34], β-D-galactofuranosidases are rarer, with just a few examples being known to date. In particular, using a large scale cloning, expression and screening approach β-D-galactofuranosidases were recently revealed in families GH5 and 43 [35], but so far little biochemical characterization has been performed [36,37], thus it is too early to say whether these enzymes display significant ability to perform transglycosylation reactions.

The GH51 α-L-arabinofuranosidase from *Thermobacillus xylanilyticus* (*Tx*Abf) is a hydrolytic rGH that is active on substrates bearing α-L-Ara*f*-moieties [38]**.** Like other rGHs, *Tx*Abf hydrolyses glyosidic bonds using a two-step displacement mechanism that involves the formation of a transient covalent glycosyl-enzyme intermediate with the catalytic nucleophile residue [39]. Formation of this catalytic species is achieved through the glycosylation step. This involves binding of the donor compound to the enzyme's active site and departure of donor leaving group, which is concomitantly protonated by a catalytic acid/base residue acting as a general acid. In the second so-called deglycosylation, step, the glycosyl-enzyme is usually broken down by a water molecule that displays nucleophilic character under the

influence of a catalytic acid/base residue acting as a general base [28,40]. However, deglycosylation can also occur when other suitable incoming species adopt nucleophilic character. When deglycosylation is mediated by an incoming glycosyl moiety, a new glycosidic bond is formed and the reaction is described as transglycosylation. Moreover, in the special case where the incoming acceptor species is the same as the donor species, the reaction is described as self-condensation. Regarding *Tx*Abf, extensive studies have revealed that this enzyme is able to perform self-condensation using L-Ara*f*, D-Xyl*p* and D-Gal*f* as substrates, the latter procuring particularly elevated yields [33,41]. Furthermore, *Tx*Abf has been the subject of protein engineering work aimed at improving its ability to perform transglycosylation [32,42,43]. However, so far no particular focus has been put on improving the self-condensation reactions performed by *Tx*Abf [33,41]. The work on self-condensation demonstrated that it was possible to synthesize α-L-arabino- and β-D-galacto-oligo-furanosides. However, the reported yields remain modest and the regioselectivity of the reaction was mainly oriented towards the formation of (1,2)-linked regioisomers. While this regioselectivity might be useful for the synthesis of certain biologically-relevant α-L-arabino- and β-D-galactofurano-oligosaccharides, it is certainly not representative of the majority of medically-important glycoconjugates.

Herein, we focus on previously generated *Tx*Abf mutants, using these to evaluate their ability to perform self-condensation. In the course of this study we have carefully characterized regioselectivity in order to reveal the potentiality of the different *Tx*Abf mutants for use as synthetic tools to access a variety of α-L-arabinofurano- and β-D-galactofurano-containing oligosaccharides (Fig 1).

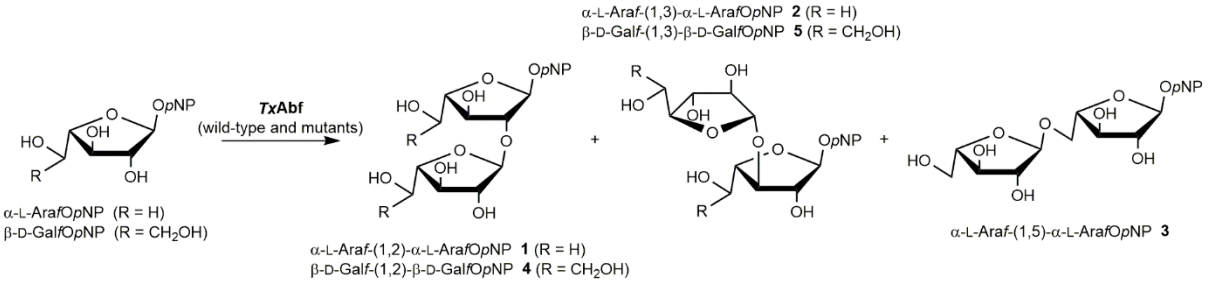


Fig. 1. Structures of the 4-nitrophenyl difuranosides synthesized through self-condensation reactions catalyzed by *Tx*Abf and its mutants. Compounds **1**-**3** and **4**-**5** were self-condensed using α-L-Ara*f*O*p*NP and β-D-Gal*f*O*p*NP as substrate respectively.

2. Materials and Methods

2.1. Substrates and Chemicals

The substrates 4-nitrophenyl α-L-arabinofuranoside (α-L-Ara*f*O*p*NP) and 4-nitrophenyl β-D-galactofuranoside (β-D-Gal*f*O*p*NP) were obtained from CarboSynth. Molecular biology reagents were purchased from New England BioLabs.

2.2. Mutagenesis, protein expression and purification

In vitro site-directed mutagenesis was achieved using the QuikChange II Site-Directed Mutagenesis Kit (Agilent). The plasmid pET24-*Tx*Abf (original pET vector was from Novagen) containing the complete *Tx*Abf coding sequence (GenBank accession no CAA76421.2) was used as the DNA template for mutagenesis. Mutations were introduced using a series of oligonucleotide primers (synthesized by Eurogentec). For recombination of mutations F26L, R69H, N216W G179F and L352M, the primers were used as previously described [44].

Double and multiple mutations were obtained through stepwise introduction of the different point mutations. At each step, mutated DNA was used to transform XL1-Blue competent cells and the success of the mutagenesis protocol was systematically verified using DNA sequencing of purified plasmid (GATC Biotech).

Expression and purification of *Tx*Abf and mutants were performed as previously described [38,45]. Briefly, a target plasmid was used to transform *E. coli* BL21 DE3 cells that were further cultured in LB medium (containing kanamycin, 50 μg.mL^{-1}) at 37 °C with shaking. For expression, 1 mM IPTG was added to cultures of *E. coli* bearing pET24-*Tx*Abf that had reached an OD_{600} of 0.5-0.6. After, these were grown for a further 4 h at 37 °C before recovering the cells using centrifugation (6,000×*g*, 15 min, 10 °C). Cells suspended in 200 mM TALON buffer (Tris-HCl pH 8 with 300 mM NaCl) were recovered a second time using centrifugation (6,000×*g*, 15 min, 10 °C) and suspended in 20 mM TALON buffer. Enzyme crude extracts were obtained by sonication (Fisherbrand™ Model FB705 Dismembrator, Fisher Scientific) of cell suspension on ice by 4 cycles of 15 s ‘on’, 45 s ‘off’, 15 s ‘on’ and 4 min ‘off’, with 40% of maximal power of the probe. Heat treatment (75 °C, 30 min) was sufficient to precipitate the majority of proteins, which were eliminated using centrifugation (6,000×*g*, 15 min, 10 °C). The subsequent soluble protein solution was applied to cobalt resin

(TALON® Metal Affinity Resin, Clontech), which specifically bound *C*-terminal $(His)_6$-tagged *Tx*Abf and mutants thereof. Protein elution was achieved using 100 mM imidazole in 20 mM TALON buffer, which procured purified *Tx*Abf (or mutants thereof), with purity being verified using SDS-PAGE. Purified *Tx*Abf solutions were concentrated and desalinated using a 10 kDa cutoff Amicon® Ultra filter (Sigma-Aldrich) and 20 mM Tris-HCl buffer, and then stored at 4 °C. Protein concentrations were determined by measuring absorbance at 280 nm using a NanoDrop ND-2000 spectrophotometer. Theoretical molecular weight and extinction coefficients were calculated using the ExPASy ProtParam tool (https://web.expasy.org/protparam/).

2.3. Enzymatic assay

Enzyme activities were measured using a discontinuous assay [45]. Reactions were performed in triplicate at 45 °C, using 5 mM β-D-Gal*f*O*p*NP as donor in 50 mM sodium phosphate buffer at pH 7 with 1 mg/mL BSA, in a final reaction volume of 350 µL. Prior to enzyme addition, reaction mixtures were pre-incubated at 45 °C. Once launched, reactions were conducted for 10-20 min, removing 40 µL samples at regular intervals. These samples were immediately mixed with 200 µL of 1 M Na_2CO_3 and placed on ice. The release of *p*NP was monitored at 401 nm using a spectrophotometer (Infinite M200 Microplate reader, Tecan) and quantified using an appropriate standard curve, prepared using pure *p*NP. Negative controls containing all of the reactants except the enzyme were used to correct for spontaneous hydrolysis of the donor substrate. Initial reaction rates were determined from the linear regions of time-dependent plots, which correspond to less than 15% consumption of the donor substrate. One unit (IU) of enzyme specific activity (SA) corresponds to the amount of enzyme releasing one µmol of *p*NP per minute. The kinetic parameters K_M, k_{cat} and k_{cat}/K_M were determined by measuring enzyme SA at various substrate concentrations. The β-D-Gal*f*O*p*NP was varied from 0.2-30 mM. A modified version of Michaelis-Menten equation (Table 2) was applied to fit the nonlinear regression plot of SA versus [S] (substrate concentration) using SigmaPlot 11.0 software.

2.4. NMR analysis

To monitor reactions, ^1H NMR spectra were collected using a Bruker Avance II spectrometer equipped with a TCI probe and operating at 500 MHz. Reactions were performed at 45 °C with α-L-Ara*f*O*p*NP or β-D-Gal*f*O*p*NP containing appropriately diluted enzymes in a total volume of 600 µL (containing 10% D_2O in 10 mM sodium phosphate buffer at pH 7 with 1

mg·mL^{-1} BSA, v/v). To initiate reactions, an aliquot of enzyme solution (10% of total reaction volume) was added to suit the 5-16 h reaction time frame. In α-L-Ara*f*O*p*NP self-condensation reaction, 0.003 µM of wild-type *Tx*Abf, 0.27 µM of R69H-N216W, 0.45 µM of R69H-N216W-L352M and 1.01 µM of F26L-R69H-N216W were used. In β-D-Gal*f*O*p*NP self-condensation reaction, 0.27 µM of wild-type *Tx*Abf, 4.00 µM of R69H-N216W and 41.80 µM of R69H-N216W-L352M were used. Time course NMR monitoring was achieved by performing pseudo-2D kinetics experiments based on a phase sensitive NOESY sequence with pre-saturation, with spectra being accumulated every 8.7 min (2 × 32 scans). Using previously determined specific activities, it was possible to adjust enzyme quantities such that reactions occurred up to maximum of 18 h, although the actual duration of the reaction was dependent on the enzyme used. The pH in 10% D_2O was measured using a glass pH electrode, applying the modified equation $pH_{10\%\ D_2O} = pH_{electrode} + 0.04$ [46]. The chemical shift reference was based on the HOD signal calibrated at 4.55 ppm at 45 °C [47].

Donor (*i.e. p*NP-glycofuranoside) consumption and the apparition of self-condensation difuranosides **1-5** were quantified by integrating the anomeric proton signals from the O*p*NP-linked furanosyl moiety (H-1a). For self-condensation products, the blinded anomeric protons associated with the non-reducing end furanosyl unit (H-1b) were not taken into account, because they overlap with the anomeric protons of the monosaccharide hydrolysis product. The selected anomeric protons and their chemical shifts are shown in Table 1. Donor substrate was quantified by integrating its relevant anomeric proton signals (Fig. 1 and Supplementary Figs. S1 and S2). The molarity of all compounds was normalized relative to the initial proton integral of the donor compound. Yields of self-condensation products were maximal yields expressed as a percentage of initial donor concentration. The mean value of the *ortho* and *meta* protons of the linked *p*NP were regarded as the sum of the donor and self-condensation products. Thereby, the overall yield of self-condensation products was calculated by subtracting the value of the integral of the anomeric donor signal from that for the average integral of *ortho* and *meta* aromatic protons of the linked O*p*NP signals (Supplementary Fig. S3). In doing so, we were able to account even for minor products, which taken alone were barely detectable.

Table 1. Chemical shift of the substrates and self-condensation products **1**-**5** (Fig. 1) used for quantifying the evolution of enzymatic reaction.[a]

Compound	δ (ppm)	
	H-1a	Aromatic H_m and H_o of *p*NP moiety
α-L-Ara*f*O*p*NP	5.78	8.20-8.18 and 7.18-7.16
1 α-L-(1,2)	5.89	
2 α-L-(1,3)	5.81	
3 α-L-(1,5)	5.79	
β-D-Gal*f*O*p*NP	5.74	8.20-8.18 and 7.18-7.16
4 β-D-(1,2)	5.82	
5 β-D-(1,3)	5.78	

[a] The chemical shifts were analyzed at 45 °C and pH 7 in buffered 10% D_2O. Chemical shifts of self-condensation difuranosides **1**-**5** were referred to previous work [2,33,41].

2.5. Modelling *Tx*Abf and R69H-N216W-L352M in complex with self-condensation products

Molecular docking experiments were carried out using YASARA software (version 12.8.1) [48]. Crystallographic data of wild-type *Tx*Abf (PDB: 2VRQ) and its mutant, bearing mutations R69H-N216W-L352M [44], were used to construct 3D models of mutant R69H-N216W-L352M [44]. Next, α-L-Ara*f* disaccharides (without *p*NP) were built using Glycam Force Field and tleap program from AMBER16 software suite [49]. For each α-L-Ara*f* disaccharide, a short minimization in implicit solvent (Generalized Born) with 250 steps of steepest descent and 250 steps of conjugated gradient was performed, using a cut-off of 99 Å**.** *p*NP aglycon was added at the reducing end and geometrically optimized using Avogadro software [50]. β-D-Gal*f* disaccharide structures were derived from the corresponding α-L-Ara*f* disaccharides by adding hydroxymethyl on each α-L-Ara*f* unit using Avogadro software**.** The complexes formed by the enzyme and the self-condensation products (**1**-**5**) were initially assembled in PyMol [51] by pair fitting onto bound α-L-Ara*f* **(**occupying subsite -1) in X-ray structure of wild-type *Tx*Abf (PDB: 2VRQ**).** Complexes were further subjected to energy minimization macro supplied in the YASARA package. The figure was then prepared using PyMol Molecular Graphics System, v0.99 (Schrödinger).

3. Results

3.1. Kinetic analysis of *Tx*Abf mutants

Measurement of steady state kinetic parameters of α-L-Ara*f*O*p*NP conversion (monitoring *p*NP release) catalyzed by *Tx*Abf and its mutants revealed that all mutations severely affected catalytic efficiency (Table 2). This was also the case for *Tx*Abf acting on β-D-Gal*f*O*p*NP. The values were reduced by more than two and three orders of magnitude in reactions containing α-L-Ara*f*O*p*NP and β-D-Gal*f*O*p*NP respectively. Comparing the k_{cat}/K_M values of reactions catalyzed by the wild-type enzyme of the two substrates (α-L-Ara*f*O*p*NP and β-D-Gal*f*O*p*NP) revealed a more than 6 000-fold difference. Regarding the reaction involving α-L-Ara*f*O*p*NP and R69H-N216W-L352M the K_M value was relatively similar to that of the reaction catalyzed by the wild-type enzyme, whereas the same reaction catalyzed by F26L-R69H-N216W is characterized by significant decreases in both k_{cat} and K_M. Assuming that deglycosylation is rate-limiting, the lowered k_{cat} value leads us to postulate that R69H-N216W-L352M is mostly impaired with respect to deglycosylation, whereas the impact of F26L-R69H-N216W appears to be more complex, affecting both steps [32]. Regarding the reaction involving R69H-N216W and β-D-Gal*f*O*p*NP, a two-phase reaction profile was observed that did not reach saturation (Fig. 2) [32,44]. Remarkably, unlike wild-type *Tx*Abf, the mutant R69H-N216W yielded a measurable K_M value (0.45 mM) for the reaction containing β-D-Gal*f*O*p*NP, which is similar to that of wild-type *Tx*Abf acting on α-L-Ara*f*O*p*NP. Nevertheless, because the k_{cat} value of the reaction is also extremely low, the overall catalytic efficiency of R69H-N216W acting on β-D-Gal*f*O*p*NP is unchanged compared to the reaction involving wild-type *Tx*Abf.

Table 2 Kinetic parameters of reactions catalyzed by *Tx*Abf and mutants thereof, using α-L-Ara*f*O*p*NP or β-D-Gal*f*O*p*NP as substrate.

Enzyme	SA_{th} (IU.mg^{-1})	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1}.mM^{-1})	N_S[a] (s^{-1}.mM^{-1})
α-L-Ara*f*O*p*NP					
*Tx*Abf[b]	615	0.72	575	795	-
*Tx*Abf[c]	145	0.25	139	556	-
R69H-N216W-L352M[c]	0.60	0.48	0.58	1.21	0.034
F26L-R69H-N216W[c]	0.16	0.01	0.15	10.81	0.034
β-D-Gal*f*O*p*NP					
*Tx*Abf[b,d]	15.6	>50	15.02	0.13	-
R69H-N216W[e]	0.065	0.45	0.06	0.14	0.008

[a] Ns is a nonspecific constant that is included in the modified Michaelis-Menten equation to account for activation of the enzyme by the self-condensation product: $SA_{app.} = SA_{th.} \cdot [S]/(K_M + [S]) + N_S \cdot [S]$ where $SA_{th.}$ is the theoretical maximum activity achieved if the enzyme operates according to the Michaelis-Menten model.

[b] The kinetic parameters were referred to previous work at 60 °C in 50 mM sodium acetate buffer pH 5.8 [52].

[c] The kinetic parameters were referred to previous work at the same operating condition (45 °C and pH 7) of enzymatic assay [32,44].

[d] The K_M value is an estimate, because its measurement was unfeasible due to the limited solubility of the substrate.

[e] Assays were carried out at 45 °C and pH 7 in 50 mM sodium phosphate buffer in triplicates. Relative errors were inferior to 10%.

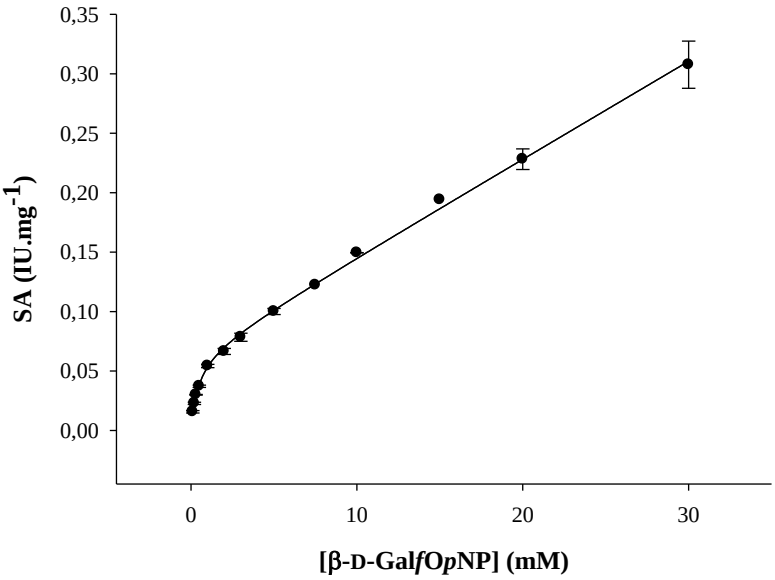


Fig. 2. Specific activity (SA) as a function of β-D-Gal*f*O*p*NP concentration for R69H-N216W.

3.2. Self-condensation profile using α-L-Ara*f*O*p*NP as donor

When using *Tx*Abf, the yield of the major self-condensation product **1**, α-L-Ara*f*-(1,2)-α-L-Ara*f*O*p*NP (Fig. 1 and Table 3), is consistent with that previously reported [41]. This compound was accompanied by trace amount of compounds **2** and **3**, corresponding to (1,3) and (1,5) linkages respectively. Compared to *Tx*Abf, the three R69H-N216W-containing mutants exhibited enhanced ability to perform self-condensation, with overall yields being 3.5- to 3.8-fold higher. Moreover, increased yield was coupled to significant changes in enzyme regioselectivity, with the proportion (% of total yield) of (1,3)-linked product **2** increasing from 30% (*Tx*Abf) to >70% in reactions catalyzed by the mutants, and even reaching 82% in the case of R69H-N216W (Table 3 and Supplementary Fig. S1). Overall, the three mutants were similar in terms of both regioselectivity and yields.

Using real time NMR spectroscopy, it was possible to monitor the evolution of each self-condensation product (Fig. 3), and thus observe secondary hydrolysis during the latter stages of the reaction. For wild-type *Tx*Abf (Fig. 3A), the maximum yield (7%) of the major product **1** was reached at 68% conversion of α-L-Ara*f*O*p*NP, meaning that hydrolysis was the predominant reaction. Regarding the R69H-N216W-containing mutants, maximal yields were achieved at later stages in the reaction, when the conversion of α-L-Ara*f*O*p*NP was higher, *i.e.* 88, 90 and 92% for F26L-R69H-N216W, R69H-N216W and R69H-N216W-L352M respectively (Fig 3 D, B and C). This means that compared to wild-type *Tx*Abf secondary

hydrolysis was delayed in reactions catalyzed by the mutants. When comparing the decomposition rate of major compound **2**, it is clear that R69H-N216W performs secondary hydrolysis (v_{HII}) at a higher rate than R69H-N216W-L352M and F26L-R69H-N216W (Supplementary Table S1). Nevertheless, considering the combined performance (*i.e.* yield of the major product **2**, its synthesis rate v_S and its SA) R69H-N216W was the best performing enzyme, although its overall self-condensation yield was slightly lower than that of the other two mutants (Supplementary Table S1 and Table 3). Additionally, R69H-N216W-L352M displayed the highest self-condensation/secondary hydrolysis (v_S/v_{HII}) ratio with a value of 3 (Supplementary Table S1).

Table 3 Maximal self-condensation yields (in %) using α-L-Ara*f*O*p*NP as donor.

Enzyme	SA ($IU.mg^{-1}$)[a]	**1** α-L-(1,2)	**2** α-L-(1,3)	**3**[b] α-L-(1,5)	Overall yield[c]
*Tx*Abf	261.79 ± 10.72	7	3	2	10
R69H-N216W	1.60 ± 0.02	6	37	8 (3)	45
R69H-N216W-L352M	0.97 ± 0.11	10	34	6 (3)	47
F26L-R69H-N216W	0.32 ± 0.01	6	38	8 (4)	48

[a] SA was measured by monitoring *p*NP release using a discontinuous assay. Assays were carried out at 45 °C in 50 mM sodium phosphate buffer pH 7 containing 5 mM α-l-ArafOpNP.
b Since maximal yield of product **3** was reached later than those of products **1** and **2**, the yield of **3** when **1** and **2** are maximum is shown in brackets.
c Maximal yields of each disaccharide are reached after different incubation times. This explains why overall yields are not equal to the sum of the maximum yield of each disaccharide.

In absolute terms, the yield of α-L-Ara*f*-(1,5)-α-L-Ara*f*O*p*NP **3** was also slightly (3 or 4-fold) increased in reactions catalyzed by the mutants. However, in relative terms the proportion of compound **3** was slightly reduced from 20% of total yield (*Tx*Abf) to as little as 6% (R69H-N216W-L352M) when maximal total transglycosylation yield was reached. Significantly, although the yield of compound **3** is low for the mutants, it continues to be produced, even when α-L-Ara*f*O*p*NP is consumed (Fig. 3B-D). This observation suggests that at later stages in the reaction, compound **3** is formed via hydrolysis of compounds **1** and **2** and the rate of hydrolysis of compound **3** is inferior to that of its synthesis.

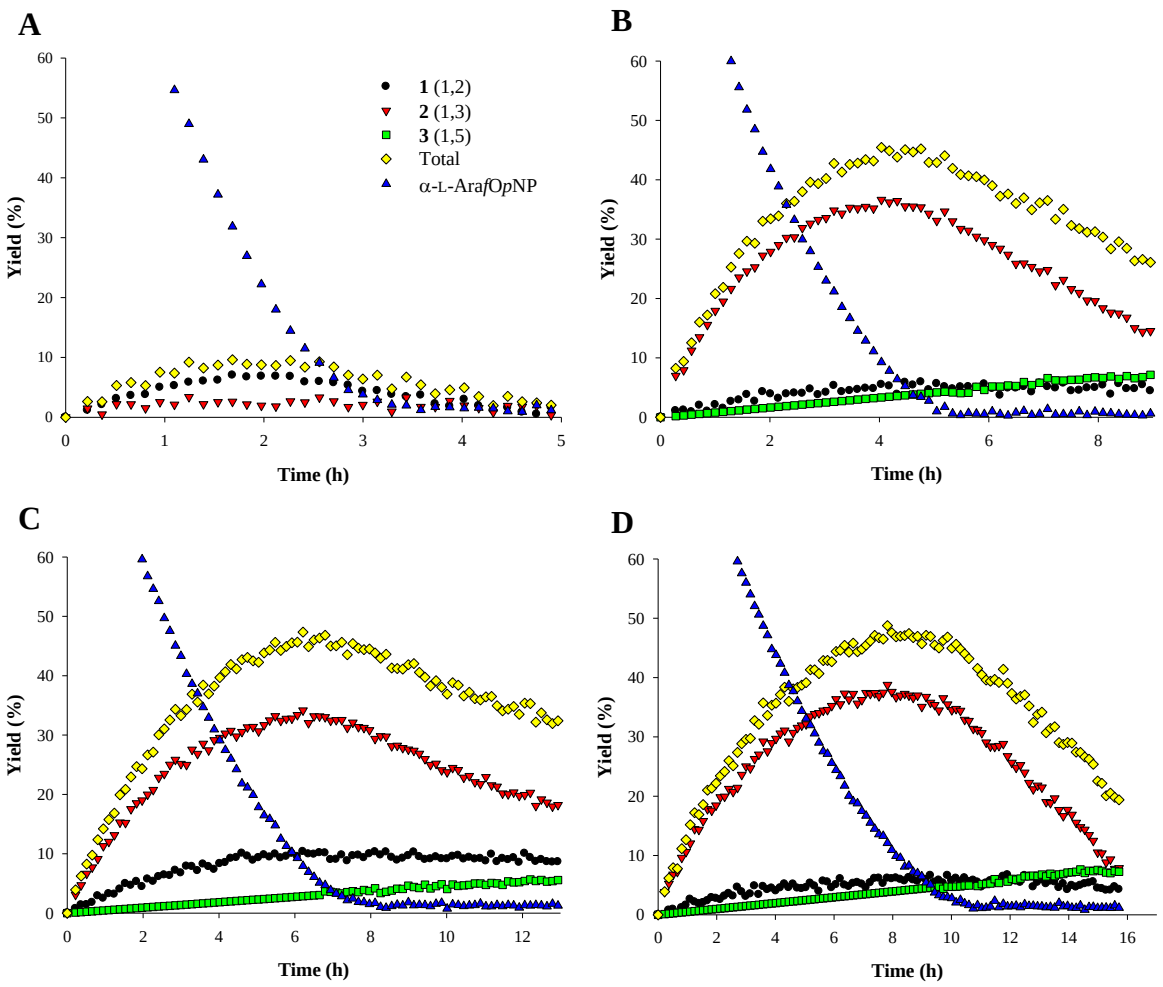


Fig. 3. Monitoring of α-L-Ara*f*O*p*NP consumption and the apparition of the different self-condensation products **1**-**3** as a function of time in reactions catalyzed by (A) *Tx*Abf, (B) R69H-N216W, (C) R69H-N216W-L352M, and (D) F26L-R69H-N216W.

3.3. Self-condensation profile using β-D-Gal*f*O*p*NP as donor

Despite being a poor substrate for *Tx*Abf-mediated hydrolysis, β-D-Gal*f*O*p*NP readily participated in the self-condensation reaction (Fig. 1), procuring a 24% overall yield of oligosaccharide products (Table 4). The major product **4**, β-D-Gal*f*-(1,2)-β-D-Gal*f*O*p*NP (18% yield, Table 4), represents 75% of the regioisomeric mixture, revealing the good regioselectivity of wild-type *Tx*Abf in this reaction. This observation is highly consistent with previous work performed both on *Tx*Abf [33] and other GH51 Abfs [53] and further confirms the natural propensity of the wild-type enzyme to hydrolyze and synthesize (1,2)-linkages. Characterization of mutants containing single substitutions (R69H and N216W) revealed that these display regioselectivities and yields similar to those of *Tx*Abf (Table 4, Fig. 4A and Supplementary Fig. S4). However, combination of these point mutations (*i.e.* R69H-N216W

and R69H-N216W-L352M) procured enzymes that double the total product yield and give regioisomeric mixtures containing a (1,3)-linked disaccharide **5** with a yield of 18% (R69H-N216W-L352M) and 22% (R69H-N216W) respectively. In this respect, the yield of product **4** was relatively unaltered, meaning that the gains in overall yield can be completely assigned to the vast improvement in the enzyme's ability to form regioisomer **5** (Table 4 and Supplementary Fig. S2).

Table 4 Maximal self-condensation yields (in %) using β-D-Gal*f*O*p*NP as donor.

Enzyme	SA (IU.mg^{-1})[a]	**4** β-D-(1,2)	**5** β-D-(1,3)	Overall yield[b]
*Tx*Abf	1.23 ± 0.09	18	2	24
R69H	0.05 ± 0.0003	18	3	24
N216W	1.33 ± 0.035	21	3	24
R69H-N216W	0.09 ± 0.001	15	22	47
R69H-N216W-L352M	0.025 ± 0.001	16	18	45

[a] SA was measured by monitoring *p*NP release using a discontinuous assay. Assays were carried out at 45 °C in 50 mM sodium phosphate buffer pH 7 using 5 mM β-D-Gal*f*O*p*NP.

[b] Overall yield is higher than the sum of compound **4** and **5** due to the presence of small amounts of β-D-(1,5) and β-D-(1,6)-linked disaccharides and traces of *p*-nitrophenyl oligo-β-D-galactofuranosides displaying degrees of polymerization >2.

In reactions catalyzed by wild-type *Tx*Abf, the maximum yield of digalactofuranosides **4** and **5** was reached when 70% of the substrate had been consumed. However, in the case of the mutants, this point was shifted to 92 (R69H-N216W-L352M) and 96% (R69H-N216W) respectively (Fig. 4). Despite this difference between reactions catalyzed by wild-type and mutant enzymes, it is noteworthy that in all three cases the self-condensation products displayed remarkable stability, meaning that secondary hydrolysis was strongly diminished, even at very low donor substrate concentrations (Fig. 4 and Supplementary Fig. S6). This was particularly noticeable for the reaction involving the mutant R69H-N216W, since product decomposition was barely perceptible. Therefore, accounting for the additional fact that R69H-N216W displayed 3.6-fold higher catalytic activity than the triple mutant (Table 4),

this enzyme can be regarded as the best performing enzyme in terms of β-D-Gal*f*O*p*NP self-condensation.

Although the major products of α-L-Ara*f*O*p*NP and β-D-Gal*f*O*p*NP self-condensation were disaccharides, in reactions catalyzed by *Tx*Abf enzymes (more for the mutants), trisaccharides were also observed. However, due to the small quantity produced and the fact that NMR signals overlapped with those of the disaccharides, these trisaccharides were neither quantified, nor fully characterized. Nevertheless, their detection further demonstrates the aptitude of the mutant enzymes to synthesize L-Ara*f* and D-Gal*f*-oligosaccharides.

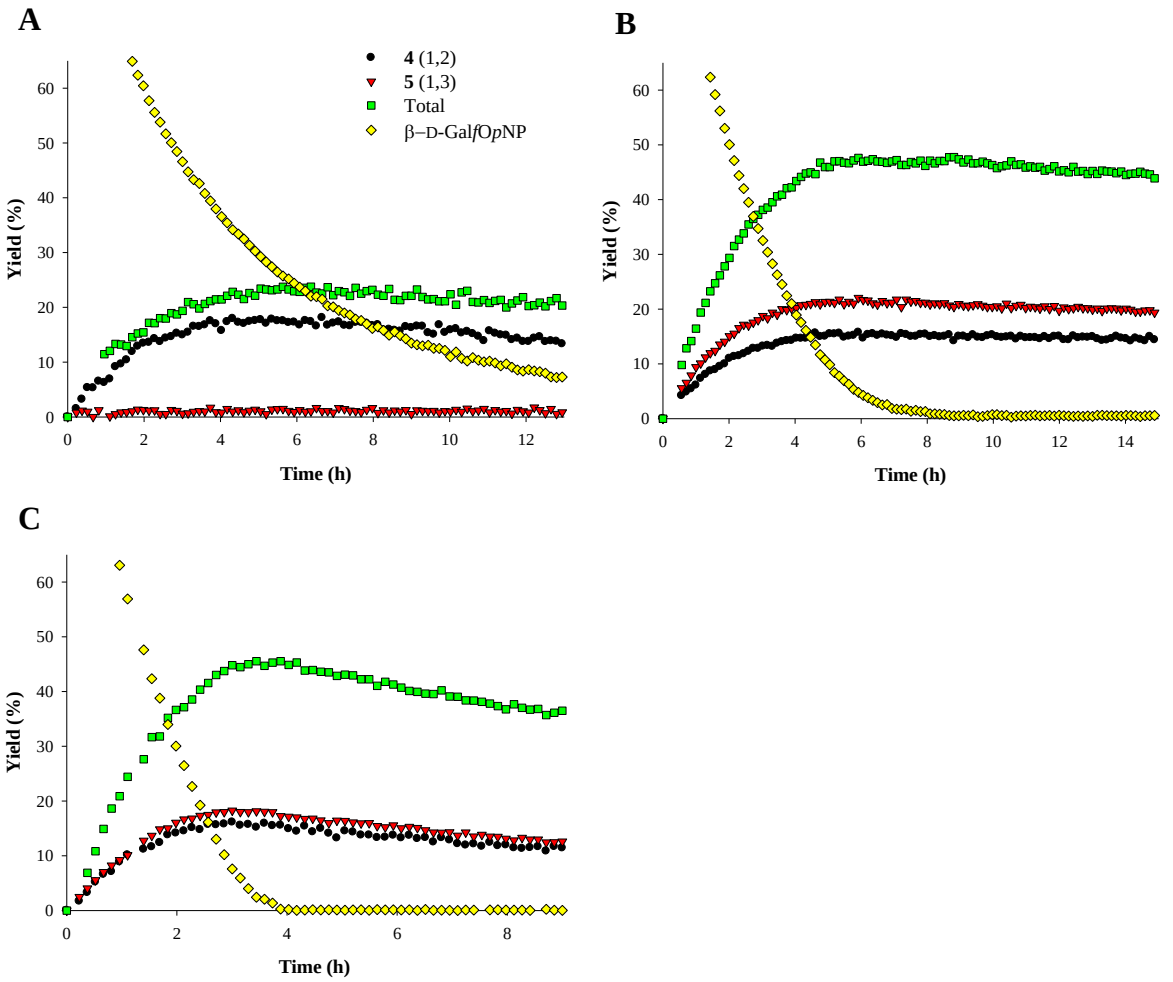


Fig. 4. Monitoring of β-D-Gal*f*O*p*NP consumption and the apparition of the different self-condensation products **4-5** as a function of time in reactions catalyzed by (A) *Tx*Abf, (B) R69H-N216W, and (C) R69H-N216W-L352M.

3.4. Probing the relationship between substrate concentration and global synthesis yield

In view of the catalytic superiority of R69H-N216W for the self-condensation of β-D-Gal*f*O*p*NP, this enzyme was employed to explore how substrate concentration influences the global yield of products (Fig. 5). Increasing substrate concentration clearly elevated both total self-condensation yield and the rate of donor conversion into self-condensation products (Fig. S5). Moreover, increasing β-D-Gal*f*O*p*NP concentration delayed the (tipping) point at which secondary hydrolysis became dominant, outweighing self-condensation. Consequently, the linear phase of self-condensation of the R69H-N216W-catalyzed reactions was extended at higher substrate concentration, meaning that the reaction is mainly under kinetic control. An overall yield of 72% was achieved when using 30 mM of β-D-Gal*f*O*p*NP. Considering substrate solubility, this is close to the maximum achievable yield in an aqueous reaction system.

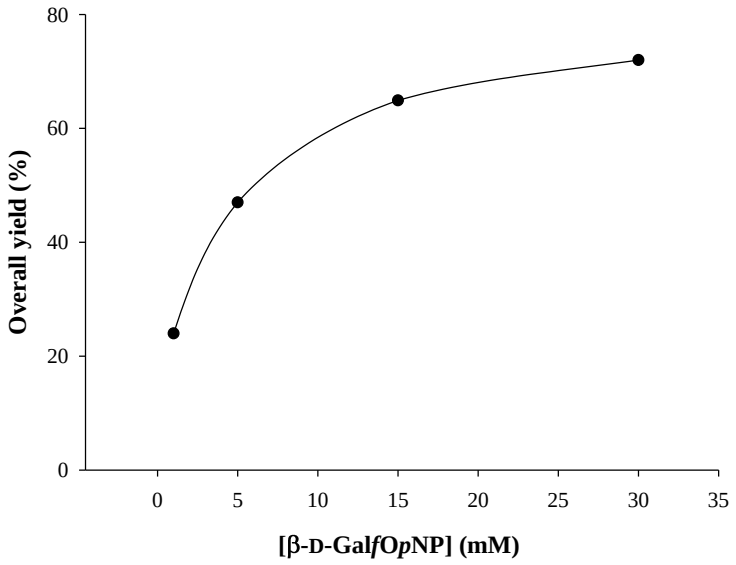
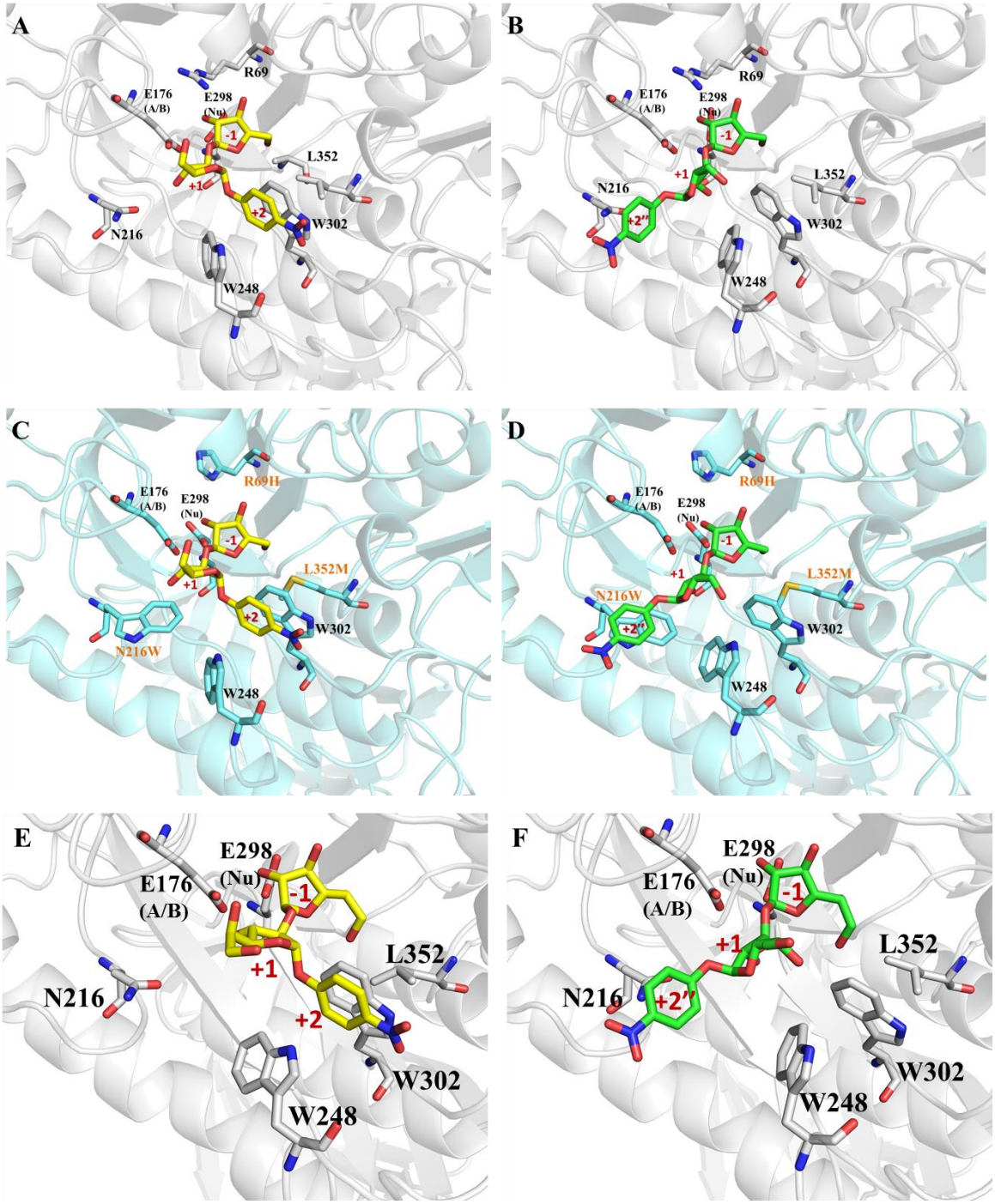


Fig. 5. Relationship between β-D-Gal*f*O*p*NP substrate concentration and overall self-condensation yield in the R69H-N216W-catalyzed reaction.

3.5. Getting molecular insight on the regioselectivity of the mutant enzymes

To investigate the regioselectivity of the mutant enzymes, 3D models of the complexes of *Tx*Abf and R69H-N216W-L352M were built using self-condensation products **1**-**5** bound in the active site (Fig. 6). Strikingly, visual analysis of 3D models of the complexed mutant enzymes revealed that mutation of Asn216 by a tryptophan residue (N216W) engenders a binding cleft that is better-adapted to accommodate the more linear (1,3)-linked difuranosides. This observation is fully consistent with the experimental results associated with both the double and triple mutants. Moreover, 3D models suggest that the nature of the furanosyl moieties (L-Ara*f* *vs* D-Gal*f*) does not significantly alter product positioning and binding interactions in the active site, consistent with the fact that the impact of N216W on regioselectivity is equivalent when either α-L-Ara*f*O*p*NP or β-D-Gal*f*O*p*NP are used as substrates. Comparing binding of the (1,2)- and (1,3)-linked compounds, occupation of subsite -1 is globally very similar, probably because the spatial confinement of this pocket-like subsite, combined with a significant number of specific enzyme-substrate interactions therein does not provide room for alternative binding modes. However, some differences are observed in subsite +1 depending on the nature of the furanosyl moiety occupying the site and on how the *p*NP moiety is accommodated (Fig. 6). In the case of (1,2)-linked compounds, 3D models suggest that both the wild-type and mutant enzymes accommodate the *p*NP moiety in subsite +2 [54], which contains residue Trp302, likely establishing a π-π stacking interaction with the *p*NP moiety (Fig. 6A, C, E and G). Regarding (1,3)-linked products, the model reveals that the *p*NP moiety lies in a different spatial position and thus occupies an alternative subsite, described here as +2''. In the wild-type enzyme, subsite +2'' contains Asn216 (Fig. 6B and F). However, in the mutant enzymes, this residue is substituted by a tryptophan, forming a hydrophobic platform that can form a π-π interaction with the *p*NP moiety (Fig. 6D and H). Consequently, this modification putatively creates a new binding mode for (1,3)-linked difuranosides that is absent in wild type *Tx*Abf. It is noteworthy that the emergence of an alternative subsite +2'' in the mutant enzymes does not alter subsite +2, which explains why the mutants do not lose their ability to form (1,2)-linked self-condensation products (Supplementary Fig. S6). Indeed, the initial interactions of *p*NP from compound **1** and **4** with Trp 248, Trp302 and putatively Lys251 are retained for R69H-N216W-L352M (Fig. S6).


A
R69
E176
(A/B)
E298
(Nu)
-1
+1
+2
L352
W302
N216
W248
B
R69
E176
(A/B)
E298
(Nu)
-1
+1
N216
+2''
L352
W302
W248
C
R69H
E176
(A/B)
E298
(Nu)
-1
+1
+2
L352M
W302
N216W
W248
D
R69H
E176
(A/B)
E298
(Nu)
-1
+1
N216W
+2''
L352M
W302
W248
E
E298
(Nu)
E176
(A/B)
-1
+1
L352
N216
+2
W302
W248
F
E298
(Nu)
E176
(A/B)
-1
+1
L352
N216
+2''
W302
W248

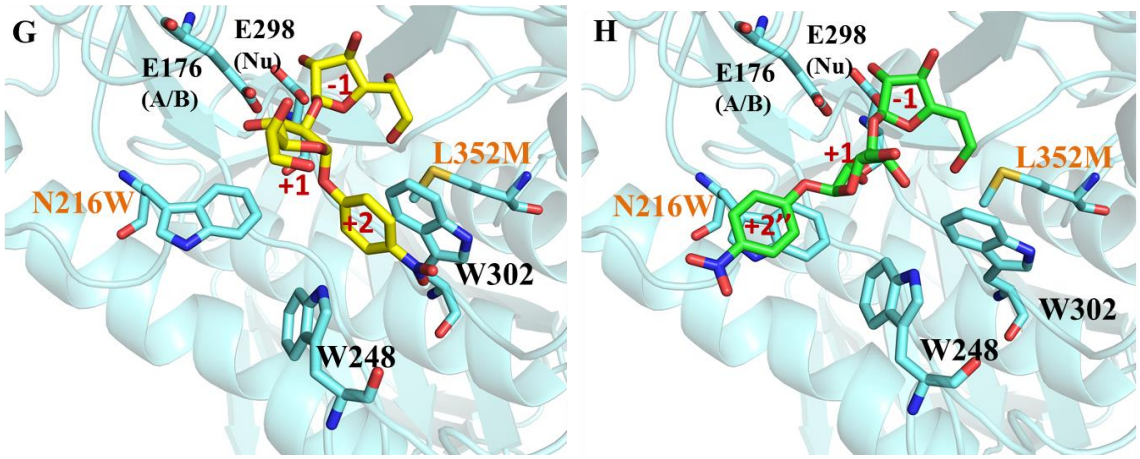


Fig. 6. Wild-type *Tx*Abf (white) and R69H-N216W-L352M (cyan) bound to products **1** or **4** (yellow) and **2** or **5** (green). (A**)** *Tx*Abf bound to its (major) product **1** and (B) *Tx*Abf bound to its (minor) product **2**, (C) R69H-N216W-L352M bound to its (minor) product **1** and (D) R69H-N216W-L352M bound to its (major) product **2** (E) *Tx*Abf bound to its (major) product **4** and (F) *Tx*Abf bound to its (minor) product **5**; (G) R69H-N216W-L352M bound to its (minor) product **4** and (H) R69H-N216W-L352M bound to its (major) product **5**. The protein is shown in cartoon, whereas the key labelled residues and the ligands are shown in stick representation.

Finally, experimental data revealed that, in the presence of α-L-Ara*f*O*p*NP, mutant enzymes displayed some higher ability to produce (1,5)-linked products. 3D modelling of α-L-Ara*f*-(1,5)-α-L-Ara*f*O*p*NP **3** shows that this disaccharide adopts a conformation that is similar (albeit more extended) to that of α-L-Ara*f*-(1,3)-α-L-Ara*f*O*p*NP **2**. As highlighted by 3D models, this suggests that compounds **2** and **3** occupy the same subsites and that the production of compound **3** is increased by the mutation N216W and the unveiling of subsite +2'' (Fig. 7). However, accounting for the experimental data, it is likely that the more extended form of compound **3** (compared to compound **2**) leads to less favorable interactions in subsite +2'', in particular with respect to residue Trp216.

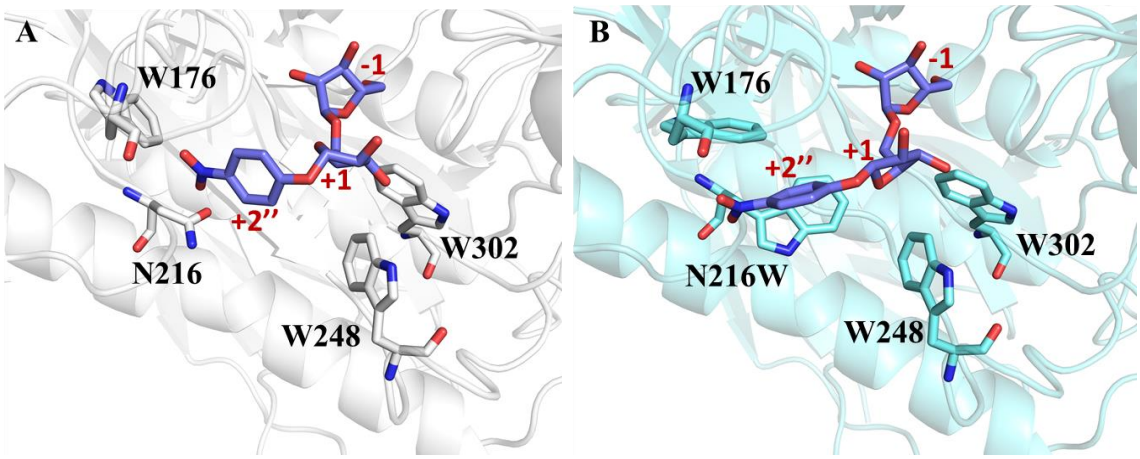


Fig. 7. Enzymes bound to the (1,5)-linked product **3** (A**)** wild-type *Tx*Abf and (B) R69H-N216W-L352M.

4. Discussion

The α-L-arabinofuranosidase, *Tx*Abf is probably the best studied member of GH51, especially regarding its ability to perform glycosynthesis [31–33,41–43,45]. So far, we have shown that this enzyme possesses a wide range of capabilities, including the ability to synthesize D-galactofuranosides and is relatively amenable to enzyme engineering [33]. In the present study, we have revisited some of the more fundamental features of *Tx*Abf, in particular its ability to catalyze self-condensation reactions using either pento- or hexo-furanosides, focusing on the way in which mutagenesis of key residues modifies this property.

4.1. Key factors favoring self-condensation

Both mutations F26L and L352M have been reported to enhance the ability of *Tx*Abf to perform transglycosylation [32,43,44]. Moreover, in the case of L352M, we have shown that modification of *Tx*Abf properties is due to quite complex long-range effects that affect both the negative and positive subsites, whereas F26L will only affect subsite -1 [44]. Here, we reveal that neither of these mutations produce major effects on self-condensation when compared to the more determining effect of the double mutation R69H-N216W. Clearly, these latter mutations are responsible for the vast overall improvements in self-condensation reported in this work and, in the case of N216W mutation, for the acquisition of enhanced aptitude to form (1,3)-linked difuranosides and α-L-(1,5)-linked diarabinofuranosides. It is noteworthy that mutation R69H alone does not enhanced the self-condensation ability of β-D-Gal*f*O*p*NP compared to the wild-type *Tx*Abf. The fine tuning of the interactions between the wild-type *Tx*Abf and single-mutant R69H with β-D-Gal*f*O*p*NP are thought to differ *i.e.*, the

global repositioning of the donor within subsite -1 due to the extra hydroxymethyl at *C*-5 and disruption of both catalytic residues features mediated by H69 respectively [32,44]. However, these two effects are not cumulative for the enhancement of the self-condensation yield in the case of unusual β-D-Gal*f*O*p*NP/mutant R69H pair.

4.2. Improving self-condensation of β-D-Gal*f*O*p*NP

One of the aims of this work was to improve the ability of *Tx*Abf to synthesize D-galactofurano-oligosaccharides. In previous work, we have shown that it is extremely difficult to determine the K_M value for reactions containing β-D-Gal*f*O*p*NP, catalyzed by the wild-type enzyme [33,52,55]. Moreover, previously performing both STD-NMR and ITC failed to reveal any measurable interactions between *Tx*Abf and β-D-Gal*f*O*p*NP [52]. Therefore, it was extremely interesting to observe that the double mutant R69H-N216W not only improves self-condensation of β-D-Gal*f*O*p*NP, but also provides the basis to measure a K_M value that is similar to that measured for the wild-type enzyme in the presence of α-L-Ara*f*O*p*NP. Since the constant K_M (= $[k_3 \cdot (k_{-1} + k_2)]/[k_1 \cdot (k_2 + k_3)]$) is composed of both an affinity component ($1/K_d = k_1/k_{-1}$) and constants related to both glycosylation (k_2) and deglycosylation (k_3) steps, it is difficult to further analyze this improvement [40,56]. However, accounting for the fact that β-D-Gal*f*O*p*NP possesses a good leaving group, we postulate that improvements in self-condensation of this compound are in part due to improved binding of the substrate.

4.3. Insights related to secondary hydrolysis

Data related to the mutants investigated herein clearly demonstrate how secondary hydrolysis of glycosynthetic products can be modulated. In the case of the synthesis of D-galactofuranosyl disaccharides, secondary hydrolysis is clearly reduced or even imperceptible. One reason for this might be suboptimal positioning of the scissile inter-glycosidic bond with respect to the catalytic residues. Indeed, comparison of the 3D models of *Tx*Abf bound to compounds **1** and **4** respectively reveals that the oxygen atom of the scissile bond in β-D-Gal*f*-(1,2)-β-D-Gal*f*O*p*NP **4** is 4.5 Å distant from Glu176 (acid/base), while this is only distant by 3.1 Å for α-L-Ara*f*-(1,2)-α-L-Ara*f*O*p*NP **1** (Fig. 8)[3]. In this regard, this difference of positioning disappears for the other self-condensation product/enzyme pairs. Therefore, it is plausible that the anomeric carbon of the scissile bond in the D-galactofuranosyl disaccharides

[3] This result need to be confirmed by new docking experiment data.

is poorly orientated for catalysis, this being the result of the additional hydroxymethyl group at the *C*-5 position on the furanose ring bound in subsite -1.

It is worth recalling that compared to R69H-N216W, the introduction of L352M or F26L reveals decreased secondary hydrolysis (v_{HII}) of product **2**. Similarly, the synthesis rates of **2** are decreased, especially for F26L-containing triple-mutant. Although exactly how this affects both self-condensation and secondary hydrolysis is unclear, the modification of the internal volume of the pocket-like subsite -1 and/or the interactions with the donor substrate might be considered to modulate the behavior of the triple-mutants.

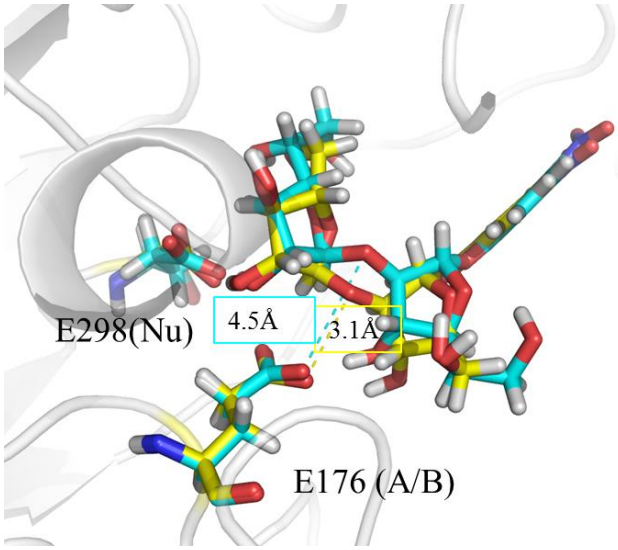


Fig. 8. Superposition of *Tx*Abf bound to **1** (yellow, α-L-Ara*f*-(1,2)-α-L-Ara*f*O*p*NP) and **4** (cyan, β-D-Gal*f*-(1,2)-β-D-Gal*f*O*p*NP). In each case, distances are measured from the inter-glycosidic oxygen (scissile bond) to the OE2 of the catalytic acid/base residue E176.

4.4. Introduction of a tryptophan at position 216 forms a new binding subsite

A key finding in this study is the emergence of new regioselectivity, yielding (1,3)-linked pentofurano- and hexofurano-disaccharides. The introduction of tryptophan at position 216 contributes to this, although as a standalone mutation N216W is insufficient to procure altered regioselectivity. Nevertheless, regarding N216W, introduction of a tryptophan provides the basis for a new π-π interaction between the planar indolyl ring and the *p*NP aglycon moiety. It is relevant that the introduction of tryptophan into other GHs has also produced changes in regioselectivity, either by forcing the repositioning of the glycoside acceptor [57], or by creating a favorable interaction with the acceptor or its linked aromatic moiety [58,59]. It is also noteworthy that in *Tx*Abf, unveiling of the new subsite +2'' does not interfere with the

original subsite +2. This infers that it could be possible through further enzyme engineering to block the subsite +2 and thus obtain an enzyme that procures mainly (1,3)-linked pentofurano- and hexofurano-disaccharides, perhaps with small traces of (1,5)-linked compounds; or inversely for the synthesis of (1,2)-linked difuranosides.

5. Conclusion

This study contributes to our understanding of how *Tx*Abf catalyzes self-condensation of pento- and hexofuranosides and provides new mutants that can be added to the growing toolbox of biocatalysts suitable for glycosynthesis. Self-condensation reactions remain under kinetic control, meaning that high yields can be obtained using higher substrate concentrations, and that regioselectivity can be quite efficiently modulated. Bearing this in mind, it should be possible in the very near future to define reactions catalyzed by mutated enzymes that will procure high yields of relatively pure (1,3)-linked regioisomers. Likewise, with the knowledge of how to enhance the formation of (1,3)-linked regioisomers, it should also be possible to repress their formation, while identifying ways to enhance the regioselective formation of (1,2) linkages.

6. Acknowledgments

The PhD fellowship of J. Zhao was supported by CSC (China Scholarship Council). NMR analyses were performed using facilities at MetaToul (Metabolomics & Fluxomics Facilities, Toulouse, France, www.metatoul.fr), which is part of the national infrastructure MetaboHUB (The French National infrastructure for metabolomics and fluxomics, www.metabohub.fr) and is supported by grants from the Région Midi-Pyrénées, the European Regional Development Fund, SICOVAL, IBiSa-France, CNRS, and INRA. G. Lippens (TBI) is gratefully acknowledged for insightful discussions and technical development of the NMR pseudo-2D kinetics experiments.

Declaration of Competing Interest

The authors declare no conflict of interest.

Author contributions

R.F. and M.J.O designed and supervised the study. J.Z. performed the biochemical experiment. J.Z., I.A. and J.E. performed the docking experiment. All authors analyzed data and contributed to write the manuscript

References

[1] Peltier P, Euzen R, Daniellou R, Nugier-Chauvin C, Ferrières V. Recent knowledge and innovations related to hexofuranosides: structure, synthesis and applications. Carbohydr Res 2008;343:1897–923. https://doi.org/10.1016/j.carres.2008.02.010.

[2] Chlubnová I, Filipp D, Spiwok V, Dvoráková H, Daniellou R, Nugier-Chauvin C, et al. Enzymatic synthesis of oligo-D-galactofuranosides and L-arabinofuranosides: from molecular dynamics to immunological assays. Org Biomol Chem 2010;8:2092–102. https://doi.org/10.1039/b926988f.

[3] Pedersen LL, Turco SJ. Galactofuranose metabolism: a potential target for antimicrobial chemotherapy. Cell Mol Life Sci 2003;60:259–66. https://doi.org/10.1007/s000180300021.

[4] Houseknecht JB, Lowary TL. Chemistry and biology of arabinofuranosyl- and galactofuranosyl-containing polysaccharides. Curr Opin Chem Biol 2001;5:677–82. https://doi.org/10.1016/S1367-5931(01)00265-4.

[5] Gruppen H, Hoffmann R., Kormelink FJ., Voragen AGJ, Kamerling J., Vliegenthart JFG. Characterisation by 1H NMR spectroscopy of enzymically derived oligosaccharides from alkali-extractable wheat-flour arabinoxylan. Carbohydr Res 1992;233:45–64. https://doi.org/10.1016/S0008-6215(00)90919-4.

[6] Seiboth B, Metz B. Fungal arabinan and L-arabinose metabolism. Appl Microbiol Biotechnol 2011;89:1665–73. https://doi.org/10.1007/s00253-010-3071-8.

[7] Classen B, Baumann A, Utermoehlen J. Arabinogalactan-proteins in spore-producing land plants. Carbohydr Polym 2019;210:215–24. https://doi.org/10.1016/j.carbpol.2019.01.077.

[8] Kawabata Y, Kaneko S, Kusakabe I, Gama Y. Synthesis of regioisomeric methyl α-L-arabinofuranobiosides. Carbohydr Res 1995;267:39–47. https://doi.org/10.1016/0008-6215(94)00290-V.

[9] Joseleau JP, Chambat G, Vignon M, Barnoud F. Chemical and 13C N.M.R. studies on two arabinans from the inner bark of young stems of *Rosa glauca*. Carbohydr Res 1977;58:165–75. https://doi.org/10.1016/S0008-6215(00)83412-6.

[10] Marino C, Baldoni L. Synthesis of D-Galactofuranose-Containing Molecules: Design of Galactofuranosyl Acceptors. ChemBioChem 2014;15:188–204. https://doi.org/10.1002/cbic.201300638.

[11] Lowary TL. Synthesis and conformational analysis of arabinofuranosides, galactofuranosides and fructofuranosides. Curr Opin Chem Biol 2003;7:749–56. https://doi.org/10.1016/j.cbpa.2003.10.005.

[12] Jones C, Todeschini AR, Agrellos OA, Previato JO, Mendonça-Previato L. Heterogeneity in the Biosynthesis of Mucin O -Glycans from *Trypanosoma cruzi* Tulahuen Strain with the Expression of Novel Galactofuranosyl-Containing

Oligosaccharides †. Biochemistry 2004;43:11889–97. https://doi.org/10.1021/bi048942u.

[13] Leal JA, Gómez-Miranda B, Prieto A, Domenech J, Ahrazem O, Bernabé M. Possible chemotypes from cell wall polysaccharides, as an aid in the systematics of Penicillium and its teleomorphic states Eupenicillium and Talaromyces. Mycol Res 1997. https://doi.org/10.1017/S0953756297004012.

[14] Prieto A, Bernabé M, Leal JA. Isolation, purification and chemical characterization of alkali-extractable polysaccharides from the cell walls of Talaromyces species. Mycol Res 1995. https://doi.org/10.1016/S0953-7562(09)80318-3.

[15] Leitão EA, Bittencourt VCB, Haido RMT, Valente AP, Peter-Katalinic J, Letzel M, et al. β-Galactofuranose-containing O-linked oligosaccharides present in the cell wall peptidogalactomannan of *Aspergillus fumigatus* contain immunodominant epitopes. Glycobiology 2003;13:681–92. https://doi.org/10.1093/glycob/cwg089.

[16] Sørum U, Robertsen B, Kenne L. Structural studies of the major polysaccharide in the cell wall of *Renibacterium salmoninarum.* Carbohydr Res 1998. https://doi.org/10.1016/S0008-6215(97)10071-4.

[17] Trincone A. Uncommon glycosidases for the enzymatic preparation of glycosides. Biomolecules 2015;5:2160–83. https://doi.org/10.3390/biom5042160.

[18] Marino C, Rinflerch A, De Lederkremer RM. Galactofuranose antigens, a target for diagnosis of fungal infections in humans. Futur Sci OA 2017;3:FSO199. https://doi.org/10.4155/fsoa-2017-0030.

[19] Konishi T, Ono H, Ohnishi-Kameyama M, Kaneko S, Ishii T. Identification of a mung bean arabinofuranosyltransferase that transfers arabinofuranosyl residues onto (1,5)-linked α-L-arabino-oligosaccharides. Plant Physiol 2006;141:1098–105. https://doi.org/10.1104/pp.106.080309.

[20] Rose NL, Completo GC, Lin SJ, McNeil M, Palcic MM, Lowary TL. Expression, purification, and characterization of a galactofuranosyltransferase involved in *Mycobacterium tuberculosis* arabinogalactan biosynthesis. J Am Chem Soc 2006;128:6721–9. https://doi.org/10.1021/ja058254d.

[21] Belanova M, Dianiskova P, Brennan PJ, Completo GC, Rose NL, Lowary TL, et al. Galactosyl Transferases in Mycobacterial Cell Wall Synthesis. J Bacteriol 2008;190:1141–5. https://doi.org/10.1128/JB.01326-07.

[22] Castilho A, Beihammer G, Pfeiffer C, Göritzer K, Montero-Morales L, Vavra U, et al. An oligosaccharyltransferase from *Leishmania* major increases the N-glycan occupancy on recombinant glycoproteins produced in Nicotiana benthamiana. Plant Biotechnol J 2018;16:1700–9. https://doi.org/10.1111/pbi.12906.

[23] Kallolimath S, Gruber C, Steinkellner H, Castilho A. Promoter Choice Impacts the efficiency of plant glyco-engineering. Biotechnol J 2018;13:1–7. https://doi.org/10.1002/biot.201700380.

[24] Montero-Morales L, Steinkellner H. Advanced plant-based glycan engineering. Front Bioeng Biotechnol 2018;9:1–8. https://doi.org/10.3389/fbioe.2018.00081.

[25] Nidetzky B, Gutmann A, Zhong C. Leloir Glycosyltransferases as Biocatalysts for Chemical Production. ACS Catal 2018;8:6283–300. https://doi.org/10.1021/acscatal.8b00710.

[26] Richards MR, Lowary TL. Chemistry and biology of galactofuranose-containing polysaccharides. ChemBioChem 2009;10:1920–38. https://doi.org/10.1002/cbic.200900208.

[27] Houseknecht JB, Lowary TL. Oligofuranosides containing conformationally restricted residues: Synthesis and conformational analysis. J Org Chem 2002;67:4150–64. https://doi.org/10.1021/jo011127p.

[28] Koshland DE. Stereochemistry and the mechanism of enzymatic reactions. Biol Rev 1953;28:416–36. https://doi.org/10.1111/j.1469-185X.1953.tb01386.x.

[29] Barker SA, Bourne EJ, Stacey M, Ward RB. Some paper-chromatographic studies with *Aspergillus niger* '152' transfructosylase. Biochem J 1958;69:60–2. https://doi.org/10.1042/bj0690060.

[30] Park YK, Almeida MM. Production of fructooligosaccharides from sucrose by a transfructosylase from *Aspergillus niger*. World J Microbiol Biotechnol 1991;7:331–4. https://doi.org/10.1007/BF00329399.

[31] Rémond C, Ferchichi M, Aubry N, Plantier-Royon R, Portella C, O'Donohue MJ. Enzymatic synthesis of alkyl arabinofuranosides using a thermostable α-L-arabinofuranosidase. Tetrahedron Lett 2002;43:9653–5. https://doi.org/10.1016/S0040-4039(02)02381-X.

[32] Bissaro B, Durand J, Planas A, Monsan P, Biarnés X, Planas A, et al. Molecular Design of Non-Leloir Furanose-Transferring Enzymes from an α-L-Arabinofuranosidase: A rationale for the engineering of evolved transglycosylases. ACS Catal 2015;5:4598–611. https://doi.org/10.1021/acscatal.5b00949.

[33] Rémond C, Plantier-Royon R, Aubrya N, O'Donohue MJJ,. An original chemoenzymatic route for the synthesis of β-D-galactofuranosides using an α-l-arabinofuranosidase. Carbohydr Res 2005;340:637–44. https://doi.org/10.1016/j.carres.2005.01.016.

[34] Cantarel BI, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycogenomics. Nucleic Acids Res 2009;37:233–8. https://doi.org/10.1093/nar/gkn663.

[35] Helbert W, Poulet L, Drouillard S, Mathieu S, Loiodice M, Couturier M, et al. Discovery of novel carbohydrate-active enzymes through the rational exploration of the protein sequences space. Proc Natl Acad Sci 2019;116:201815791. https://doi.org/10.1073/pnas.1815791116.

[36] Matsunaga E, Higuchi Y, Mori K, Yairo N, Oka T, Shinozuka S, et al. Identification and Characterization of a Novel Galactofuranose-Specific β-D-Galactofuranosidase from Streptomyces Species. PLoS One 2015;10:e0137230. https://doi.org/10.1371/journal.pone.0137230.

[37] Matsunaga E, Higuchi Y, Mori K, Yairo N, Toyota S, Oka T, et al. Characterization of a PA14 domain-containing galactofuranose-specific β-D-galactofuranosidase from *Streptomyces* sp. Biosci Biotechnol Biochem 2017;81:1314–9. https://doi.org/10.1080/09168451.2017.1300518.

[38] Debeche T, Cummings N, Connerton I, Debeire P, O'Donohue MJ. Genetic and Biochemical Characterization of a Highly Thermostable alpha-L-Arabinofuranosidase from *Thermobacillus xylanilyticus.* Appl Environ Microbiol 2000;66:1734–6. https://doi.org/10.1128/AEM.66.4.1734-1736.2000.

[39] Debeche T, Bliard C, Debeire P, O'Donohue MJ. Probing the catalytically essential residues of the α-L-arabinofuranosidase from *Thermobacillus xylanilyticus*. Protein Eng Des Sel 2002;15:21–8. https://doi.org/10.1093/protein/15.1.21.

[40] Bissaro B, Monsan P, Fauré R, O'Donohue MJ. Glycosynthesis in a waterworld: new insight into the molecular basis of transglycosylation in retaining glycoside hydrolases. Biochem J 2015;467:17–35. https://doi.org/10.1042/BJ20141412.

[41] Rémond C, Plantier-Royon R, Aubry N, Maes E, Bliard C, O'Donohue MJ. Synthesis of pentose-containing disaccharides using a thermostable α-L-arabinofuranosidase. Carbohydr Res 2004;339:2019–25. https://doi.org/10.1016/j.carres.2004.04.017.

[42] Arab-Jaziri F, Bissaro B, Tellier C, Dion M, Fauré R, O'Donohue MJ. Enhancing the chemoenzymatic synthesis of arabinosylated xylo-oligosaccharides by GH51 α-L-arabinofuranosidase. Carbohydr Res 2015;401:64–72. https://doi.org/10.1016/j.carres.2014.10.029.

[43] Arab-Jaziri F, Bissaro B, Dion M, Saurel O, Harrison D, Ferreira F, et al. Engineering transglycosidase activity into a GH51 α-L-arabinofuranosidase. N Biotechnol 2013;30:536–44. https://doi.org/10.1016/j.nbt.2013.04.002.

[44] Zhao J, Tandrup T, Bissaro B, Barbe S, André I, Dumon C, et al. Probing the determinants of the transglycosylation/hydrolysis partition in a retaining α-L-arabinofuranosidase. Unpublished 2020.

[45] Bissaro B, Saurel O, Arab-jaziri F, Saulnier L, Milon A, Tenkanen M, et al. Mutation of a pH-modulating residue in a GH51 α-L-arabinofuranosidase leads to a severe reduction of the secondary hydrolysis of transfuranosylation products. Biochim Biophys Acta - Gen Subj 2014;1840:626–36. https://doi.org/10.1016/j.bbagen.2013.10.013.

[46] Glasoe PK, Long FA. Use of glass electrodes to measure acidities in deuterium oxide 1,2. J Phys Chem 1960;64:188–90. https://doi.org/10.1021/j100830a521.

[47] Gottlieb HE, Kotlyar V, Nudelman A. NMR chemical shifts of common laboratory solvents as trace impurities. J Org Chem 1997;62:7512–5. https://doi.org/10.1021/jo971176v.

[48] Krieger E, Vriend G. YASARA View - molecular graphics for all devices - from smartphones to workstations. Bioinformatics 2014;30:2981–2. https://doi.org/10.1093/bioinformatics/btu426.

[49] Case DA, Betz RM, Cerutti DS, Cheatham Iii TE, Darden TA, Duke RE, et al. AMBER 2016. Univ California, San Fr 2016;810.

[50] Hanwell MD, Curtis DE, Lonie DC, Vandermeersch T, Zurek E, Hutchison GR. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. J Cheminform 2012;4:17. https://doi.org/10.1186/1758-2946-4-17.

[51] DeLano, W.L. (2002) The PyMOL Molecular Graphics System. Delano Scientific, San Carlos. n.d.

[52] Arab-Jaziri F, Bissaro B, Barbe S, Saurel O, Débat H, Dumon C, et al. Functional roles of H98 and W99 and β2α2 loop dynamics in the α-L-arabinofuranosidase from Thermobacillus xylanilyticus. FEBS J 2012;279:3598–611. https://doi.org/10.1111/j.1742-4658.2012.08720.x.

[53] Chlubnová I, Králová B, Dvořáková H, Hošek P, Spiwok V, Filipp D, et al. The

versatile enzyme Ara51 allowed efficient synthesis of rare pathogen-related β-D-galactofuranosyl-pyranoside disaccharides. Org Biomol Chem 2014;12:3080–9. https://doi.org/10.1039/c3ob42519c.

[54] Paës G, Skov LK, O'Donohue MJ, Rémond C, Kastrup JS, Gajhede M, et al. The structure of the complex between a branched pentasaccharide and *Thermobacillus xylanilyticus* GH-51 arabinofuranosidase reveals xylan-binding determinants and induced fit. Biochemistry 2008;47:7441–51. https://doi.org/10.1021/bi800424e.

[55] Euzen R, Lopez G, Nugier-Chauvin C, Ferrières V, Plusquellec D, Rémond C, et al. A Chemoenzymatic approach for the synthesis of unnatural disaccharides containing D-galacto- or D-fucofuranosides. European J Org Chem 2005;2005:4860–9. https://doi.org/10.1002/ejoc.200500525.

[56] Johnson KA. A century of enzyme kinetic analysis, 1913 to 2013. FEBS Lett 2013;587:2753–66. https://doi.org/10.1016/j.febslet.2013.07.012.

[57] Tran V, Hoffmann L, Rabiller C, Tellier C, Dion M. Rational design of a GH1 β-glycosidase to prevent self-condensation during the transglycosylation reaction. Protein Eng Des Sel 2010;23:43–9. https://doi.org/10.1093/protein/gzp068.

[58] Madhuprakash J, Singh A, Kumar S, Sinha M, Kaur P, Sharma S, et al. Structure of chitinase D from Serratia proteamaculans reveals the structural basis of its dual action of hydrolysis and transglycosylation. Int J Biochem Mol Biol 2013;4:166–78.

[59] Rosengren A, Reddy SK, Sjöberg JS, Aurelius O, Logan DT, Kolenová K, et al. An Aspergillus nidulans β-mannanase with high transglycosylation capacity revealed through comparative studies within glycosidase family 5. Appl Microbiol Biotechnol 2014;98:10091–104. https://doi.org/10.1007/s00253-014-5871-8.

Supplementary Information

Synthesis of α-L-Ara*f* and β-D-Gal*f* series difuranosides using mutants of a GH51 α-L-arabinofuranosidase

Jiao Zhao, Jérémy Esque, Isabelle André, Michael J. O'Donohue*, Régis Fauré*

TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France

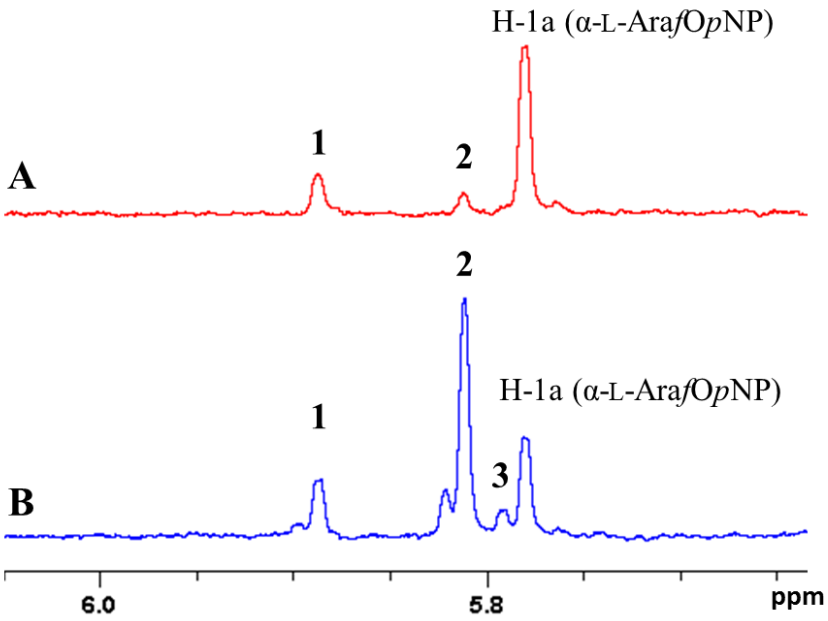


Fig. S1. ^{1}H NMR spectra of anomeric proton signals (H-1a) from the O*p*NP-linked furanosyl moiety of self-condensation products (**1**, **2** and **3**), and remaining α-L-Ara*f*O*p*NP, catalyzed by (**A**) *Tx*Abf and (**B**) R69H-N216W-L352M.

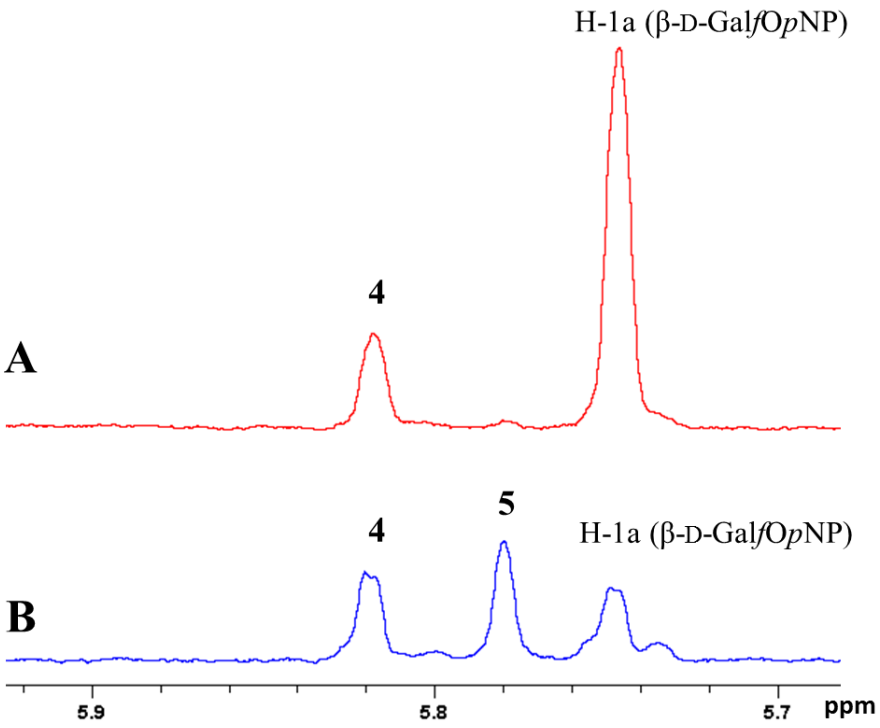


Fig. S2. ^{1}H NMR spectra of anomeric proton signals (H-1a) from the O*p*NP-linked furanosyl moiety of self-condensation products (**4** and **5)**, and remaining β-D-Gal*f*O*p*NP, catalyzed by (**A**) *Tx*Abf and (**B**) R69H-N216W-L352M.

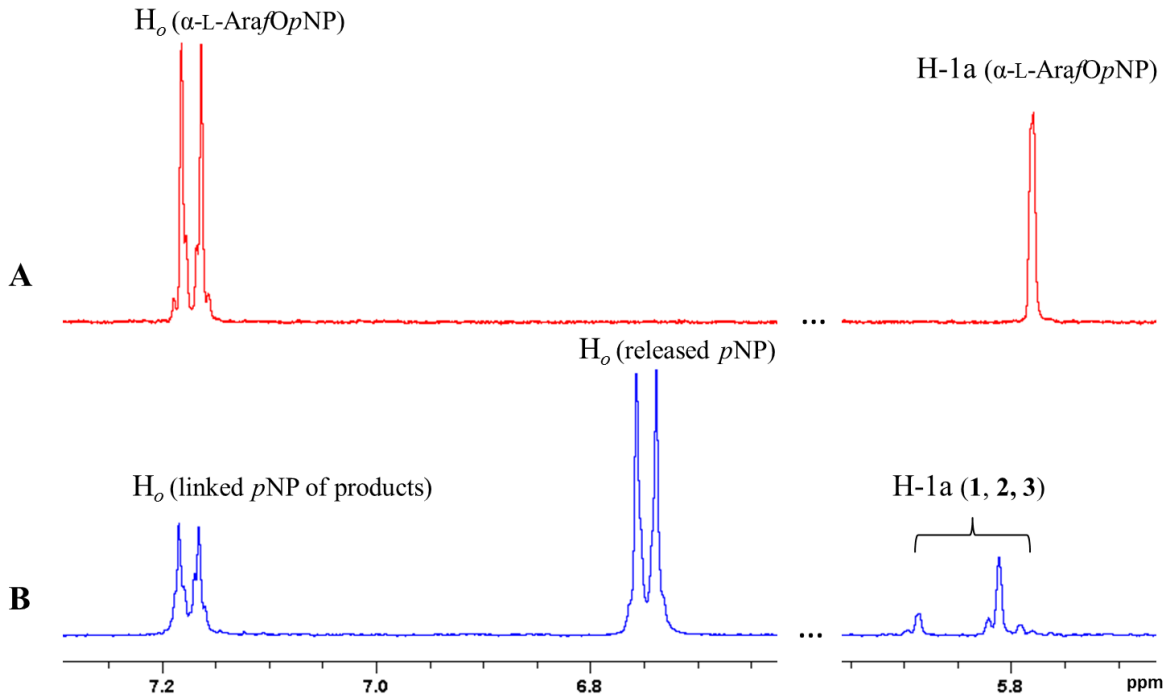


Fig. S3. NMR monitoring of self-condensation by *Tx*Abf mutants. ^{1}H NMR significant aromatic and anomeric signals of α-L-Ara*f*O*p*NP and self-condensation products (**1**, **2** and **3**: **A**) 5 mM α-L-Ara*f*O*p*NP; **B**) after 8 h incubation with R69H-N216W-L352M (0.45 µM). Both spectra were performed at 45 °C and pH 7 in buffered 10% D_2O.

Aromatic H_m (not show in this figure) and H_o signals of both self-condensation products and donor substrate displayed undifferentiated chemical shifts, meaning that the sum of both self-condensation products and remaining α-L-Ara*f*O*p*NP can be monitored by integrating their superimposed aromatic signals. At the meantime, the remaining α-L-Ara*f*O*p*NP can be extracted from its anomeric signal, indicating that the overall yield of self-condensation products can be alternatively calculated by substracting the value of the integral of the anomeric donor signal from that for the average integral of *ortho* and *meta* aromatic protons of the linked O*p*NP signals.

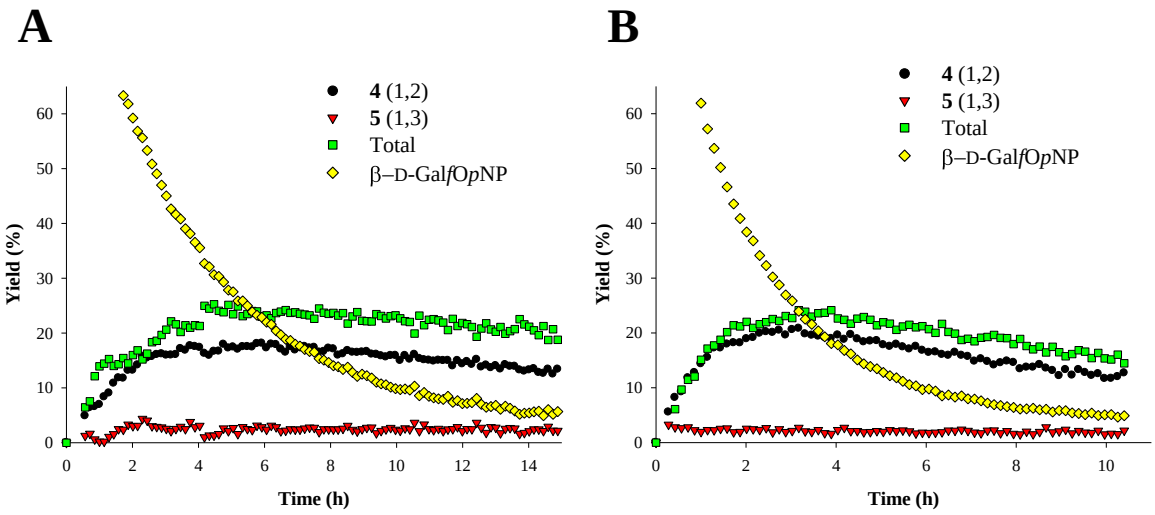


Fig. S4. Monitoring of β-D-Gal*f*O*p*NP consumption and the apparition of the different self-condensation products **4-5** as a function of time in reactions catalyzed by single-mutant (**A**) R69H and (**B**) N216W.

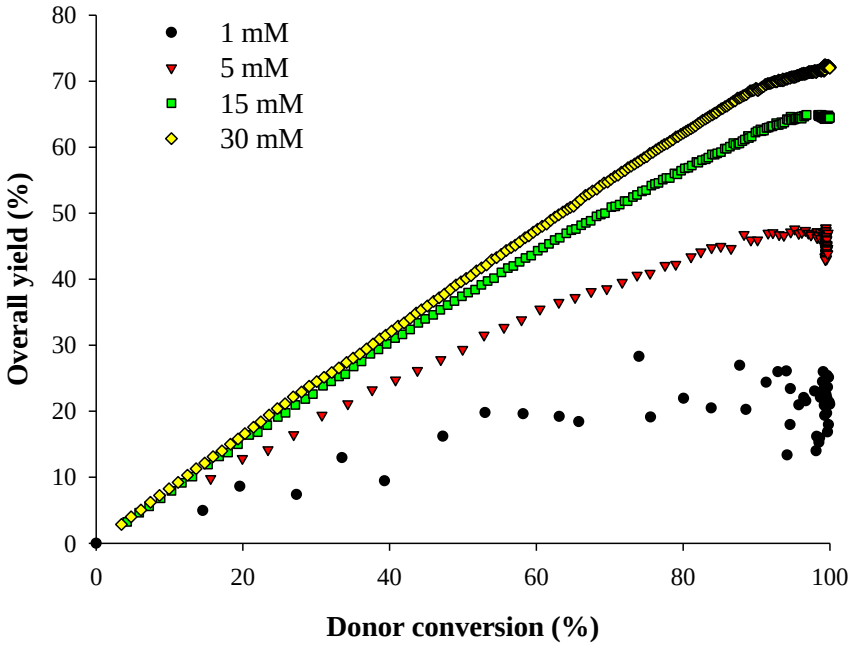


Fig. S5. Overall self-condensation yield as a function of donor consumption in R69H-N216W-catalyzed reaction using different concentrations of β-D-Gal*f*O*p*NP.

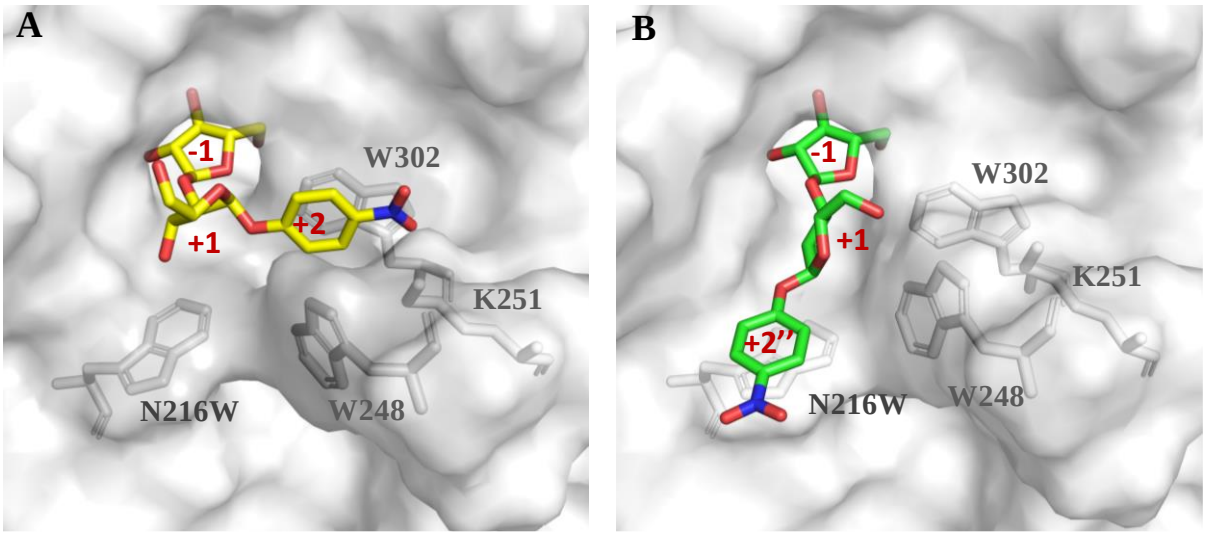


Fig. S6. R69H-N216W-L352M bound to (**A**) α-L-Ara*f*-(1,2)-α-L-Ara*f*O*p*NP **1** and (**B**) α-L-Ara*f*-(1,3)-α-L-Ara*f*O*p*NP **2** reveals the existence of two alternative binding modes.

Table S1

Apparent synthesis and secondary hydrolysis rates of the main regioisomer **2**, α-L-Ara*f*-(1,3)-α-L-Ara*f*O*p*NP, for R69H-N216W-containing mutants.

Enzyme	**Synthesis rate v_S ($mM.h^{-1}.\mu M^{-1}$)**	**2nd Hydrolysis rate v_{HII} ($mM.h^{-1}.\mu M^{-1}$)**	**v_S/v_{HII}[a]**
R69H-N216W	1.28	0.47	2.7
R69H-N216W-L352M	0.51	0.17	3.0
F26L-R69H-N216W	0.22	0.12	1.8

[a] Apparent self-condensation (v_S, initial phase of the plot)/secondary hydrolysis (v_{HII}, after maximum yield) rates were derived from the time-course monitoring of self-condensation product **2** yield. This ratio depicts the ability of the enzymes to perform self-condensation towards the synthesis of the main regioisomer **2** while considering their secondary hydrolysis capabilities.

Chapter IV.

Investigation of transgalactofuranosylation activity in an α-L-arabinofuranosidase

The discovery many years ago that *Tx*Abf can use β-D-Gal*f* as a donor and procure high yields of transglycosylation products was the subject of some excitement and the source of numerous questions. In the paper published in 2005 (Rémond, 2005), it was already suggested that β-D-Gal*f* somehow binds unproductively, penalizing water-mediated deglycosylation. Subsequently few opportunities arose to allow us to pursue this work, even though progress was made by another research group at the University of Rennes. Therefore, one of the explicit aims of this doctoral work was to pursue the study of transgalactofuranosylation and attempt to provide new examples of how *Tx*Abf might be used for the synthesis of biologically-relevant glycoconjugates present in pathogenic microorganisms.

Once again, capitalizing on the availability of recently-acquired, potent *Tx*Abf mutants, such as R69H-N216W, the research presented here used site-saturation mutagenesis to probe transgalactofuranosylation activity. Significant results in this work are related to the fact that some mutants improved regioselectivity, a finding that is thoroughly investigated using STD-NMR. From a practical standpoint, this work also revealed the potential of *Tx*Abf mutants for the preparation of two biologically relevant glycoconjugates that are present in *Actinobacillus pleuropneumoniae* and *Steptococcus thermophiles Sfi39* (for β-D-Gal*f*-(1,3)-α-D-Glc*p*), and *Trypanosoma cruzi* (for β-D-Gal*f*-(1,4)-α-D-Glc*p*NAc). Overall, these results confirmed that the engineering of the donor subsite is a useful way to diversify *Tx*Abf for use as a tool to access D-Gal*f*-containing disaccharides.

Investigation of transgalactofuranosylation activity in an α-L-arabinofuranosidase

Jiao Zhao,[a], Jérémy Esque,[a] Olivier Saurel,[b] Isabelle André,[a] Michael J. O'Donohue[*a] and Régis Fauré[*a]

[a]*TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France. E-mail: michael.odonohue@inra.fr, regis.faure@insa-toulouse.fr; Tel: +33 5 6155 9410*

[b]*Institut de Pharmacologie et de Biologie Structurale, Université de Toulouse, CNRS, Université Paul Sabatier, Toulouse, France*

Abstract

The GH51 α-L-arabinofuranosidase from *Thermobacillus xylanilyticus* (*Tx*Abf) possesses versatile catalytic properties, displaying not only the ability to hydrolyze glycosidic linkages but also to synthesize furanosyl-containing oligosaccharides. Herein, *Tx*Abf mutant R69H-N216W was investigated to evaluate its performances for transgalactofuranosylation reactions. Overall yield of transgalactofuranosylation onto α-D-Glc*p* moiety reached 65%, being increased 2-fold compared to the wild-type enzyme. Site saturation mutagenesis was applied to *Tx*Abf in order to target four hydrophobic residues in the vicinity of the C-5 hydroxymethyl group of the β-D-Gal*f*O*p*NP donor. Consequently, two triple-mutant, R69H-N216W-L314N and R69H-N216W-L352G, displayed additional enhancement of both transgalactofuranosylation yield (74% and 73%) and regioselectivity (91% and 87%) towards the synthesis of the biologically relevant β-D-Gal*f*-(1,3)-α-D-Glc*p* motif. Correlated with an increased affinity for the donor, STD-NMR effects demonstrated that the recognition towards D-Gal*f* moiety was improved for template R69H-N216W and its two L314N- and L352G-containing triple-mutants. Distinct donor-enzyme binding patterns seem to drive the increased yields and regioselectivity. Overall, these results confirmed that the engineering of the donor subsite might be used for the diversification of *Tx*Abf as a powerful tool for the synthesis D-galactofuranosyl-containing disaccharides.

Keywords: retaining glycoside hydrolase, transglycosylation, site-saturation mutagenesis, STD-NMR, D-galactofuranosyl-containing oligosaccharides

Abbreviations

Abfs, α-L-arabinofuranosidases; α-L-Ara*f*, α-L-arabinofuranosyl unit; α-L-Ara*f*O*p*NP, 4-nitrophenyl α-L-arabinofuranoside; A_{STD}, STD amplification factor ; β-D-Gal*f*, β-D-galactofuranosyl unit; β-D-Gal*f*O*p*NP, 4-nitrophenyl β-D-galactofuranoside; GlfT, galactofuranosyl transferases; rGH, retaining glycoside hydrolase; *p*NP, 4-nitrophenol; rGH, retaining glycoside hydrolase; SA, specific activity; SSM, site-saturation mutagenesis; STD, saturation transfer difference; *Tx*Abf, α-L-arabinofuranosidase from *Thermobacillus xylanilyticus*; T/H, transglycosylation/hydrolysis ratio.

1. Introduction

D-Galactofuranose (D-Gal*f*), the five membered ring form of D-galactose, is widespread in naturally occurring glycoconjugates found in bacteria, protozoa, fungi, plants and archaebacteria.[1–4] Importantly, D-Gal*f* is often present in pathogenic microorganisms, forming part of cell wall glycoconjugates that act as antigenic epitopes, trigging immunogenic responses in humans.[5,6] Taking into account the fact that mammals are unable to synthesize D-Gal*f*, it is possible to consider this sugar as a specific marker of pathogens and thus an interesting target for the development of diagnostic tools and medical applications strategies.[1,6–9]

The synthesis of D-Gal*f*-containing oligosaccharides can be achieved using either chemical or biosynthetic apporaches.[1,9] The former involves tedious protection/deprotection cycles and activation strategies to control both the regioselectivity and stereochemistry of the target glycosidic linkage.[10,11] Regarding biosynthesis, *in vivo* and *in vitro* approaches are possible, both of these generally involving the use of galactofuranosyl transferases (GlfT).[9] Using such a strategy, it has been shown that GlfTs readily catalyse the *in vitro* elongation of D-galactofuranosyl chains, generating biologically relevant products.[12–14] Nevertheless, GlfTs use UDP-Gal*p* as starting material and require the intervention of UDP-galactopyranose mutase to convert it into UDP-Gal*f*, which is the donor for GlfTs, meaning that the overall reaction is relatively costly and complex.[4,9] An alternative way to synthesize D-galactofuranose-based conjugates is to use retaining glycoside hydrolases (GHs) that perform transglycosylation.[15–17] Ideally, to achieve this one should target β-D-galactofuranosidases (EC 3.2.1.146).[18] However, at this point in time very little is known about these enzymes, since only a small number of β-D-galactofuranosidases have so far been identified.[19] These

include *exo*- and *endo*-β-D-galactofuranosidase from fungi,[20–23] bacteria,[24] and protozoa.[25] The most recently discovered β-D-galactofuranosidases belong to families GH2, 5 and 43 of the CAZy classification (www.cazy.org; www.cazypedia.org),[18,19,26,27] but the glycosynthetic abilities of these enzymes have not yet been assessed.

Previous work has revealed that well-studied GH51 α-L-arabinofuranosidases (Abfs) possess some ability to catalyze the synthesis of Gal*f*-containing conjugates,[2,15,28,29] the rationale behind this being the chemical similarity of α-L-arabinofuranose (L-Ara*f*) and D-Gal*f* (the latter being the 5-hydroxymethyl analog of the former) coupled to catalytic promiscuity of the enzymes. Using Abfs, it has thus far been possible to demonstrate the *in vitro* synthesis of homo-disaccharides such as β-D-Gal*f*-(1,2)-β-D-Gal*f*O*p*NP, β-D-Gal*f*-(1,6)-β-D-Gal*f*O*p*NP and other D-galactofuranosides containing up to five β-D-Gal*f* moieties.[2,15] Moreover, the synthesis of other D-Gal*f*-containing conjugates has been achieved using diverse acceptors groups such as pyranosides[15,28], alcohols[29,30], thiols [31], and acids[29].

Among the Abfs that can synthesize D-galactofuranoside conjugates, the Abf produced by *Thermobacillus xylanilyticus* (*Tx*Abf) is a particularly well-characterized and studied example. This enzyme is a retaining GH, meaning that it hydrolyses and synthesizes glycosidic bonds using a double displacement mechanism that involves the formation of a transient glycosyl-enzyme intermediate.[32–34] Upon formation of glycosylated *Tx*Abf, hydrolysis can occur if a water molecule attacks the intermediate. However, if a glycoside acceptor attacks the intermediate then transglycosylation is the primary outcome of the reaction. In the case where a single, suitably-activated sugar is present in the reaction mixture, *Tx*Abf catalyzes self-condensation, which is a special case of transglycosylation in which the donor and acceptor molecules are identical and a homo-disaccharide is formed.[15,35] Likewise, *Tx*Abf performs self-condensation when using either α-L-Ara*f*O*p*NP or β-D-Gal*f*O*p*NP as substrate. However, the latter reaction leads to higher disaccharide yields, indicating that hydrolysis is less favorable when β-D-Gal*f*O*p*NP is used as substrate. To explain this difference in reaction outcome, it is assumed that the D-galactofuranosyl-*Tx*Abf intermediate is somehow less apt for water-mediated glycosylation, probably because the hydroxymethyl moiety at the C-5 position of the D-galactofuranoside moiety prevents it from binding in the same manner as α-L-Ara*f* in subsite -1. This assumption is supported by the observation that the K_M value related to the hydrolysis of β-D-Gal*f*O*p*NP is extremely high (>50 mM) compared to that of α-L-Ara*f*O*p*NP hydrolysis and that STD-NMR analysis fails to detect any STD effects or binding when β-D-Gal*f*O*p*NP is used as substrate for *Tx*Abf.[15,36,37]

Previously, studies of the *Tx*Abf mutant R69H-N216W-L352M revealed that this enzyme performs transglycosylation using α-L-Ara*f*O*p*NP as donor and xylotriose as acceptor, procuring up to 70% yield depending on the substrate concentration.[33,38] Moreover, the same mutant enzyme can perform self-condensation of β-D-Gal*f*O*p*NP, with yields reaching 45% (compared to 10% for wild-type *Tx*Abf) [35]. Other work focused on *Tx*Abf and its reaction on β-D-Gal*f*O*p*NP pinpointed several residues (F26, L30, L314 and L352) in the vicinity of the C-5 hydroxymethyl group that might cause unfavorable interactions, and thus determine the specific positioning of the D-Gal*f* moiety in subsite -1 compared to that of L-Ara*f*.[33,34,37] Therefore, combing these results, herein we use the *Tx*Abf mutant R69H-N216W as a prototype for further characterization and engineering, the aim being to determine whether the site-saturation mutagenesis of specific residues lying in the vicinity of subsite -1 can alter and/or improve upon the catalytic properties of this promising glycosynthetic *Tx*Abf variant.

2. Results

2.1. Evaluation of mutant R69H-N216W towards enhanced transgalactofuranosylation

To further study the catalytic properties of R69H-N216W, its ability to perform transgalactofuranosylation reactions using β-D-Gal*f*O*p*NP (donor) and different acceptors was evaluated, in some cases targeting reactions that can produce biologically relevant glycoconjugates (Table 1). The use of α-D-Glc*p*O*p*NP (acceptor) procured the highest yield (51% for β-D-Gal*f*-(1,3)-α-D-Glc*p*O*p*NP) and the best resolved (in terms of separation of ^{1}H NMR signals) product regioisomers. Additionally, except for β-D-Glc*p*NAcO*p*NP, the other acceptors also procured products, albeit in lower yields. It is noteworthy, that most products were (1,2)-, (1,3)- or (1,4)-linked regioisomers, although due to poor signal resolution the presence of (1,6)-linked products couldn't be completely ruled out.

Table 1 Transgalactofuranosylation profile using β-D-Gal*f*O*p*NP (5 mM) as donor and different acceptors (10 mM) with template enzyme R69H-N216W[a]

Acceptor	**Transglycosylation products linkages and yields (%)**[b]				**Natural occurrence**[c]
	β-D-(1,2)	**β-D-(1,3)**	**β-D-(1,4)**	**β-D-(1,6)**	
α-D-Glc*p*O*p*NP	7	51	7 (linkage unknown)		(1,3) *Actinobacillus. pleuropneumoniae* [39]; (1,3) and (1,6) *Streptococcus thermophilus* SFi39[40,41]
β-D-Glc*p*O*p*NP	7	~ 7	<7 (putative)	-	(1,3) *Streptococcus pneumoniae* 33F[42]
α-D-Gal*p*O*p*NP	Low	<32 (+ D-Gal)	25	Low (putative)	(1,3) *Fibrobacter succinogenes*[43–45] (1,4) *Agelas longissima*[44] *A. pleuro pneumoniae serotypes* 3 and 8[39]
β-D-Gal*p*O*p*NP	15	≤15 (putative: (1,3) or (1,4))		Low	(1,3) *Streptococcus pneumoniae 10B/C*[46] *Leuconostoc mesenteroides ssp. cremoris PIA2*[47]
α-D-Glc*p*NAcO*p*NP	22, 5, 2 (linkage undetermined)				(1,4) *Trypanosoma cruzi*[48–50] (1,3) *Shigella* B7[51], *Klebsiella pneumoniae*[52]
β-D-Glc*p*NAcO*p*NP	10 (linkage undetermined)				(1,3) *Klebsiella pneumoniae*[52], *Shigella* B3[51]

[a] Transglycosylation reactions were performed at 45 °C and pH 7 in buffered 10% D_2O. R69H-N216W was used at 2 µM. [b] Anomeric proton signals of the products were extracted from published data.[28] [c] Products that are naturally produced by microbial sources.

2.2. Site-saturation mutagenesis on R69H-N216W and screening of transglycosylating mutants

Previously, we reported that *Tx*Abf has very low affinity for β-D-Gal*f*O*p*NP (K_M >50 mM), with STD-NMR producing a barely detectable signal.[15,36,37] Therefore, to understand how the extra hydroxylmethyl group at C-5 perturbs β-D-Gal*f*O*p*NP binding a docking experiment was performed on R69H-N216W (Fig. 1). This revealed that three leucines (L30, L314 and L352)

and a phenylalanine (F26) surround the extra hydroxymethyl group at C-5, forming a putative hydrophobic pocket. Therefore, attempting to further enhance the synthetic potency of R69H-N216W, site-saturation mutagenesis (SSM) was performed on this template targeting F26, L30, L314 and L352.

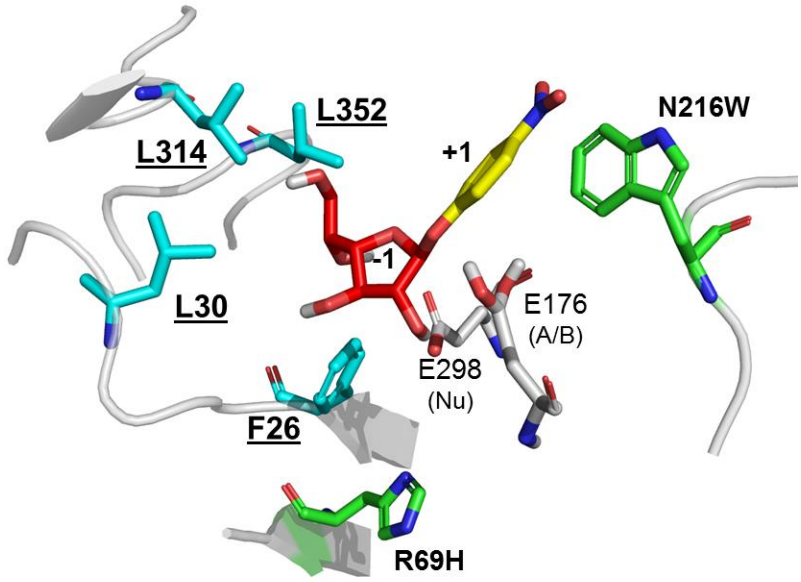


Fig. 1. Docking of β-D-Gal*f*O*p*NP (in red for D-Gal*f* and yellow for *p*NP) with R69H-N216W model (constructed from R69H-N216W-L352M[38] and L352 from wild-type (PDB: 2VRK)) Residues F26, L30, L314 and L352 surrounding the extra hydroxymethyl group at the C-5 position of β-D-Gal*f*O*p*NP are highlighted (in cyan) as well as the catalytic acid/base (A/B) E176 and nucleophile (Nu) E298 residues (in grey), and mutated residues R69H and N216W (in green).

Following construction of the four SSM libraries, preliminary screening was performed using α-D-Glc*p*O*p*NP as the acceptor, consistent with the fact that this compound procures highest product yields in combination with β-D-Gal*f*O*p*NP (donor). Out of a total of 76 possible mutants (*i.e.*, NNK potentially generates 20 variants), 16 were detected. While accounting for the limited sample screening (although 95% coverage of the protein sequence space requires 94 transformants for NNK libraries), [53] it is noteworthy that the mutation of L314 procured 6 active variants, whereas mutation of F26 only procured mutant F26M (Supplementary Table S1). Presumably, this reflects the tolerance of the different positions to modification and indicates that F26 is not very amenable to mutation without consequent loss of activity.[54] furthermore, in the case of multiple occurrences of identical mutants, measured hydrolysis

and transglycosylation activities (*i.e.*, rates of *p*NP release during the first step of the mechanism but in the absence and the presence of acceptor respectively) were quite similar and generated low standard deviations, indicating that the screening method was robust (Supplementary Table S1). Remarkably, although several library isolates displayed modified values relative to the library template, none displayed a significantly higher value of activity ratio. For further work, 7 isolates derived from the mutation of L30, L314 and L352 respectively were targeted for purification (Fig. 2 and Table S1). Except for L30G, all these mutants displayed increased SA_H values), with mutants L314G/N and L352G/C displaying the highest values.

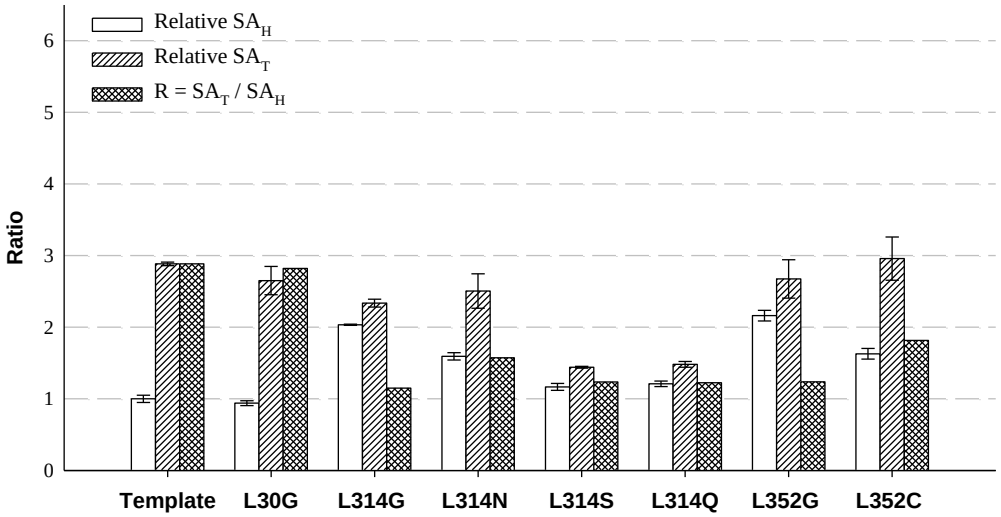


Fig. 2 Relative specific activity (SA) of the purified enzyme in both hydrolysis (SA_H) and transglycosylation (SA_T) mode. Template represents for R69H-N216W, all the other mutants are recombinants based on the template enzyme.

2.3. Transglycosylation profile of candidate mutants

Evaluation of the transglycosylation capability of the 7 mutants described above revealed that all procured overall yields similar (63 ± 3) to that of R69H-N216W and all produced regioisomeric mixtures (Table 2). In each case the main product was β-D-Gal*f*-(1,3)-α-D-Glc*p*O*p*NP, which was further purified and fully characterized by NMR and mass spectrometry. Remarkably, mutants R69H-N216W-L314N and R69H-N216W-L352G clearly procured higher (up to 3-fold for R69H-N216W-L314N) yields of the (1,3)-linked disaccharide compared to *Tx*Abf and up to 18% more than R69H-N216W. Additionally, compared to *Tx*Abf a lower secondary hydrolysis of the (1,3)-linked product was observed for

the R69H-N216W-containing mutants (Supplementary Fig. S1 and Table S2). The more remarkable than overall yield changes was the ability of the mutants to synthesize (1,3)- and (1,2)- products respectively. Expressing this as a ratio (*i.e.* (1,3)-linked yield/(1,2)-linked yield) reveals that while wild-type *Tx*Abf and R69H-N216W favor the synthesis of the (1,3)-linked product with ratios of 1.7 and 8.5 respectively, the mutants perform even more biased reactions with ratios on the range 9.2 to 54 (for R69H-N216W-L30G). In this respect, it is remarkable that in the case of R69H-N216W-L314N the (1,3)-linked regiosiomer represented 94% of the products (*cf.* 78% in the case of R69H-N216W). This is underlined when the composite yield of (1,2)- and (1,3)-linked products is compared with that of other regioisomers. While the value of this ratio is 7.1 for the reaction involving R69H-N216W, it is 31 (highest value) for R69H-N216W-L314N and 6.1 (lowest value) for R69H-N216W-L30G. Therefore, comparing all reactions catalyzed by the mutant enzymes, R69H-N216W-L314N is most regioselective, followed by R69H-N216W and R69H-N216W-L30G. In the latter, the final product is essentially composed of three regioisomers, with the (1,2)-linked product being almost absent. Instead, in addition to the (1,3)-linked product, two other regioisomers (3 and 6% yields respectively) were detected, although it was not possible to determine the linkages. Unlike transgalactofuranosylation mode, comparing R69H-N216W to the triple mutants in self-condensation mode failed to reveal any improvement (Supplementary Table S3 and Fig. S3).

In the course of this study, the relationship between transglycosylation yield, substrate concentration and acceptor/donor ratio were examined (Supplementary Table S4). Accordingly, reactions involving R69H-N216W-L314N and either 15 mM donor and 30 mM acceptor, or 5 mM donor and 20 mM acceptor procured enhanced (70 and 74% respectively) overall yields of the (1,3)-linked product, without affecting overall regioselectivity (maintain at least at 90%). Using this information, preparative scale synthesis of β-D-Gal*f*-(1,3)-α-D-Glc*p*O*p*NP was performed, using R69H-N216W-L314N and a 1-4 donor/acceptor ratio (5 and 20 mM respectively). This procured an overall yield of purified product of 53%, a result that was achieved in 3 steps using an acetylation/deacetylation cycle for purification.

Table 2 Maximal transglycosylation yield of the candidate mutants[a]

Enzyme	Transgalactofuranosylation yield (%)				
	β-D-(1,2) 5.11 ppm	**β-D-(1,3)** 5.26 ppm	**β-D-(1,m)**[b] 4.88 ppm	**β-D-(1,n)**[b] 5.41 ppm	**Overall**[c]
wt	12	20	1	-	33
R69H-N216W	6	51	8	-	65
R69H-N216W-L30G	1	54	3	6	64
R69H-N216W-L314G	4	48	6	-	58
R69H-N216W-L314N	2	60	2	-	64
R69H-N216W-L314S	5	46	7	-	58
R69H-N216W-L314Q	4	50	8	-	62
R69H-N216W-L352G	2	59	5	-	66
R69H-N216W-L352C	4	54	8	-	66

[a] Transglycosylation reactions were performed with 5 mM β-D-Gal*f*O*p*NP as donor and 10 mM α-D-Gal*p*O*p*NP as acceptor at 45°C and pH 7 in buffered 10% D_2O. Transglycosylation products were *para*-nitrophenyl disaccharides that transferring β-D-Gal*f* onto different hydroxyl position of α-D-Glc*p*O*p*NP acceptor. Chemical shifts were determined at 45°C.

[b] Undetermined linkage, β-D-Gal*f*-(1,4/6)-α-D-Glc*p*O*p*NP regioisomers.

[c] Total self-condensation yields were under 7%.

2.4. Kinetic analysis

To further characterize the mutants created in this study, kinetic parameters were studied for reactions catalyzed by a R69H-N216W, R69H-N216W-L314N and R69H-N216W-L352G (Table 3). These reveal that all mutants display quite altered properties when compared to *Tx*Abf. In hydrolysis mode the K_M value was measurable (0.45-4.13 mM) and comparable to that obtained (0.25 mM) for reactions involving *Tx*Abf and α-L-Ara*f*O*p*NP,[35,38] suggesting that recognition of β-D-Gal*f*O*p*NP is enhanced. The catalytic efficiencies (k_{cat}/K_M) of the mutants remains low and comparable to that of *Tx*Abf, low catalytic (hydrolysis) efficiency

being a hallmark of transglycosylases.[16] Regarding k_{cat}, it is noteworthy that its value for (hydrolysis) reactions catalyzed by either R69H-N216W-L314N or R69H-N216W-L352G was approximately 4 to 7 times higher than that of R69H-N216W. However, when k_{cat} was measured for transglycosylation mode, values were relatively similar, but approximately 2 times higher for R69H-N216W.

Table 3 Steady-state kinetic parameters on β-D-Gal*f*O*p*NP in hydrolysis and transglycosylation mode[a]

Enzyme	Reaction mode	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($s^{-1}\cdot mM^{-1}$)	Ns[d] ($s^{-1}\cdot mM^{-1}$)
wt[c]	Hydrolysis	>50	15.02 ± 0.46	0.13	-
R69H-N216W		0.45 ± 0.08	0.06 ± 0.004	0.14	0.008
R69H-N216W-L314N		2.82 ± 0.30	0.27 ± 0.01	0.10	-
R69H-N216W-L352G		4.13 ± 0.33	0.41 ± 0.01	0.10	-
R69H-N216W	Transglycosylation[b]	4.81 ± 0.21	0.76 ± 0.01	0.16	-
R69H-N216W-L314N		3.38 ± 0.44	0.34 ± 0.01	0.10	-
R69H-N216W-L352G		3.69 ± 0.49	0.47 ± 0.02	0.13	-

[a] Assays were carried at 45°C and pH 7 in 50 mM sodium phosphate buffer in triplicates.

[b] In transglycosylation mode, the concentration of α-D-Glc*p*O*p*NP acceptor was kept constant at 30 mM.

[c] The kinetic parameters were referred to previous work at 60 °C in 50 mM sodium acetate buffer pH 5.8.[15,36,37] The K_M value is an estimate, because its measurement was unfeasible due to the limited solubility of the substrate.

[d] A modified Michaelis-Menten equation $SA = SA_{max}[S]/(K_M + [S]) + Ns \cdot [S]$ that contains a nonspecific constant Ns was used for R69H-N216W in hydrolysis mode.

A particularly noteworthy feature of the kinetic analyses is the two-phase reaction trajectory (SA_H against [β-D-Gal*f*O*p*NP]) described by R69H-N216W when operating in hydrolysis mode (Supplementary Fig. S2A). At low substrate concentration model (<1 mM) the reaction can be modelled using Michaelis-Menten theory. However, at higher concentrations the reaction deviates to a linear trajectory that fails to reach saturation.[33,35,38] This unusual

behaviour was not observed for reactions catalysed by R69H-N216W-L314N or R69H-N216W-L352G (Fig. S2A), thus to better understand it a further study was performed using all three mutants, probing the relationship between self-condensation yield and β-D-Gal*f*O*p*NP concentration (Supplementary Fig. S3). This revealed that, irrespective of the enzyme used, increasing substrate concentrations led to increased yield of self-condensation, implying that in transglycosylation mode, all the reactions obey Michaelis-Menten kinetics (Supplementary Fig. S2B). As already mentioned, comparison of individual kinetic parameters of reactions catalysed by the three mutants confirms that when operating in transglycosylation mode all mutants display similar behaviour (Table 3). The use of high acceptor concentration (30 mM) clearly altered the reaction catalyzed by R69H-N216W, increasing the K_M and k_{cat} values 11- and 13-fold respectively, bringing these nearer to the values measured for the other two mutants.

2.5. Enzyme–substrate interactions by STD-NMR

To probe enzyme†-substrate/product interactions, saturation transfer difference (STD)-NMR was employed. To perform this experiment, we successfully created and purified four catalytically-impaired enzymes (starting from *Tx*Abf, R69H-N216W, R69H-N216W-L314N and R69H-N216W-L30G), in which the acid/base residue (E176) was substituted by an alanine. Incubation of these variants (hereafter denoted using the symbol †) with either β-D-Gal*f*O*p*NP or β-D-Gal*f*-(1,3)-α-D-Glc*p*O*p*NP (*i.e.*, donor substrate and transglycosylation product, respectively) confirmed the impotency of the variants, since no hydrolysis was detected. Moreover, incubation of the different enzymes with β-D-Gal*f*-(1,3)-α-D-GlcO*p*NP only revealed very weak STD effects. This is indicative of fairly low binding towards the transglycosylation product for all the enzymes, consistent with similar observations reported previously.[55] Regarding wild-type *Tx*Abf†, consistent with our previous report and the high K_M value (>50 mM),[37] measurements performed using β-D-Gal*f*O*p*NP revealed that the STD signals of β-D-Gal*f* protons were barely detectable. Nevertheless, after prolonged overnight measurements, weak STD intensities could be amplified and used to compare with data related to the mutants.

Using an internal normalization method to process the NMR data (*i.e.*, setting at 100% the most intensive glycosyl proton signal for each enzyme-donor pair, Fig. 3A), the STD effect of H-4 β-D-Gal*f*O*p*NP was most intense in the case of wild-type *Tx*Abf†, whereas that of H-3

was of similar intensity irrespective of the enzyme† involved. These results indicate a slightly different positioning of H-4 in the case of *Tx*Abf†-β-D-Gal*f*O*p*NP, consistent with the conclusion that the double mutation R69H-N216W allows alternative positioning of the β-D-Gal*f* moiety. Regarding R69H-N216W†, H-2, H-3 and H-4 all displayed relative STD effects in the range 83-100% (normalized using the H-3 signal), while the anomeric proton H-1, H-5, H-6a and H-6b were characterized by weaker STD intensities (39-61%). In the case of the two triple mutants, the H-3 signal (set to 100%) was also the strongest, but signals associated with H-2 were reduced to 70-77% compared to R69H-N216W† (93%) and those linked to H-4 were enhanced, being almost 100%. Furthermore, it is notable that the STD intensities of signals arising from H-6b is reversed in the case of triple mutants when compared to those of wild-type *Tx*Abf† or R69H-N216W†. This observation, which is indicative of different donor positioning, is consistent with the altered regioselectivities R69H-N216W-L314N and R69H-N216W-L352G when catalyzing transglycosylation.

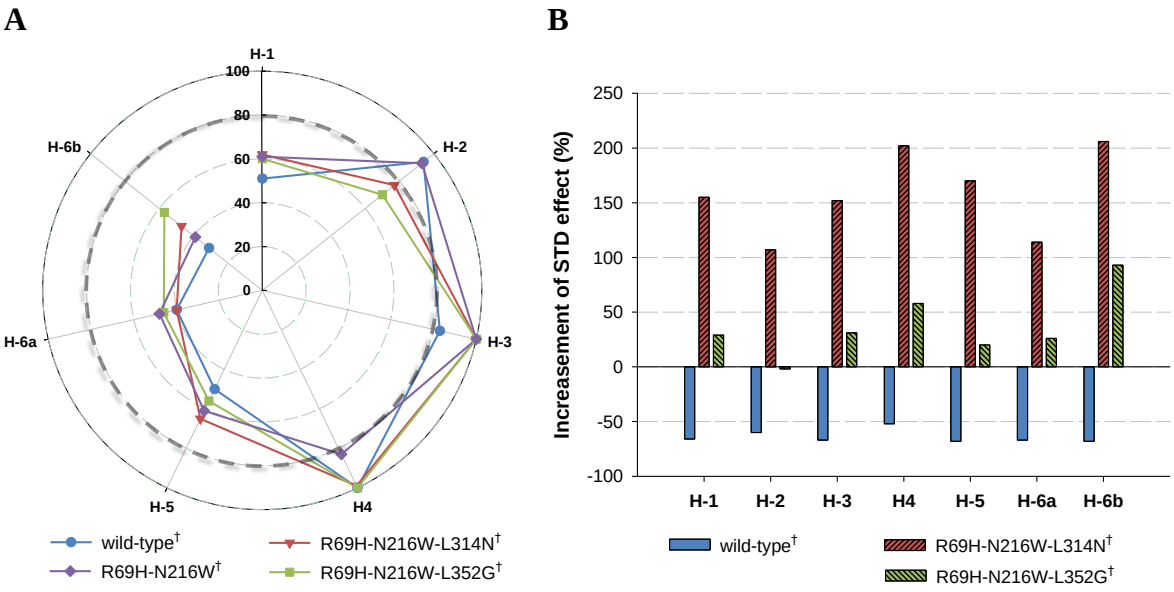


Fig. 3. STD-NMR analysis of enzyme-ligand interactions. (A) STD-NMR effect fingerprint, for each enzyme-donor substrate couple, between galactofuranosyl protons of β-D-Gal*f*O*p*NP and inactivated enzymes†. The values were normalized relative to the most intensive glycosyl proton signal (H-4 for wild-type† and H-3 for mutants†); (B) Standardized STD-NMR effects expressed for each glycosyl proton with inactivated enzymes†, expressed as the increasement (%) of the STD-NMR effects measured between the template enzyme R69H-N216W and β-D-Gal*f*O*p*NP (*i.e.* set R69H-N216W as reference at 0).

An alternative treatment of the NMR data involving the increase (%) of STD-NMR effect (I_{STD}/I_0) for each glycosyl proton with respect to the two triple mutants and wild-type *Tx*Abf was used to compare with the homologous one in R69H-N216W† (*i.e.* R69H-N216W was set at 0) (Fig. 3B). Overall, STD-NMR effect increasement of the three mutants revealed that the enzyme-donor interactions were significantly enhanced when compared to those displayed by wild-type *Tx*Abf†, consistent with the lowered K_M values measured for reactions involving the three mutants and β-D-Gal*f*O*p*NP. Moreover, the triple mutants R69H-N216W-L314N† and R69H-N216W-L352G† displayed even stronger overall interactions (*e.g.*, 152 and 31% higher signals for H-3 respectively, with H-3 being the most intensive glycosyl proton signal for each mutant†-donor pair and thus representative of the overall effect) when compared to R69H-N216W†. Remarkably, the epitope profiles of R69H-N216W-L314N† and R69H-N216W-L352G† were quite similar and differed from that of R69H-N216W†, displaying their own STD-NMR effect fingerprint, notably at H-2 (107% higher and 2% lower), H-4 (202 and 58% higher respectively), H6-a (114 and 26% higher) and H-6b (almost 206 and 93% higher respectively). These distinct binding patterns displayed by the triple mutants with donor substrate possibly correlate with the enhanced regioselectivities of these two enzymes compared to that of R69H-N216W (Table 1).

Finally performing titration experiments using different β-D-Gal*f*O*p*NP concentrations provided a means to determine values for the dissociation constant (K_d). Although it was impossible to determine a K_d value for *Tx*Abf†, the three mutant enzymes† all procured reliable values. Furthermore, plotting the relationship between the STD amplification factor (A_{STD}) of H-3 and β-D-Gal*f*O*p*NP concentration (Supplementary Fig. S4) revealed in the case of R69H-N216W† that the curve did not reach saturation, thus explaining why a much higher K_d value is measured for this mutant compared with the two others. Notably, the K_d values of the three mutants (29.6 ± 7.84, 7.0 ± 0.86 and 17.0 ± 2.10 mM for R69H-N216W†, R69H-N216W-L314N† and R69H-N216W-L352G†, respectively) do not follow the same trend as the K_M values (Table 3), an observation that underlines the fact that K_M is a composite constant describing more than just enzyme-substrate affinity.

3. Discussion

The molecular determinants of the partition between transglycosylation and hydrolysis (T/H) in *Tx*Abf has been the subject of numerous studies. These have provided a rich corpus of knowledge and mutants that display remarkable ability to synthesize α-L-Ara*f*-containing

oligosaccharides.[33,38] However, the previously identified ability of *Tx*Abf to donate D-Gal*f* moieties to various acceptor groups is less well studied.[15,35] Therefore, in the present study we have sought to address this shortcoming, using both an existing mutant and new ones created using site-saturation mutagenesis to probe structure-function relationships.

3.1. The mutant R69H-N216W fulfills the basic requirements for transgalactofuranosylation

Previous work reveals that efficient transglycosylases or TGs (*i.e.*, retaining GHs that display high transglycosylation activity compared to hydrolysis) are usually quite poor catalysts, displaying low (or moderate) catalytic efficiency (k_{cat}/K_M) when compared with their hydrolytic counterparts.[16,33,35,56] This general rule applies well to *Tx*Abf, because when it hydrolyses α-L-Ara*f*-based conjugates, such as α-L-Ara*f*O*p*NP, it exhibits good catalytic activity (e.g. for α-L-Ara*f*O*p*NP, k_{cat} = 139 s^{-1} and k_{cat}/K_M = 556 $s^{-1}.mM^{-1}$). Likewise, its ability to perform transglycosylation using α-L-Ara*f*O*p*NP is quite moderate, meaning that the T/H partition is clearly in favor of hydrolysis. On the other hand, β-D-Gal*f*O*p*NP is a poor substrate for hydrolysis (very high K_M and low k_{cat}/K_M values) but shifts the T/H partition towards transglycosylation. Surprisingly, when *Tx*Abf is mutated, introducing substitutions R69H and N216W, the hydrolytic efficiency on β-D-Gal*f*O*p*NP is unaltered, although turnover of the substrate is lowered. This is because the mutant causes concomitant lowering of K_M, a modification that has profound consequences for its properties. Somehow, mutation of residues 69 and 216 alters the way in which β-D-Gal*f*O*p*NP is bound in the active site, causing transglycosylation to increase up to 2-fold compared with *Tx*Abf. Although K_d values of the three R69H-N16W-containing mutants† are one order of magnitude higher than their corresponding K_M values, the results presented herein clearly indicate β-D-Gal*f*O*p*NP binding is enhanced. It is noteworthy that when either *Tx*Abf syntheses β-D-Gal*f*-(1,2)-α-D-Glc*p*OBn (15% yield) or when *Ct*Araf51 (another GH51 α-L-arabinofuranosidase from *Ruminiclostridium thermocellum*) is mutated to enhance transglycosylation, K_M values are also lowered.[29] Considering that reduced hydrolysis is a prerequisite to tip T/H partition in favor of transglycosylation, improving donor binding in the active site might appear to be counter intuitive. However, tighter binding does not necessarily lead to more productive binding for hydrolysis. In conclusion, we suggest that mutation of residues 69 and 216 in *Tx*Abf generates an altered binding mode for β-D-Gal*f*O*p*NP that is characterized by higher

enzyme-substrate affinity, but an unfavorable configuration for water-mediated deglycosylation of the subsequent D-galactofuranosyl-enzyme intermediate.

3.2. Mutants L314N and L352G introduce subtle donor binding features that increase regioselectivity

An outstanding result in this work is the improvement of the ability of *Tx*Abf to synthesize β-D-Gal*f*-(1,3)-α-D-Glc*p*, a motif which is found within the repeating unit of *exo*-polysaccharide from *Streptococcus thermophilus* SFi39[40] and lipopolysaccharide from *Actinobacillus pleuropneumoniae*.[39] This is consistent with recent findings that suggest that tryptophan 216 (present in R69H-N216W-containing mutants) forms a hydrophobic platform that allows alternative donor binding and thus modified regioselectivity.[35] Clearly, this new binding mode is exploited and enhanced when further mutations at positions 314 (L314N) or 352 (L352G) are introduced, since regioselectivity toward the (1,3)-linked D-galactofuranoside is increased. The STD-NMR indicates that this new regioselectivity is the result of subtle changes in β-D-Gal*f*O*p*NP binding that, in the case of the triple mutants, probably do not affect affinity, but rather positioning of the C-5 hydroxymethyl group and the anomeric proton. We postulate that this translates into less flexible positioning of the anomeric carbon and thus reduced regiodiversity during the reaction. However, another result arising from this work underlines the fact that acceptor positioning is also a determining factor for regioselectivity. The mutant enzymes described herein were able to synthesize β-D-Gal*f*-(1,4)-α-D-Glc*p*NAc, a glycomotif presents in *O*-linked oligosaccharide moieties found in *Trypanosoma cruzi* (the causative agent of Chagas' disease),[48] or even β-D-Gal*f*-(1,3)-α-D-Glc*p*NAc that presents in O antigens of *Shigella* (causes diarrhea or bacilliary dysentery).[51] Nevertheless, unlike the reaction using α-D-Glc*p*O*p*NP as the acceptor, the triple mutants apparently do not alter regioselectivity compared to the double mutant when α-D-Glc*p*NAcO*p*NP is used as acceptor (Supplementary Table S5).[52]

4. Conclusion

Focusing on the transgalactofuranosylation activity of *Tx*Abf, this work has shed new light on how this activity can be modulated using mutagenesis. However, the key finding is that the previously identified mutant R69H-N216W displays very impressive ability to perform transgalactofuranosylation and that other mutations in the vicinity of the C-5 position of the

donor D-galactofuranoside do not provide the means to surpass this performance. Nevertheless, they do provide increased regioselectivity for the synthesis of the biologically relevant β-D-Gal*f*-(1,3)-α-D-Glc*p* motif. Likewise, the synthesis of D-Gal*f*-(1,3)-α-D-Glc*p*Nac is demonstrated. Finally, regarding transglycosylase engineering, this work underlines the already established fact that transglycosylases are sluggish catalysts and that flipping from hydrolysis to transglycosylation involves altered donor binding that can also affect regioselectivity.

5. Materials and methods

5.1. Substrates and Chemicals

The substrates 4-nitrophenyl glycosides were purchased from CarboSynth. Routine experimental work was performed using chemicals purchased from Sigma-Aldrich and molecular biology reagents purchased from New England BioLabs.

5.2. Creation of site-saturation mutant libraries and site-directed mutagenesis

Four site-saturation mutagenesis gene libraries were constructed using pET24-*Tx*Abf-R69H-N216W as DNA template. Template DNA R69H-N216W was constructed through stepwise introduction of the different point mutations as previously reported [33]. SSM was performed using appropriate primer pairs designed with degenerated codons NNK (MNN for the reverse complement, synthesized by Eurogentec). To perform mutagenesis, Phusion® High-Fidelity DNA Polymerase (NEB) was employed following the supplier recommendations. Reactions were conducted following the protocol: initial denaturation step at 98 °C for 1 min, followed by 20 cycles at 98 °C for 30 s, 55 °C for 30 s, 72 °C for 4 min, and a final elongation step at 72 °C for 10 min. Afterwards, the remaining parental DNA template was digested by the addition of *Dpn*I (20 IU) at 37 °C for 3 h. 2 μL of aliquots were then used to transform chemically in XL1-Blue commercial competent cell. After growth overnight, all colonies (>200 cfus) were pooled together to extract the total plasmid DNA as library, in which 12 colonies were randomly picked and their plasmids were sequenced to verify the library diversity (GATC Biotech). Mutants were obtained with the following primers (the mutated codon is underlined and codon base change is indicated using bold letters): F26SSM: 5'-CGGCCATNNKTCGGAACATCTC-3'; L30SSM, 5'-

CATTTCTCGGAACATNNKGGGCGATGCATCTAC-3'; L314SSM, 5'-GAATCCGGGCTTCNNKTATCAGCAGAACTCG-3'; L352SSM, 5'-GCTCGTCAACGTGNNKCAATCCGTCATCC-3'.

Site-directed mutagenesis for inactivated mutants was achieved using the QuikChange II Site-Directed Mutagenesis Kit (Agilent) as previously described.[37] The enzymes for STD-NMR experiments were combined with acid/base mutation E176A that led to inactivated enzymes†. This mutation was achieved using the following primer: E176A, 5'-GGCGTCGGCAAC**GCC**AACTGGGGCTGC-3'. Mutated DNA was used to transform XL1-Blue competent cells and the success of the mutagenesis protocol was systematically verified using DNA sequencing of purified plasmid.

5.3. Mutant library screening

Competent BL21 star (DE3) cells were transformed using plasmid DNA from 4 SSM libraries and placed on selective LB agar medium (containing 50 $\mu g.mL^{-1}$of kanamycin) within a Q-tray format (37 °C, 16 h). To cover 95% of the diversity contained within the NNK libraries,[53] 94 colonies were randomly selected from each library (>200 cfus grew for each library) and dispersed in 200 µL of LB- kanamycin medium in 96-well microplates and grown at 30 °C with shaking (700 rpm) for 16 h, together with R69H-N216W as standard in the remaining two wells. Meanwhile the precultures (10 µL) were used to inoculate with 500 µL LB-kanamycine medium in 96-deepwell plates (2 mL/well) and grown overnight (30 °C, 700 rpm). Afterwards, new precultures (100 µL) were prepared and used to inoculate 900 µL aliquots of fresh LB-kanamycine in 96-deepwell plates. Cells were precultured twice to obtain homologous cell density across the different wells. Following growth for 2 h at 37 °C with shaking (700 rpm), protein expression was induced by the addition of IPTG (0.5 mM final) and incubation was pursued for 18 h at 18 °C with shaking (700 rpm). Cells were harvested by centrifugation (2,250×g, 20 min, 10 °C) and then suspended in 300 µL of 20 mM Tris-HCl buffer pH 8 containing 0.5 $mg.mL^{-1}$ lysozyme and 0.5 $IU.mL^{-1}$ DNase, and incubated at 30 °C for 1 h followed by freeze/thaw treatment (-80 °C for 30 min and then 30 °C for 30 min). Suspensions were heat-treated (75 °C, 30 min) and clarified lysates were obtained by centrifugation (2,250×g, 50 min, 4 °C). Clear lysates were transferred into fresh 96-well micro plates and stored at 4 °C for subsequent library screening. Large-scale enzyme expression and purification of *Tx*Abf and mutants were performed as previously described.[55]

The 96 cleared lysates (including the two control template duplicates), representing individual libraries, were used to simultaneously perform 96 enzyme assays using a liquid handling system (Genesis RSP 200, TECAN). Lysates were used to perform two different assays, measuring activity on β-D-Gal*f*O*p*NP (5 mM) in the absence and presence of α-D-Glc*p*O*p*NP (10 mM). The results of these assays procured activity values for v_H (hydrolysis mode) and v_T (transglycosylation) respectively. The activity ratio R (v_T/v_H) was also calculated to search for mutants whose activity increases in the presence of acceptor (α-D-Glc*p*O*p*NP). Substrate solutions were prepared in 50 mM sodium phosphate buffer at pH 7 containing 1 mg.mL^{-1} BSA. To perform reactions, the substrate solution (with or without acceptor) was first pre-heated at 45 °C for 15 min, before transfer to the reservoir of the liquid handling system. Once the reactions were initiated by adding 135 µL of substrate solution into 15 µL of lysates, the reaction microplate was immediately transferred into a spectrophotometer (Infinite M200 Microplate reader, TECAN) for continuous monitoring at 42°C (the maximum temperature of the spectrophotometer) and pH 7, thus allowing direct measure of absorbance (at 401 nm) due to free *p*NP. For reaction monitoring, absorbance was measured every 4 mins over a 2-h period, shaking the microplate for 5 s before each reading. The activity (IU.mL^{-1}) of each lysate was calculated from the linear regions of time-dependent plots.

Expression and purification of *Tx*Abf mutants

Expression and purification of *Tx*Abf and mutants were performed as previously described.[35]

5.4. Enzyme kinetics

Enzyme activities were measured using a discontinuous assay.[55] Reactions operating in hydrolysis or transglycosylation mode were performed in triplicate at 45 °C, using 5 mM β-D-Gal*f*O*p*NP (donor) and, when relevant, 10 mM different acceptors prepared in 50 mM sodium phosphate buffer at pH 7 with 1 mg/mL BSA. Prior to enzyme addition, reaction mixtures were pre-incubated at 45 °C. Once launched, reactions were conducted for 10-30 min, removing 40 µL samples at regular intervals. These samples were immediately mixed with 200 µL of 1 M Na_2CO_3 placed on ice. The release of *p*NP was monitored at 401 nm and quantified using an appropriate standard curve that was prepared using pure *p*NP. Negative controls containing all of the reactants except the enzyme were used to correct for spontaneous hydrolysis of the donor substrate. Initial reaction rates were determined from the linear regions of time-dependent plots, which correspond to less than 15% consumption of the

donor substrate. One unit (IU) of enzyme specific activity (SA) corresponds to the amount of enzyme releasing one µmol of *p*NP per minute. SA_H and SA_T designate the specific activity measured when the reaction was operated in hydrolysis and tranglycosylation mode (*i.e.*, in absence or in presence of acceptor) respectively.

The kinetic parameters K_M, k_{cat} and k_{cat}/K_M were determined by measuring enzyme SA at various substrate concentrations. In hydrolysis mode, the β-D-Gal*f*O*p*NP concentration was varied from 0.1-30 mM. In transglycosylation mode, donor kinetic parameters were measured while maintaining the acceptor α-D-Glc*p*O*p*NP concentration at 30 mM. Data describing SA as a function of substrate concentration ([S]) were fitted to the Michaelis-Menten model $SA = SA_{max}\cdot[S]/(K_M + [S])$ using a non-linear regression module embedded in SigmaPlot 11.0 software. When appropriate, a modified version of the Michaelis-Menten equation $SA = SA_{max}\cdot[S]/(K_M + [S]) + N_S\cdot[S]$ was also used for data fitting.

5.5. Time-Course ^{1}H NMR analysis

To monitor self-condensation and transglycosylation reactions, ^{1}H NMR spectra were collected using a Bruker Avance II spectrometer equipped with a TCI probe and operating at 500MHz. Reactions were performed at 45 °C with β-D-Gal*f*O*p*NP in the absence/presence of acceptors and enzyme in a total volume of 600 µL (containing 10% D_2O in 10 mM sodium phosphate buffer pH 7 with 1 $mg\cdot mL^{-1}$ BSA, v/v) in 5 mm NMR tubes. Time course NMR monitoring was achieved by performing pseudo-2D kinetics experiments based on a phase sensitive NOESY sequence with pre-saturation, with spectra being accumulated every 8.7 mins (2 × 32 scans). Using previously calculated specific activities it was possible to adjust enzyme quantities (0.69 µM for *Tx*Abf, 1.91 µM for R69H-N216W, 2.55 µM for R69H-N216W-L30G, 4.16 µM for R69H-N216W-L314G, 4.34 µM for R69H-N216W-L314N, 4.86 µM for R69H-N216W-L314S, 4.34 µM for R69H-N216W-L314Q, 3.12µM for R69H-N216W-L352G, and 3.47 µM for R69H-N216W-L352C), such that reactions occurred within a 18 h timeframe. The pH in 10% D_2O was measured using a glass pH electrode and the modified equation $pH_{10\%\ D2O} = pH_{electrode} + 0.04$.[57] The chemical shift reference was based on the HOD signal calibrated at 4.55 ppm at 45 °C. Donor (β-D-Gal*f*O*p*NP) and the apparition of products were quantified by integrating the relevant anomeric proton signals (chemical shifts are shown in Table 2), and the molarity was normalized by the initial integral of β-D-Gal*f*O*p*NP. The synthesis yields (in %) were determined by plotting product concentration against initial donor concentration. In hydrolysis mode, the mean value of the *ortho*- and

meta-protons of the linked *p*NP were regarded as the sum of the donor and self-condensation products.

5.6. Preparative-scale enzymatic synthesis of *p*ara-nitrophenyl β-D-galactofuranosyl-(1,3)-α-D-glucopyranoside

A 83 mL aliquot of 50 mM sodium phosphate buffer pH 7, containing 1 mg·mL^{-1} BSA, 5 mM β-D-Gal*f*O*p*NP (123 mg, 0.41 mmol, 1 eq.), 20 mM α-D-Glc*p*O*p*NP (498 mg, 1.65mmol, 4 eq.), and 2.3 µM R69H-N216W-L314N, was incubated for 9 h at 45 °C with stirring and then stopped by heat inactivation (10 min at 95 °C) and freezing in liquid nitrogen before lyophilization. The main transglycosylation product, β-D-Gal*f*-(1,3)-α-D-Glc*p*O*p*NP, was purified after acetylation (20 mL, in a 1/1 acetic anhydride/pyridine mixture, v/v, and in the presence of a catalytic amount of 4-dimethylaminopyridine) using automated Reveleris® flash column chromatography and a 20-100% (v/v) gradient of ethyl acetate in petroleum ether. Deacetylation performed at 0 °C in a mixture of dry methanol and dichloromethane (20 mL, 1/1, v/v) with sodium methoxide (1 M in methanol, 0.5 equiv) and acidic workup with Amberlite IR-120 (H+) resin afforded the expected β-D-Gal*f*-(1,3)-α-D-Glc*p*O*p*NP (101 mg, 0.22 mmol, 53% overall yield in 3 steps).^{1}H NMR (600 MHz, D_2O): δ 8.27-8.23 (m, 2H, $CH_{m\text{-}p\text{NP}}$), 7.31-7.28 (m, 2H, $CH_{o\text{-}p\text{NP}}$), 5.8 (d, 1H, $J_{1/2}$ = 3.7, H-1^{Glcp}), 5.32 (d, 1H, $J_{1/2}$ = 1.7, H-1^{Galf}), 4.19 (dd, 1H, $J_{2/3}$ = 3.3, H-2^{Galf}), 4.13-4.08 (m, 2H, H-4^{Galf} and H-3^{Galf}), 4.05 (t, 1H, $J_{3/4}$ = 9.6, H-3^{Glcp}), 3.90 (dd, 1H, $J_{2/3}$ = 9.6, H-2^{Glcp}), 3.86-3.83 (m, 1H, H-5^{Galf}), 3.73-3.63 (m, 5H, H-5^{Glcp}, H-6a^{Glcp}, H-6b^{Glcp}, H-6a^{Galf} and H-6b^{Galf}), 3.57 (t, 1H, $J_{4/5}$ = 9.6, H-4^{Glcp}). ^{13}C NMR (150 MHz, D_2O): δ 161.2 ($C_{i\text{-}p\text{NP}}$), 142.3 ($C_{p\text{-}p\text{NP}}$), 126.0 ($C_{m\text{-}p\text{NP}}$), 116.7 ($C_{o\text{-}p\text{NP}}$), 108.3 (C-1^{Galf}), 96.6 (C-1^{Glcp}), 82.9 (C-4^{Galf}), 81.2 (C-2^{Galf}), 79.4 (C-3^{Glcp}), 76.7 (C-3^{Galf}), 72.8 (C-5^{Glcp} 5a-C), 70.9 (C-2^{Glcp}), 70.6 (C-5^{Galf}), 67.6 (C-4^{Glcp}), 62.8 (C-6^{Galf}), 60.1 (C-6^{Glcp}). The spectral data were in agreement with those reported.[28] HR-ESI-MS *m/z* = 486.1217, calculated for $C_{18}H_{25}NO_{13}Na$ $[M+Na]^+$: 486.1218).

NMR spectra were recorded at 298 K on a Bruker Avance 600 spectrometer equipped with a TCI CryoProbe. Coupling constants (*J*) are reported in Hz, chemical shifts (δ) are given in ppm. Multiplicities are reported as follows: d= doublet, t= triplet, m= multiplet, dd= doublet of doublets. Analysis and assignments were made using 1D (^{1}H, ^{13}C and Jmod) and 2D (COSY, HSQC, HMBC and TOCSY) spectra. High-resolution mass spectrometry experiment was performed at the CRMPO (Centre régional de mesures physiques de l'Ouest, University of Rennes, France) on a Bruker MaXis 4G spectrometer, equipped with an electrospray ionization source operating in positive ion detection mode (ESI+).

5.7. STD-NMR experiments

All protein (inactivated enzyme) solutions were exchanged against buffered D_2O (99.90%, containing 20 mM Tris-HCl pD 8) using two cycles of 10-fold dilution, followed by concentration on 10 kDa cutoff Amicon® Ultra filter (Millipore). The ligand β-D-Gal*f*O*p*NP and β-D-Gal*f*-(1,3)-α-D-Glc*p*O*p*NP were prepared in D_2O (99.90%). Samples were prepared in 550 µL of D_2O (5 mm NMR tubes) containing 5 µM of protein and 250 µM of ligand with a constant molar ratio of 1:50 (protein: ligand). STD-NMR experiments were performed at 298 K with a Bruker Avance 600 spectrometer equipped with a TCI CryoProbe.[58] To achieve this, proteins were saturated on resonance at -0.5 ppm with a field strength of 56 Hz and total saturation time of 3s and 6K of repetition scans. On-resonance and off-o-resonance (+30 ppm) saturations were performed alternatively with the same experiment bet stored separately. Control experiments without proteins were also performed. I_0 was measured representing the integral of ligand proton without saturation while I_{STD} is the subtraction of integral with and without saturation. For the epitope mapping, STD effect (I_{STD}/I_0) for each proton was normalized using the largest STD effect of the sugar moiety as a reference (*i.e*. the H-3 proton of the D-Gal*f* unit was set to 100% for mutants). In order to compare the impact of mutations on STD effects, I_{STD}/I_0 of each proton were normalized by the I_{STD}/I_0 of relevant proton measured with template enzyme R69H-N216W†. All data were acquired and processed using Topspin v3.5 software.

To determine the K_d of each enzyme, a series of titration experiments were performed with different molar ratio of protein and ligand. Experiments were processed at the same condition as epitope mapping, except that the number of scan was reduced to 512 (duration 1.2 h). For each enzyme, 6 experiments were processed when maintaining enzyme at 5 µM and varying ligand from 250 µM to 1mM, meaning the ligand access ([ligand]/[protein]) ranges from 50-2000. STD amplification factor (A_{STD}) was calculated as follows: $A_{STD} = (I_{STD}/I_0) \times$ ligand excess. For each proton, A_{STD} were plotted against ligand concentrations and K_d was determined by employing Michaelis-Menten model $A_{STD} = \alpha_{STD} \cdot [ligand]/(K_d + [ligand])$ to fit the regression plot, where α_{STD} is the maximum value of A_{STD} and K_d is the dissociation constant.[59,60] K_d for each enzyme is calculated by the average K_d value of H-1, H-3, H-4, H-5, H-6a and H-6b. K_d of H-2 was excluded because its signal is too close to the water signal and its intensity was affected by the suppression of residual water signal thus resulted in a large deviation on K_d.

5.8. Molecular docking

To perform molecular docking of β-D-Gal*f*O*p*NP substrate in the active site of R69H-N216W mutant, we constructed first a 3D model of the mutant using the 3D coordinates of the triple mutant (R69H-N216W-L352M [38] and wild-type *Tx*Abf (PDB: 2VRK). Both X-ray structures were superimposed and 3D coordinates of the L352 residue from wild-type *Tx*Abf were extracted and used to replace the corresponding Methionine in the triple mutant. Regarding the β-D-Gal*f*O*p*NP substrate, charges and parameters were determined for β-D-Gal*f* and *p*NP separately. β-D-Gal*f* parameters were derived from GLYCAM 06j-1 Force Field [61], whereas *p*NP parameters were extracted from GAFF[62] with some of them made compatible with GLYCAM 06j-1 Force Field, especially for bonding two different moieties. For both β-D-Gal*f* and *p*NP, RESP charges were obtained after fitting partial charges (ESP) computed from Hartree−Fock theory (HF/6-31G*) by using Gaussian G09.[63] For β-D-Gal*f*, charges from α-L-Ara*f* were used for the common part and the charges for the additional hydroxymethyl in β-D-Gal*f* were refined using RESP charges from Gaussian calculations. β-D-Gal*f*O*p*NP was built using tleap program from AMBER18 software suite [64] and subjected to energy-minimization (250 steps of steepest descent and then 250 steps of conjugate gradient) in implicit solvent (Generalized Born, igb5) with a cut-off of 12 Å. Next, the complex was generated by superimposing the R69H-N216W mutant onto inactive *Tx*Abf in complex with α-L-Ara*f* (PDB: 2VRQ). The β-D-Gal*f* moiety of β-D-Gal*f*O*p*NP substrate was then pair fitted onto α-L-Ara*f* in 2VRQ structure. The whole complex model was finally energy-minimized (500 steps of steepest descent and then 500 steps of conjugate gradient) in explicit solvent, with standard protonation of titrable residues at physiological pH except for the acid/base (E176) which was protonated accordingly to the reaction mechanism. Superimposition and visualization were performed with PyMol.[65]

6. Acknowledgments

J. Zhao was supported by CSC (China Scholarship Council). NMR analyses were performed using facilities at MetaToul (Metabolomics & Fluxomics Facilities, Toulouse, France, www.metatoul.fr), which is part of the national infrastructure MetaboHUB (The French National infrastructure for metabolomics and fluxomics, www.metabohub.fr) and is supported by grants from the Région Midi-Pyrénées, the European Regional Development Fund, SICOVAL, IBiSa-France, CNRS, and INRA. We thank the ICEO facility dedicated to

enzyme screening and discovery, and part of the Integrated Screening Platform of Toulouse (PICT, IBiSA) for providing access to the Tecan automated liquid handling equipment.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

R.F. and M.J.O designed and supervised the study. J.Z. performed the biochemical experiment. O.S. and J.Z. performed the STD-NMR, I.A. and J.E. performed the docking experiment. All authors analyzed the data and contributed to write the manuscript

References

1 P. Peltier, R. Euzen, R. Daniellou, C. Nugier-Chauvin and V. Ferrières, *Carbohydr. Res.*, 2008, **343**, 1897–1923.

2 I. Chlubnová, D. Filipp, V. Spiwok, H. Dvoráková, R. Daniellou, C. Nugier-Chauvin, B. Králová, V. Ferrières, *Org. Biomol. Chem.*, 2010, **8**, 2092–2102.

3 L. L. Pedersen and S. J. Turco, *Cell. Mol. Life Sci.*, 2003, **60**, 259–266.

4 J. B. Houseknecht and T. L. Lowary, *Curr. Opin. Chem. Biol.*, 2001, 5, 677–682.

5 M. V. Arruda, W. Colli and B. Zingales, *Eur. J. Biochem.*, 1989, **182**, 413–421.

6 C. Marino, A. Rinflerch and R. M. De Lederkremer, *Futur. Sci. OA*, 2017, **3**, FSO199.

7 I. Chlubnová, B. Sylla, C. Nugier-Chauvin, R. Daniellou, L. Legentil, B. Kralová and V. Ferrières, *Nat. Prod. Rep.*, 2011, **28**, 937.

8 A. Trincone, *Biomolecules*, 2015, 5, 2160–2183.

9 M. R. Richards and T. L. Lowary, *ChemBioChem*, 2009, **10**, 1920–1938.

10 M. M. Nielsen and C. M. Pedersen, *Chem. Rev.*, 2018, **118**, 8285–8358.

11 P. H. Seeberger, *Acc. Chem. Res.*, 2015, **48**, 1450–1463.

12 N. L. Rose, G. C. Completo, S. J. Lin, M. McNeil, M. M. Palcic and T. L. Lowary, *J. Am. Chem. Soc.*, 2006, **128**, 6721–6729.

13 M. B. Poulin, R. Zhou and T. L. Lowary, *Org. Biomol. Chem.*, 2012, **10**, 4074–4087.

14 K. Yamatsugu, R. A. Splain and L. L. Kiessling, *J. Am. Chem. Soc.*, 2016, **138**, 9205–9211.

15 C. Rémond, R. Plantier-Royon, N. Aubrya, M. J. J. O'Donohue, *Carbohydr. Res.*, 2005, **340**, 637–644.

16 B. Bissaro, P. Monsan, R. Fauré and M. J. O'Donohue, *Biochem. J.*, 2015, **467**, 17–35.

17 N. H. Abdul Manas, R. Md. Illias and N. M. Mahadi, *Crit. Rev. Biotechnol.*, 2018, **38**, 272–293.

18 V. Lombard, H. Golaconda Ramulu, E. Drula, P. M. Coutinho and B. Henrissat, *Nucleic Acids Res.*, 2014, **42**, D490–D495.

19 E. Matsunaga, Y. Higuchi, K. Mori, N. Yairo, S. Toyota, T. Oka, K. Tashiro and K. Takegawa, *Biosci. Biotechnol. Biochem.*, 2017, **81**, 1314–1319.

20 M. Rietschel-Berst, N. H. Jentoft, P. D. Rick, C. Pletcher, F. Fang and J. E. Gander, *J. Biol. Chem.*, 1977, **252**, 3219–3226.

21 A. W. Van Bruggen-Van der Lugt, H. J. Kamphuis, G. A. De Ruiter, P. Mischnick, J. H. Van Boom and F. M. Rombouts, *J. Bacteriol.*, 1992, **174**, 6096–6102.

22 E. Matsunaga, Y. Higuchi, K. Mori, K. Tashiro and K. Takegawa, *Genome Announc.*, 2017, **5**, 3–4.

23 M. A. Cousin, S. Notermans, P. Hoogerhout and J. H. Van Boom, *J. Appl. Bacteriol.*, 1989, **66**, 311–317.

24 N. Ramli, M. Fujinaga, M. Tabuchi, K. Takegawa and S. Iwahara, *Biosci. Biotechnol. Biochem.*, 1995, **59**, 1856–1860.

25 L. C. Miletti, K. Mariño, C. Marino, W. Colli, M. J. M. Alves and R. M. d. Lederkremer, *Mol. Biochem. Parasitol.*, 2003, **127**, 85–88.

26 W. Helbert, L. Poulet, S. Drouillard, S. Mathieu, M. Loiodice, M. Couturier, V. Lombard, N. Terrapon, J. Turchetto, R. Vincentelli and B. Henrissat, *Proc. Natl. Acad. Sci.*, 2019, **116**, 201815791.

27 E. Matsunaga, Y. Higuchi, K. Mori, N. Yairo, T. Oka, S. Shinozuka, K. Tashiro, M. Izumi, S. Kuhara and K. Takegawa, *PLoS One*, 2015, **10**, e0137230.

28 I. Chlubnová, B. Králová, H. Dvořáková, P. Hošek, V. Spiwok, D. Filipp, C. Nugier-Chauvin, R. Daniellou and V. Ferrières, *Org. Biomol. Chem.*, 2014, **12**, 3080–3089.

29 Q. Pavic, A. Pillot, O. Tasseau, L. Legentil and S. Tranchimand, *Org. Biomol. Chem.*, 2019, **17**, 6799–6808.

30 C. Rémond, M. Ferchichi, N. Aubry, R. Plantier-Royon, C. Portella and M. J.

O'Donohue, *Tetrahedron Lett.*, 2002, **43**, 9653–9655.

31 M. Almendros, D. Danalev, M. François-Heude, P. Loyer, L. Legentil, C. Nugier-Chauvin, R. Daniellou and V. Ferrières, *Org. Biomol. Chem.*, 2011, **9**, 8371.

32 G. Paës, L. K. Skov, M. J. O'Donohue, C. Rémond, J. S. Kastrup, M. Gajhede and O. Mirza, *Biochemistry*, 2008, **47**, 7441–51.

33 B. Bissaro, J. Durand, A. Planas, P. Monsan, X. Biarnés, A. Planas, P. Monsan, M. J. O'Donohue, R. Fauré, *ACS Catal.*, 2015, **5**, 4598–4611.

34 F. Arab-Jaziri, B. Bissaro, M. Dion, O. Saurel, D. Harrison, F. Ferreira, A. Milon, C. Tellier, R. Fauré and M. J. O'Donohue, *N. Biotechnol.*, 2013, **30**, 536–544.

35 J. Zhao, J. Esque, I. André, M. J. O'Donohue and R. Fauré, *Prep.*

36 R. Euzen, G. Lopez, C. Nugier-Chauvin, V. Ferrières, D. Plusquellec, C. Rémond and M. O'Donohue, *European J. Org. Chem.*, 2005, **2005**, 4860–4869.

37 F. Arab-Jaziri, B. Bissaro, S. Barbe, O. Saurel, H. Débat, C. Dumon, V. Gervais, A. Milon, I. André, R. Fauré and M. J. O'Donohue, *FEBS J.*, 2012, **279**, 3598–3611.

38 J. Zhao, T. Tandrup, B. Bissaro, S. Barbe, I. André, C. Dumon, L. Lo Leggio, M. J. O'Donohue and R. Fauré, *Unpublished*.

39 M. B. Perry, L. L. MacLean and E. Vinogradov, *Biochem. Cell Biol.*, 2005, **83**, 61–69.

40 J. Lemoine, F. Chirat, J. M. Wieruszeski, G. Strecker, N. Favre and J. R. Neeser, *Appl. Environ. Microbiol.*, 1997, **63**, 3512–3518.

41 G. Stevenson, B. Neal, D. a N. Liu, M. Hobbs, N. H. Packer, M. Batley, J. W. Redmond, L. Lindquist and P. Reeves, *J. Bacteriol.*, 1994, **176**, 4144–4156.

42 F. L. Lin, E. Vinogradov, C. Deng, S. Zeller, L. Phelan, B. A. Green, K. U. Jansen and V. Pavliak, *Carbohydr. Res.*, 2014, **383**, 97–104.

43 E. Vinogradov, E. E. Egbosimba, M. B. Perry, J. S. Lam and C. W. Forsberg, *Eur. J. Biochem.*, 2001, **268**, 3566–3576.

44 F. Cafieri, E. Fattorusso, Y. Mahajnah and A. Mangoni, *Liebigs Ann. der Chemie*, 1994, **1994**, 1187–1189.

45 G. R. Pettit, J. Xu, D. E. Gingrich, M. D. Williams, D. L. Doubek, J.-C. Chapuis and J. M. Schmidt, *Chem. Commun.*, 1999, 915–916.

46 J. Yang, M. H. Nahm, C. A. Bush and J. O. Cisar, *J. Biol. Chem.*, 2011, **286**, 35813–35822.

47 M. V. Svensson, X. Zhang, E. Huttunen and G. Widmalm, *Biomacromolecules*, 2011, **12**, 2496–2501.

48 V. M. Mendoza, R. Agusti, C. Gallo-Rodriguez and R. M. de Lederkremer, *Carbohydr. Res.*, 2006, **341**, 1488–1497.

49 C. Jones, A. R. Todeschini, O. A. Agrellos, J. O. Previato and L. Mendonça-Previato, *Biochemistry*, 2004, **43**, 11889–11897.

50 L. Mendonça-Previato, L. Penha, T. C. Garcez, C. Jones and J. O. Previato, *Glycoconj. J.*, 2013, **30**, 659–666.

51 B. Liu, Y. A. Knirel, L. Feng, A. V. Perepelov, S. N. Senchenkova, Q. Wang, P. R. Reeves and L. Wang, *FEMS Microbiol. Rev.*, 2008, **32**, 627–653.

52 B. R. Clarke, O. G. Ovchinnikova, S. D. Kelly, M. L. Williamson, J. E. Butler, B. Liu, L. Wang, X. Gou, R. Follador, T. L. Lowary and C. Whitfield, *J. Biol. Chem.*, 2018, **293**, 4666–4679.

53 M. T. Reetz, D. Kahakeaw and R. Lohmer, *ChemBioChem*, 2008, **9**, 1797–1804.

54 D. Teze, J. Zhao, M. Wiemann, Z. G. Kazi, R. Lupo, M. E. Rønne, G. Carlström, J. Ø. Duus, Y.-H. Sanejouand, M. J. O'Donohue, E. N. Karlsson, R. Fauré, H. Stålbrand and B. Svensson, *ChemRxiv*, , DOI:10.26434/chemrxiv.11538708.v1.

55 B. Bissaro, O. Saurel, F. Arab-jaziri, L. Saulnier, A. Milon, M. Tenkanen, P. Monsan, M. J. O'Donohue and R. Fauré, *Biochim. Biophys. Acta - Gen. Subj.*, 2014, **1840**, 626–636.

56 D. Teze, F. Daligault, V. Ferrières, Y. H. Sanejouand and C. Tellier, *Glycobiology*, 2015, **25**, 420–427.

57 P. K. Glasoe and F. A. Long, *J. Phys. Chem.*, 1960, **64**, 188–190.

58 M. Mayer and B. Meyer, *J. Am. Chem. Soc.*, 2001, **123**, 6108–6117.

59 A. Viegas, J. Manso, F. L. Nobrega and E. J. Cabrita, *J. Chem. Educ.*, 2011, **88**, 990–994.

60 J. Angulo, P. M. Enríquez-Navas and P. M. Nieto, *Chem. - A Eur. J.*, 2010, **16**, 7803–7812.

61 K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. González-Outeiriño, C. R. Daniels, B. L. Foley and R. J. Woods, *J. Comput. Chem.*, 2008, **29**, 622–655.

62 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput.*

Chem., 2004, **25**, 1157–1174.

63 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci and G. A. Petersson, *Wallingford, CT*, 2009, **32**, 5648–5652.

64 D. A. Case, I. Y. Ben-Shalom, S. R. Brozell, D. S. Cerutti, T. E. Cheatham III, V. W. D. Cruzeiro, T. A. Darden, R. E. Duke, D. Ghoreishi and M. K. Gilson, *Univ. California, San Fr.*

65 W. . DeLano, *PyMOL Mol. Graph. Syst. Delano Sci. San Carlos.*

Supplementary information

Investigation of transgalactofuranosylation activity in an α-L-arabinofuranosidase

Jiao Zhao[a], Jérémy Esque[a], Olivier Saurel[b], Isabelle André[a], Michael J. O'Donohue[a,*] and Régis Fauré[a,*]

[a]TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France

[b]Institut de Pharmacologie et de Biologie Structurale, Université de Toulouse, CNRS, Université Paul Sabatier, Toulouse, France

Correspondance

E-mail: michael.odonohue@inra.fr (M.J.O.); regis.faure@insa-toulouse.fr (R.F.); Tel: +33 5 6155 9410

Table S1 Relative activity of lysates of mutants highlighted in each SSM library

Position	Substitution	Relative hydrolysis activity (A_H)	Relative transglycosylation activity (A_T)	Activity ratio (A_T/A_H)	Occurrence (%)
R69H-N216W (Template)		1.0 ± 0.09	3.0 ± 0.28	3.0	-
F26[a]	**M**	1.0 ± 0.19	2.0 ± 0.27	1.9	4/185
L30[a]	**G**	1.2 ± 0.13	3.9 ± 0.30	3.3	7/186
	A	1.4 ± 0.12	3.2 ± 0.31	2.2	7/186
	V	1.4 ± 0.07	1.8 ± 0.28	1.3	7/186
	S	1.2	1.5	1.3	1/186
L314[a]	**G**	2.8 ± 0.44	3.3 ± 0.48	1.2	5/92
	N	1.8 ± 0.11	2.8 ± 0.66	1.6	3/92
	S	1.7 ± 0.00	2.6 ± 0.40	1.4	2/92
	A	1.9	2.5	1.3	1/92
	H	1.9	3.0	1.6	1/92
	Q	1.7	1.7	1.0	1/92
L352[a]	**G**	4.4 ± 0.24	5.7 ± 0.43	1.3	4/92
	N	2.6 ± 0.24	4.3 ± 0.20	1.7	4/92
	V	2.8 ± 0.34	4.6 ± 0.59	1.6	3/92
	C	2.0	4.9	2.4	1/92
	F	2.9	4.7	1.6	1/92

[a] Parental clones (*i.e.*, transformants that display no mutation) represent low proportion of each SSM library, 6, 8, 9 and 7% for F26, L30, L314 and L352 SSM respectively.

Table S2. Apparent synthesis and secondary hydrolysis rates of the main regioisomer β-D-Gal*f*-(1,3)-α-D-Glc*p*O*p*NP for *Tx*Abf and mutants thereof.

Enzyme	Synthesis rate v_T ($mM.h^{-1}.\mu M^{-1}$)	2nd Hydrolysis rate v_{HII} ($mM.h^{-1}.\mu M^{-1}$)	v_S/v_{HII}[a]
wt	1.02	0.25	4.01
R69H-N216W	0.37	0.02	17.50
R69H-N216W-L314N	0.40	0.03	12.89
R69H-N216W-L352G	0.46	0.03	13.72

[a] Apparent transglycosylation (v_T, initial phase of the plot)/secondary hydrolysis (v_{HII}, after maximum yield) rates were derived from the time-course monitoring of transgalactofuranosylation product β-D-Gal*f*-(1,3)-α-D-Glc*p*O*p*NP yield. This ratio depicts the ability of the enzymes to perform transgalactofuranosylation towards the synthesis of the main regioisomer β-D-(1,3)-linked product while considering their secondary hydrolysis capabilities.

Table S3. Maximal self-condensation yield in hydrolysis mode[a]

***Tx*Abf**	**Self-condensation yield Y_{max} (%)**					**Relative SA_H[c]**
	(1,2)[b]	**(1,3)[b]**	**(1,5)[b]**	**(1,6)[b]**	**Total**	
wt	18	2	-	-	24	13.2
R69H-N216W	15	22	3	3	47	1.0
R69H-N216W-L30G	-	13	-	4	22	0.9
R69H-N216W-L314G	11	13	2	1	32	2.0
R69H-N216W-L314N	18	12	3	3	41	1.6
R69H-N216W-L314S	8	18	-	3	34	2.0
R69H-N216W-L314Q	11	15	-	4	33	1.6
R69H-N216W-L352G	14	10	5	2	34	2.2
R69H-N216W-L352C	10	13	5	2	34	1.6

[a] Self-condensation reactions were carried out using 5 mM β-D-Gal*f*O*p*NP at 45°C and pH 7 in buffered 10% D_2O.

[b] Self-condensation products were *p*-nitrophenyl di-β-D-galactofuranosides.

ND: the yield is too low to be detected.

[c] Relative SA_H was determined as an activity ratio in hydrolysis mode based on the SA_H of template R69H-N216W.

Table S4. Influence of the substrate concentrations and ratio on transglycosylation yields for *Tx*Abf and mutants thereof.[a]

[Donor] (mM)	[acceptor] (mM)	Enzyme	Transgalactofuranosylation yield (%)				
			(1,2)	(1,3)	(1,m)	Overall	Regioselectivity (%)[b]
5	10	wt	12	20	1	33	61
		R69H-N216W	6	51	8	65	78
		R69H-N216W-L314N	2	60	2	64	93
		R69H-N216W-L352G	2	59	5	66	89
15	30	R69H-N216W	7	60	8	73	81
		R69H-N216W-L314N	3	70	5	77	90
		R69H-N216W-L352G	3	68	8	77	87
5	20	R69H-N216W	8	65	9	82	79
		R69H-N216W-L314N	3	74	5	81	91
		R69H-N216W-L352G	3	73	7	84	87

[a] All assays were carried out using 5 mM β-D-Gal*f*O*p*NP as donor and 10 mM α-D-Glc*p*O*p*NP as acceptor at 45 °C and pH 7 in buffered 10% D_2O.

[b] The proportion of β-D-Gal*f*-(1,3)-α-D-GlcO*p*NP was calculated against the overall transglycosylation yield, indicating the purity of the main product.

Table S5. Transglycosylation profile using α-D-Glc*p*NAcO*p*NP as acceptor

	Transglycosylation yield (%)	
Enzyme	5.07 ppm[48]	4.89 ppm[52]
R69H-N216W	5	20
R69H-N216W-L314N	5	9
R69H-N216W-L352G	4	17

All assays were carried out using 5 mM β-D-Gal*f*O*p*NP as donor and 10 mM α-D-Glc*p*NAcO*p*NP as acceptor at 45 °C and pH 7 in buffered 10% D_2O. Chemical shifts were measured at 45 °C. However, linkage attributions were undetermined.

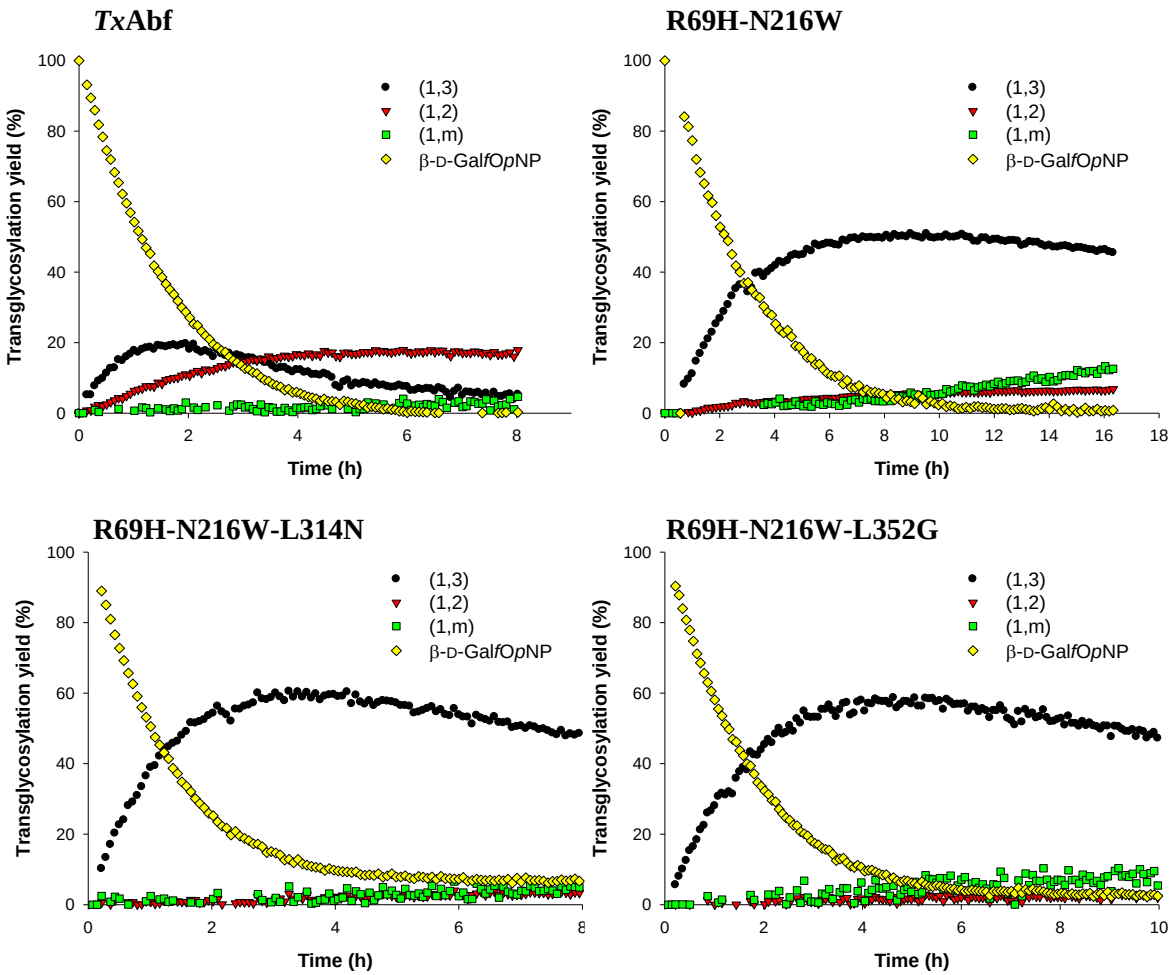


Fig. S1 Monitoring of β-D-Gal*f*O*p*NP consumption and the apparition of the different transgalactofuranosylation products as a function of time in reactions catalyzed by *Tx*Abf and mutants thereof. All assays were carried out using 5 mM β-D-Gal*f*O*p*NP as donor and 10 mM α-D-Glc*p*O*p*NP as acceptor at 45 °C and pH 7 in buffered 10% D_2O.

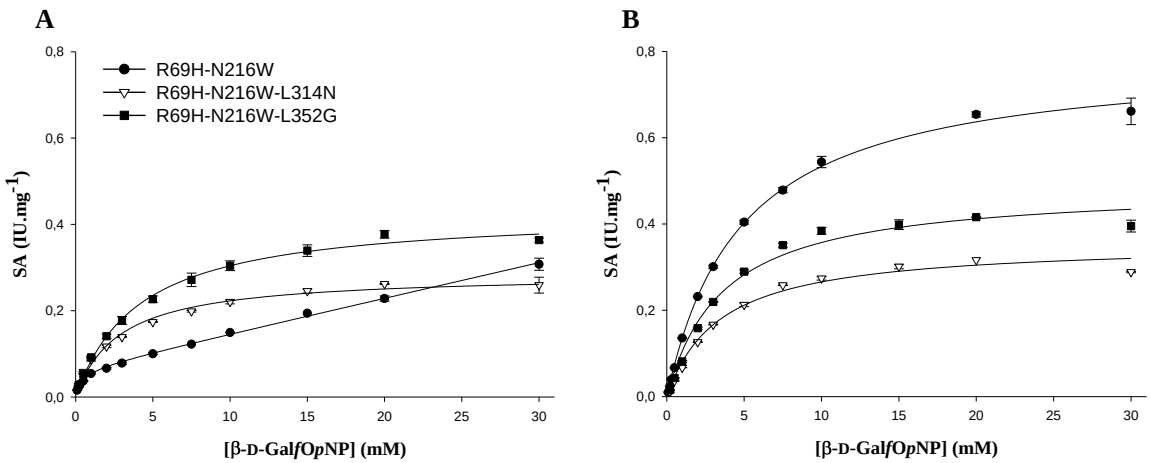


Fig. S2 Specific activity (SA) as a function of β-D-Gal*f*O*p*NP concentration for R69H-N216W (●), R69H-N216W-L314N (▽) and R69H-N216W-L352G (■) in (A) hydrolysis mode and (B) transglycosylation mode (acceptor α-D-Glc*p*O*p*NP was fixed at 30 mM).

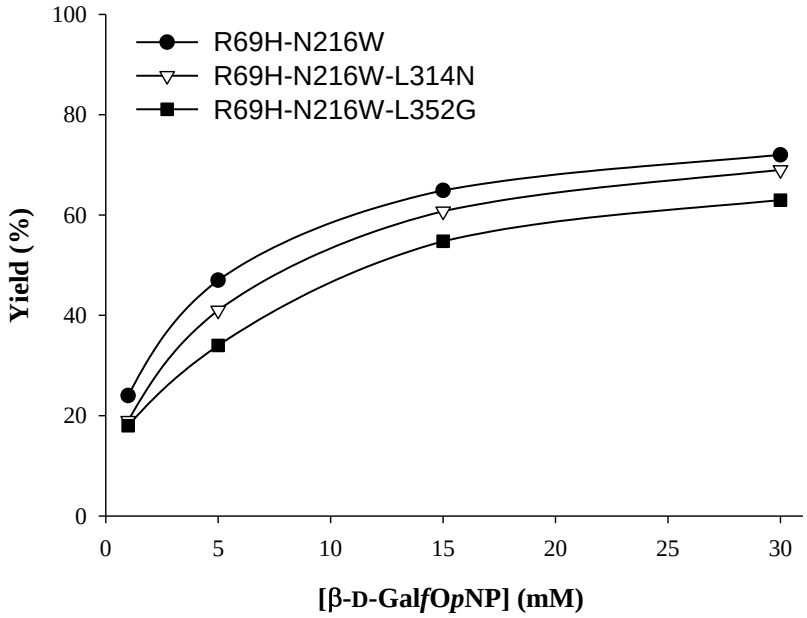


Fig. S3 Overall self-condensation yield as function of β-D-Gal*f*O*p*NP concentration in the reaction catalyzed by R69H-N216W, R69H-N216W-L314N and R69H-N216W-L352G. All assays were carried out at 45 °C and pH 7 in buffered 10% D_2O.

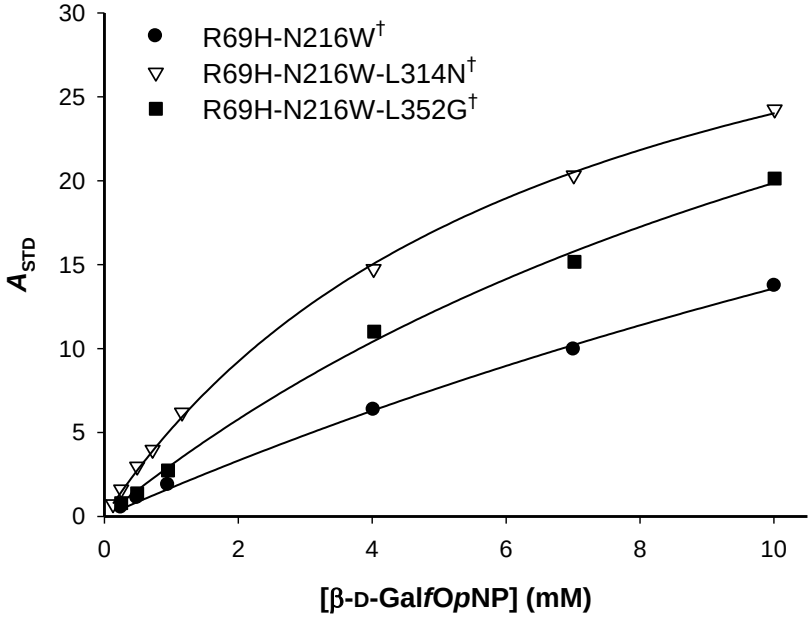


Fig. S4 STD amplification factor (A_{STD}), based on the most intense signal (H-3 of D-Gal*f* moiety), as a function of β-D-Gal*f*O*p*NP ligand concentration for R69H-N216W, R69H-N216W-L314N and R69H-N216W-L352G.

Chapter V.

Rational enzyme design without structural knowledge: a sequence-based approach for efficient generation of glycosylation catalysts

One of our ultimate goals is to acquire sufficient understanding of how the T/H partition is established in *Tx*Abf in order to then transfer this knowledge to other related rGHs. In this respect, our quest is shared by a considerable community of European researchers. Through combined efforts and many years of research, it is now possible to lay down a certain number of guidelines. However, engineering transglycosylases is still a long process that is more adapted to the interests of enzymologists than those of carbohydrate chemists who simply seek suitable catalysts to access target glycoconjugates. Therefore, without slackening the pace of in-depth research in the field of transglycosylases, there is a real need for faster, less labour intensive engineering approaches to deliver tailored-for-purpose glycosynthetic tools within a reasonable time delay.

Building on the pioneering work performed by Teze et al (2015) and good collaboration between different research groups, this publication describes how a refined conservation-based engineering approach can be used to quickly generate a small library of enzyme candidates for glycosynthesis, starting from almost any GH-A clan member. To exemplify this, the method was performed on five different GH-A members. The results show that in each case the straightforward sequence-based approach, which doesn't require prior structural knowledge or intensive screening process, provides a means to quickly generate a series of mutants that display increased transglycosylation. The work also demonstrates that mutational hotspots in one enzyme are often transposable to other related enzymes. Clearly, the results presented hereafter will be useful to both glycochemists seeking fast access to new glycosynthetic and to enzymologists wishing to initiate more in-depth studies on new GH candidate enzymes.

Rational enzyme design without structural knowledge: a sequence-based approach for efficient generation of glycosylation catalysts

David Teze[1,*], Jiao Zhao[2], Mathias Wiemann[3], Kazi Zubaida Gulshan Ara[4], Rossana Lupo[1], Mette Errebo Rønne[1], Göran Carlström[5], Jens Duus[6], Yves-Henri Sanejouand[7], Michael J. O'Donohue[2], Eva Nordberg-Karlsson[4], Régis Fauré[2], Henrik Stålbrand[3], Birte Svensson[1,*]

[1]Department of Biotechnology and Biomedicine. Technical University of Denmark, Kongens Lyngby, Denmark.

[2]TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France.

[3]Department of Biochemistry and Structural Biology. Lund University, Lund, Sweden.

[4]Biotechnology, Department of Chemistry. Lund University, Lund, Sweden.

[5]Centre for Analysis and Synthesis, Department of Chemistry. Lund University, Lund, Sweden

[6]Department of Chemistry. Technical University of Denmark, Kongens Lyngby, Denmark.

[7]Université de Nantes, CNRS, UFIP, France. Correspondence should be addressed to D.T. (datez@dtu.dk) or B.S. (bis@dtu.dk).

Abstract

Oligosaccharides are ubiquitous in Nature, being involved in most biological processes. Nevertheless, due to the costly and challenging nature of glycochemistry, glycobiology is dogged by the relative scarcity of synthetic, defined oligosaccharides. Enzyme-catalysed glycosylation using glycoside hydrolases is an alternative route to obtain oligosaccharides, but is hampered by the innate hydrolytic activity of the enzymes. Protein engineering methods, commonly employed to diminish hydrolysis in glycoside hydrolases, usually require prior structural knowledge of the target enzyme and the use of powerful computing methods, and/or relies on extensive *in vitro* screening methodologies. Here we describe a straightforward strategy that involves rapid *in silico* analysis of protein sequences. The method pinpoints a small number (<10) of candidates for mutational analysis aimed at diminishing hydrolysis and thus tipping the reaction balance toward transglycosylation. Requiring no other significant prior knowledge of the target enzyme, the results reveal that the method is quite generic, allowing the improvement of glycoside hydrolases that act on

different α-/β-pyranosides or furanosides. Moreover, data presented suggest that mutational hotspots that are validated in one enzyme can be transposed to other closely related enzymes without the need for further analysis.

Keywords: biocatalysis, sequence conservation, rational design, oligosaccharide synthesis, transglycosylation, glycoside hydrolase

1. Introduction

Glycosides are ubiquitous and abundant in Nature, being essential for a variety of biological interactions and processes. Nevertheless, progress in glycobiology is hampered by the lack of synthetic carbohydrates, an issue related to their complexity. Carbohydrates are composed of polyhydroxylated units that exist in different forms (e.g. pyranoside or furanoside conformations), interlinked in a variety of ways, with the anomeric centres displaying either α- or β-anomeric configurations[1]. Faced with this high degree of complexity organic chemistry has developed numerous glycosylation methodologies[2,3]. These generally involve several synthetic steps, including protection-deprotection cycles, are characterised by relatively poor overall yields, and generate a considerable amount of waste products. This is in stark contrast to polynucleotides and polypeptides, both of which are accessible *via* automated chemical synthesis processes and through *in vivo* biological synthesis. Unfortunately, unlike these biopolymers, carbohydrates cannot be obtained using straightforward, generic technologies[4] amenable to automation.

Enzyme-catalysed glycosylation offers an alternative to chemical methods. The natural choice for this are glycosyltransferases (GTs) that are well-represented in a variety of families in the CAZy database (http://www.cazy.org/)[5], with each family potentially harbouring numerous specificities[6]. Nevertheless, GTs have proven to be rather difficult to handle *in vitro* and often require expensive nucleotide-glycosides donors[7]. Therefore, glycoside hydrolases (GH) offer an alternative for glycosynthesis. These enzymes are particularly abundant and their diversity in terms of bond breaking ability is matched only by carbohydrate complexity itself. GHs catalyse hydrolysis using a variety of mechanisms, but the majority operate in two-step displacement of the anomeric configuration, thus yielding a product whose anomeric configuration is identical to that of the substrate[8,9]. GHs operating by such a mechanism are termed “retaining” GHs and represent 68% of all GHs (727 714 classified modules) in the CAZy database, grouped in 82 out of its 160 GH families (as of December 23th, 2019). Allowing for rare exceptions[10], all other GHs are

inverting. A result of the double displacement mechanism is that retaining GHs possess the intrinsic potential to catalyse transglycosylation, thus to synthesize glycosidic bonds. Even though most retaining GHs have strong hydrolytic activity and weak, often undetectable transglycosylation activity, some display significant levels of transglycosylation. This reaction is under kinetic control and modulated by a number of factors related to reaction conditions[11], including acceptor concentration, water activity, substrate activation, temperature, pH, and enzyme properties. Accordingly, engineering of GHs has proven to be a potent way to obtain evolved transglycosylases[12].

The most generic GH engineering approach described to date is the so-called glycosynthase strategy[13–16]. This involves the creation of a crippled enzyme, in which a catalytic carboxylate (the nucleophile) is replaced by a catalytically impotent moiety. The resulting mutant enzyme is fed with a strongly activated substrate that mimics the reaction intermediate (e.g. an α-glycosyl fluoride for a β-active wild-type (WT) enzyme), and turns the enzyme into an inverting glycosynthase. Albeit powerful, this method relies on the availability of a suitably reactive, but sufficiently stable substrate. Moreover, glycosynthases are intrinsically impotent biocatalysts that display extremely low activity, thus requiring large quantities of enzyme[13–18].

An alternative strategy to convert GHs into efficient transglycosylases is to increase the transglycosylation/hydrolysis (T/H) ratio while conserving the retaining mechanism. To achieve this, a considerable number of studies have employed rational or random protein engineering methodologies. However, such approaches require either in-depth structural and biochemical knowledge and, in some cases, use of sophisticated computational methods (rational design); or the creation of large libraries, introducing location-agnostic modifications, coupled to a powerful phenotypic screen (random mutagenesis)[19]. Nevertheless, by simply targeting a small number of conserved active-site residues in several retaining GHs[20–22], we previously demonstrated that transglycosylation capability can be improved without extensive screening using structural information coupled with sequence conservation analysis. Likewise, others have successfully applied our method to related enzymes, leading to significant improvements in transglycosylation capability[23,24].

Herein we demonstrate how a freshly refined sequenced-based approach can be used to improve the T/H ratio (i.e. enhance the transglycosylation capability) in a variety of inherently hydrolytic GHs. This strategy involves mutation of an enzyme's most conserved residues, replacing these by structural analogues (e.g., Tyr into Phe, Asp into Asn), while

avoiding mutation of the catalytic residues (when known) and residues that are distant from the subsite -1[25] (if known). Moreover, any conserved glycines and prolines are excluded[21]. Using this strategy, we demonstrate successful application to GHs from families 2, 10, 20, 29 and 51. This approach is: i) fast, requiring the generation and analysis of ≤ 10 variants per enzyme; ii) generic, being applicable to a variety of glycosidic bond-forming reactions; iii) procures highly efficient transglycosylases; and iv) allows for mutation transfer.

2. Multiple sequence alignment and residue conservation

The method described herein relies on identifying conserved residues. While the concept of amino acid conservation is central to molecular evolution theory, to date there is no agreed, precise definition of what it signifies. To circumvent this obstacle, we have devised a methodology that pinpoints residues that are significantly more conserved compared to others within a given sequence. For this approach to be successful, a ranking method is required that is robust even in extreme cases where the target sequences are either highly diverse, meaning that the overall stringency of conservation is low, or alternatively where they are highly homogenous. The three-step methodology begins with the collection of a large number of sequences, which are then clustered to reduce redundancy[26,27]. After clustering, iterative multiple sequence alignments (MSA) are performed to select sequences that are above an increasing identity threshold. At each step, sequences displaying more than 10% identity, the residues are ranked by decreasing conservation, and their conservation is plotted against the logarithm of their rank. When a significant conservation drop is observed after rank 10–15 (Fig. 1), the sequence pool is saved. The corresponding MSA is then analysed to identify the residues hereafter designated as "conserved". Iterative MSA and conservation analysis are performed using clustal-omega[28] and an in-house script.

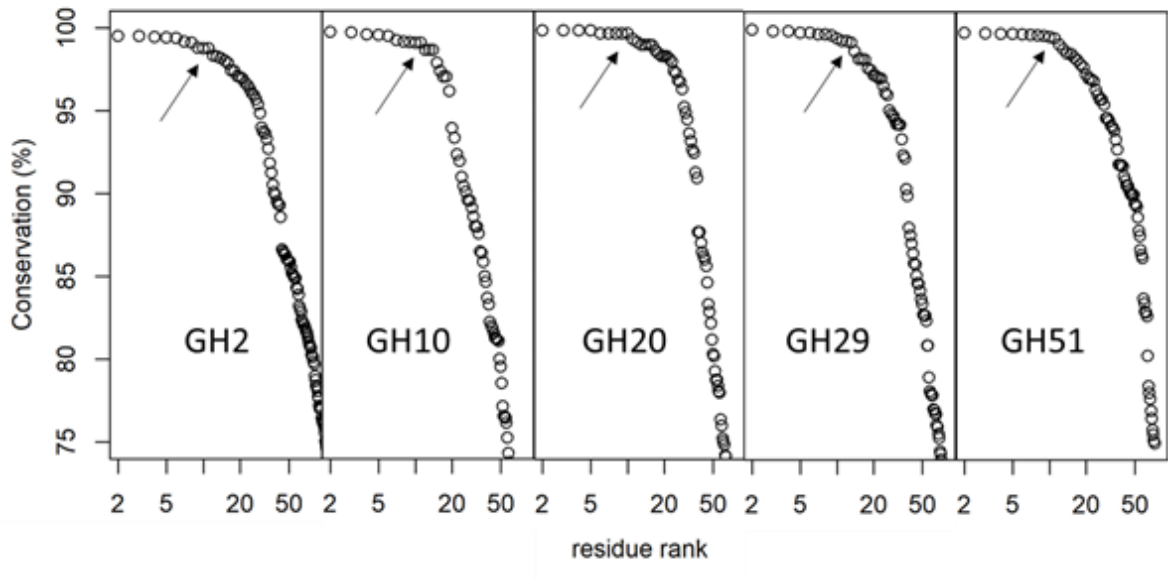


Fig. 1 | Residue conservation analysis. Residues in the MSA are ranked according to decreasing conservation, and their conservation is plotted as function of the logarithm of their rank. From left to right, results obtained from the analysis of the final MSA on GH families 2, 10, 20, 29 and 51 respectively. The arrows indicate the last rank considered as "significantly more conserved" than the rest of the residues in one sequence.

To test the methodology, GHs from families 2, 10, 20, 29 and 51 were selected. These cover a wide range of glycosidic bond features, each representing a distinctive synthetic challenge. The outcomes of the sequence analysis are summarized in Table 1. Comparing GH10 and GH51 shows that similar sequence pool size and identical stringency criteria generate different heterogeneity levels (21 vs 35% average identity – termed "mean ID%"), while the case of GH2 and GH20 reveal that similar heterogeneity can be obtained from a variable number of selected sequences (9301 vs 585 sequences retained from 20 000 retrieved). Nevertheless, in all cases the methodology yields 6–12 candidates for mutation and experimental evaluation.

GH family	Enzyme target	Species	Retrieved sequences	ID%[a] threshold	Sequences kept	Mean ID%[a,b]	Conserved residues	Candidates evaluated[c]
GH2	*Cf*Man2A	*C. fimi*	20 000	12	9301	22.6	14	12
GH10	Xyn10A	*R. marinus*	9 288	15	4431	21.1	11	9
GH20	*Bb*HI	*B. bifidum*	20 000	10	585	22.5	9	6
GH29	AlfB	*L. casei*	15 328	20	1194	24.0	13	7
GH51	*Tx*Abf	*T. xylanilyticus*	11 553	15	3272	34.6	11	8

Table. 1 | MSA and residue conservation analysis. [a]ID%: percentage of identity between sequences. [b]Average identity with the reference sequence. [c]Not evaluated Gly, Pro and identified catalytic residues, and His are mutated in Asn or Phe.

3. GH2, β-mannosidase *Cf*Man2A

Compared to the synthesis of α-D-mannosidic (or even β-D-glucosidic and β-D-galactosidic) bonds, the synthesis of β-D-mannosyl-containing compounds is complicated by the axial 2-OH of D-mannosyl moieties. In the case of β-mannosidase-mediated synthesis, this constraint leads to a different conformational itinerary ($^{1}S_{5} \rightarrow [B_{2,5}]^{\ddagger} \rightarrow {}^{O}S_{2}$) compared to that adopted by most other β-pyranosidases-catalysed reactions ($^{1}S_{3} \rightarrow [^{4}H_{3}]^{\ddagger} \rightarrow {}^{4}C_{1}$)[29]. Overall, the challenge of synthesizing β-mannosides[30] has led to interest in enzymatic synthesis[31–33] and several GH2 mannosidases have been successfully tested[34–36]. Furthermore, a glycosynthase variant of the *Cellulomonas fimi* GH2 β-mannosidase (*Cf*Man2A)[37] proved to be a proficient synthetic catalyst when fed with donor α-D-mannosyl fluoride and various acceptors[31,38]. Herein, we investigate whether *Cf*Man2A-mediated transmannosylation can be enhanced without using the drastic glycosynthase strategy.

*Cf*Man2A variant	Yield (%)	Enzyme (**μg·mL^{-1}**, nM)
WT	**18**	**9.8** (104)
R360K	**37**	**30.9** (330)
W362H	**18**	**4.7** (50)
N428T	**43**	**34.3** (360)
S474A	**32**	**33.2** (350)
H496Q	**15**	**5.3** (56)
W612H	**37**	**23.9** (253)
Q613N	**16**	**15.3** (162)
W623H	**17**	**2.8** (30)

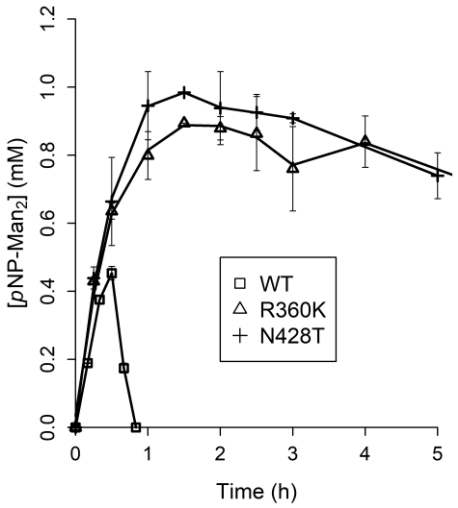


Fig. 2 | GH2 engineering. Left, maximum yields of *p*NP-Man$_2$ synthetized by *Cf*Man2A forms. Protein concentrations used to obtain the yields within 5 h are indicated in μg·mL^{-1} and (nM). **Right,** HPLC monitoring of the formation of *p*NP-Man$_2$ by *Cf*Man2A-WT and two of its best mutants.

The transglycosylation ability of wild-type (WT) *Cf*Man2A and mutated variants (collectively "*Cf*Man2A forms") was evaluated with 5 mM *p*-nitrophenyl- -D-mannopyranoside (β-D-Man*p*O*p*NP) as substrate, acting as both donor and acceptor. The formation of β-D-Man*p*-(1→3)-β-D-Man*p*O*p*NP and β-D-Man*p*-(1→4)-β-D-Man*p*O*p*NP (both are referred to as *p*NP-Man$_2$) from 5 mM β-D-Man*p*O*p*NP was monitored using HPLC (Figs. 2 and S1). Nuclear magnetic resonance spectroscopy (NMR) and mass-spectrometry (MS) (Fig. S1, Tables S1 and S2) were used to determine the chemical structure of the products. Four variants (R360K, N428T, S474A and W612H) showed markedly improved yields for *p*NP-Man$_2$ synthesis (32–43%) compared to *Cf*Man2A-WT (18%), while W169H, D170N, D386N and H496F were discarded due to low activities (<5 nkat·mg^{-1}).

It is noteworthy that unlike mutants generated using the glycosynthase approach, the four *Cf*Man2A mutants described herein were catalytically-active and thus could be used at loadings that were only moderately greater (2 to 3.5-fold, Fig. 2) than that used for *Cf*Man2A-WT. Moreover, compared to WT enzyme, monitoring revealed that the mutants performed less secondary product hydrolysis and thus ensured high product concentrations even over prolonged reaction times (Fig. 2). This result is significant because secondary hydrolysis often impedes enzyme-mediated synthesis of D-mannoside conjugates[39,40].

4. GH10, β-*endo*-xylanase *Rm*Xyn10A_CM

Compared with most hexopyranoses, the pentopyranose configuration of the D-xylosyl unit confers enhanced flexibility and allows two distinct conformational itineraries that are compatible with GH hydrolysis: $^{1}S_{3} \rightarrow [^{4}H_{3}]^{\ddagger} \rightarrow {}^{4}C_{1}$ and $^{2}S_{O} \rightarrow [^{2,5}B]^{\ddagger} \rightarrow {}^{5}S_{1}$[29,41]. Mechanistically interesting as it may be, the real challenge posed by this enzyme is related to its *endo*-activity. This implies the use of a substrate that does not possess an exceptionally good leaving group, and thus does not provide kinetic control of the reaction, and also the necessity to monitor multiple products in order to assess transglycosylation efficiency. This is because the reaction allows the transfer of oligosaccharide moieties onto an acceptor in a single catalytic step and provides the means to synthesize higher oligosaccharides displaying degrees of polymerisation (DP) > 10. *Endo*-1,4-β-xylanase is the predominant enzyme class in GH10 family and the subject of numerous studies[42–45]. Both negative and positive subsites[25] in GH10 xylanases have been probed to understand their influence on substrate binding and catalysis[46–49] (including *Rm*Xyn10A–CM, the catalytic module of the *endo*-xylanase from *Rhodothermus marinus* studied here)[50]. However, very few studies have focused on the evaluation of the transglycosylation reaction, these being limited to mutational studies of the aglycone subsites in a few GH10 xylanases[46,49].

***Rm*Xyn10A–CM variant**	**Maximal X_8 area (nC·min)**	**Enzyme (μg·mL^{-1}, nM)**
WT	**4.9**	**10.2** (21)
H69F	**8.4**	**43.6** (90)
H69N	**11.6**	**37.7** (78)
W73H	**8.3**	**41.3** (86)
N118T	**13.4**	**39.5** (82)
N169T	**5.5**	**10.5** (22)
H204F	**8.9**	**35.6** (74)
H204N	**6.9**	**36.3** (75)
D234N	**6.6**	**39.1** (81)
W284H	**10.7**	**38.5** (80)

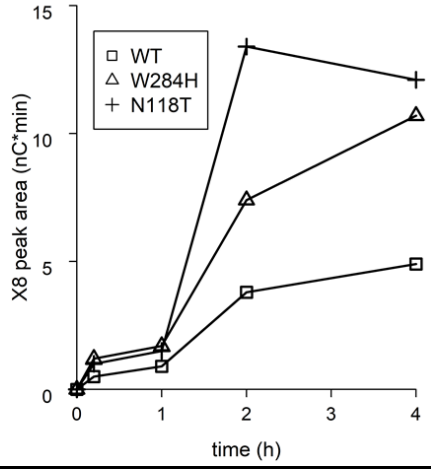


Fig. 3 | GH10 engineering. Left, Xylooctaose (X_8) synthesised by *Rm*Xyn10A–CM forms. Protein concentrations used to obtain transglycosylation within 4 h are indicated in μg·mL^{-1} and (nM). **Right,** monitoring of X_8 synthesis by *Rm*Xyn10A–CM-WT and two of its best mutants using HPAEC-PAD. Residue numbering is based on the single catalytic module.

To characterise transglycosylation activity, xylotetraose (X_4) was used as both donor and acceptor. All *Rm*Xyn10A–CM forms were able to synthesise xylo-oligosaccharides (XOS) larger than X_4, predominantly accumulating X_8 during the 4 h time frame of the experiment (Figs. 3, S2 and S3). Notably, mutants H69N, N118T and W284H synthesized XOS with higher DP during the disproportionation of 20 mM X_4 (Fig. S2) than the WT enzyme. N118 can potentially hydrogen bond with the xylose residue in the −1 subsite of *Rm*Xyn10A (Fig. S5). The high transglycosylation activity of N118T indicates that the mutation alters the ability of the neighbouring catalytic E119 to extract a proton from a water molecule. Further analysis revealed that using X_4 as substrate, *Rm*Xyn10A–CM variants synthesized XOS of DP 6−11 (Fig. S2, S3, Table S3). Importantly, compared to the WT enzyme, all mutants display greater ability to form XOS exhibiting DP ≥ 8 (Table S3), these being the result of multiple transglycosylation events. These findings reveal that the conserved sequence approach led to *Rm*Xyn10A mutants with significantly enhanced transglycosylation abilities compared to the wild-type enzyme.

5. GH20, β-hexosaminidase *Bb*HI

During catalysis, most retaining GHs generate a glycosyl-enzyme intermediate whose anomeric configuration is opposite to that of both the substrate (donor) and the products. However, GH families 18, 20, 25, 56, 84, 85 and 123 use a substrate-assisted mechanism, in which an equatorial *N*-acetyl in position C-2 of the donor acts as the nucleophile[51]. This leads to the formation of a non-covalent oxazoline- or oxazolinium ion-enzyme intermediate[52] (oxazolinium ion in the case of GH20). Importantly, this type of mechanism is predominant in GH-catalyzed reactions involving β-D-GlcNAc or β-D-GalNAc, sugars that are highly prevalent in biological systems. Hence, to assess if the methodology presented here is only relevant when a glycosyl-enzyme intermediate is formed, a GH20 enzyme was targeted.

Previously, GH20 hexosaminidases have been extensively used for transglycosylation[53,54]. Therefore, herein we focused on *Bb*HI from *Bifidobacterium bifidum*, sp *infantis*. This enzyme has the ability to catalyse the synthesis of Lacto-*N*-Triose or LNT2 (β-D-Glc*p*NAc-(1→3)-β-D-Gal*p*-(1→4)-D-Glc), a widely sought-after glycomotif that is present in human milk oligosaccharides[7,55]. Six mutants and the *Bb*HI-WT were analysed for their ability to synthesize LNT2 from 10 mM β-D-Glc*p*NAcO*p*NP and 40 mM lactose (Fig. 4).

The mutant H603F displayed the highest T/H ratio, but also a drastically reduced activity. With the highest enzyme loading (15 µM), the reaction was incomplete after 12 h, thus further characterization of *Bb*HI-H603F was abandoned.

***Bb*HI variant**	**Yield (%)**	**Enzyme (µg·mL^{-1},** µM)
WT	**16**	**93** (0.53)
R577K	**36**	**229** (1.3)
H603F	**–**	**–**
D606N	**9**	**229** (1.3)
W801H	**17**	**810** (4.6)
W882H	**66**	**862** (4.9)
D884N	**12**	**582** (3.3)

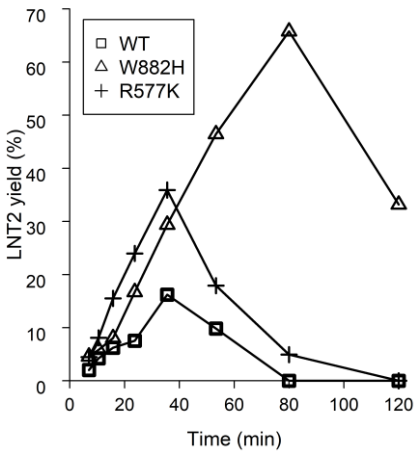


Fig. 4 | GH20 engineering. Left, maximum yields of LNT2 synthetized by *Bb*HI forms. Protein concentrations used to obtain the yields within 2 h are indicated in µg·mL^{-1} and µM (in brackets). **Right,** HPAEC-PAD monitoring of LNT2 production by *Bb*HI-WT and its two best mutants.

The variants R577K and W882H displayed high yields of LNT2 (66% for W882H), although the latter was also poorly active. Nevertheless, the mutant R577K represents an interesting compromise, since compared to *Bb*HI-WT, it generated a markedly higher yield (36 vs 16%) of LNT2 at a reasonable enzyme loading (1.3 µM). Importantly, these results demonstrate that our strategy is not restricted to retaining GHs catalysing transglycosylation through a covalent glycosyl-enzyme intermediate. Conversely, the application of our approach to another GH20, from *Ewingella americana*, failed to generate mutants able to perform transglycosylation using β-D-Glc*p*NAcO*p*NP as donor and lactose as an acceptor (data not shown). Although this might indicate that the strategy is not fully generic, we believe that this was unsuccessful due to the fact that the WT GH20 from *Ewingella americana* displays no detectable ability to transglycosylate lactose (no other acceptors tested). This emphasises that the existence of innate, albeit weak, transglycosylation activity is possibly a prerequisite for success when using our approach.

6. GH29, α-L-fucosidases AlfB and AlfC

L-Fucose is the most common L-sugar in animals[56], invariably connected through an axial glycosidic linkage. Accordingly, a conformational itinerary $^{1}C_4 \rightarrow [^{3}H_4]^{\ddagger} \rightarrow {}^{3}S_1$ is followed in enzymatic hydrolysis[51]. A few studies have been undertaken to improve GH29-mediated transfucosylation[57,58], including one in which directed evolution was used[59], albeit with varying success. For this study, we focused on AlfB from *Lactobacillus casei*. This enzyme is reported to synthesize α-L-Fuc*p*-(1→3)-D-GlcNAc[60,61], the α-L-Fuc*p*-(1→3) motif being particularly common with 8 out of 13 known human fucosyltransferases being 3-fucosyltransferases[56]. We assessed AlfB-WT and seven of its mutants for their ability to synthesise α-L-Fuc*p*-(1→3)-D-GlcNAc from 20 mM 2-chloro-4-nitrophenyl α-L-fucopyranoside (α-L-Fuc*p*OCNP) and 20 mM D-GlcNAc (Figs. 5 and S7).

Variant	Yield (%)	Enzyme (μg·mL^{-1} , μM)
AlfB-WT	**18**	**65** (1.4)
H21F	–	–
Y45F	**21**	**20** (0.4)
H80F	**57**	**250** (5.2)
H81F	**1.3**	**465** (10)
Y124F	**15**	**740** (16)
D129N	**22**	**117** (2.5)
W130H	**31**	**750** (16)
AlfC-WT	**17**	**150** (1.0)
H87F	**26**	**4.500** (30)
W137H	**22**	**500** (3.3)

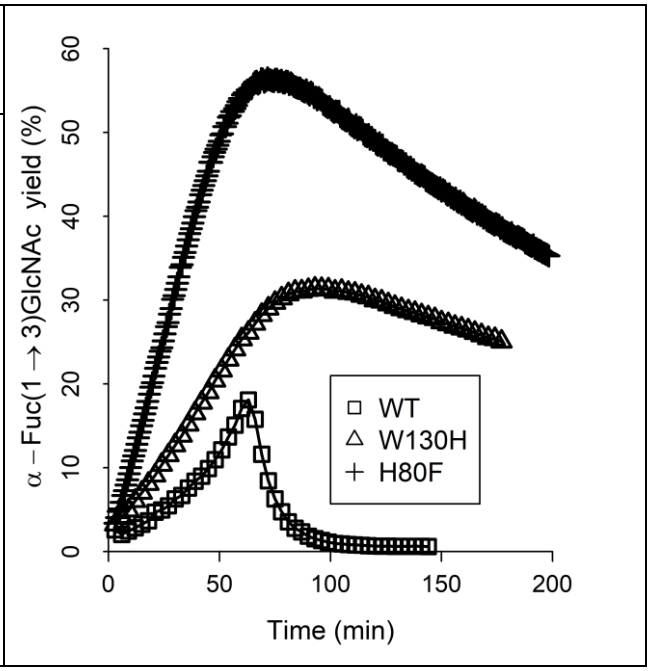


Fig. 5 | GH29 engineering. Left, maximum yields of α-L-Fuc*p*-(1→3)-D-GlcNAc and α-L-Fuc*p*-(1→6)-D-GlcNAc synthetized by AlfB and AlfC forms, respectively. Protein concentrations used to obtain maximum yields within 3 h given in μg·mL^{-1} and (μM). **Right,** NMR monitoring of the product formation with AlfB-WT and its two best mutants.

The mutant H21F presented barely detectable activity and was eliminated. Out of the six remaining variants, two presented markedly higher yields than AlfB-WT. The H80F mutation procured very high transfucosylation yield (57%), while displaying high

transglycosylation rate (> 11 mM of α-L-Fuc*p*-(1→3)-D-GlcNAc were obtained within 1 h with 5 µM enzyme) and complete regioselectivity. It is noteworthy that mutation of its neighbouring residue (H81F) almost eliminated transglycosylation. Therefore, these two adjacent, conserved histidines possess opposite but determinant roles in regulating the T/H ratio. The presence of α-L-Fuc*p*OCNP prevents secondary hydrolysis of the product. Therefore, in reactions catalysed by AlfB-WT the transglycosylation product is rapidly hydrolysed after complete consumption of the donor substrate (Figs. S6 and S7). However, in the case of mutants, particularly for AlfB-W130H, secondary hydrolysis was significantly decreased.

The two most successful mutations, H80F and W130H were transposed to another fucosidase from *Lactobacillus casei*, AlfC, which is reported to synthesize α-L-Fuc*p*-(1→6)-D-GlcNAc[60,61]. Although AlfB and AlfC are only distantly related (below 30% identity), analysis of AlfC-H87F and AlfC-W137H revealed that these also procured increased disaccharide yields in reactions containing equimolar (8 mM) amounts of α-L-Fuc*p*OCNP and D-GlcNAc (Figs. 5, S8 and S9). Significantly, this result illustrates how successful mutations generated in one enzyme can be transposed to other GH family members without the need to perform further analyses. In addition to the reduction of the sequence space and increased relevance, it is a key advantage of targeting conserved residues.

7. GH51, α-L-arabinofuranosidase *Tx*Abf

Five-carbon furanose rings are notoriously more flexible and thermodynamically less stable than their six-carbon counterparts, meaning that unlike pyranoses, furanoses can interconvert between different ring conformations. This reflects similar energy states for the different conformations[62,63], a point that complicates the organic synthesis of furanosides. Accounting for the fact that catalysis mediated by any given GH involves a specific conformational itinerary for the donor substrate, the lowered energy barriers between furanoside conformations[62] implies that furanosidase-catalysed reactions display altered mechanisms when compared to those catalysed by pyranosidases. Therefore, it is of interest to investigate to which extent the GH engineering approach described herein applies to furanosidases.

The α-L-arabinofuranosidase from *Thermobacillus xylanilyticus* (*Tx*Abf) belonging to the GH51 family[64] was used as a model furanosidase. Displaying inherent ability to perform transfuranosylation[65,66], *Tx*Abf has been the target of several studies aimed at improving its transglycosylation capability[22,67–69]. All eight mutants generated in this study showed improved ability to synthesize arabinoxylo-oligosaccharides (AXOS). Compared to *Tx*Abf-WT (9% yield), AXOS yields for the variants were in the range 19–82% (Fig. 6 and Table S4), with R69K, N175T, D297N and H240F/N being the best performers (37–82% overall yields). The transglycosylation activity (specific activity in transglycosylation mode, SA_T, Table S5) of the mutants was lower (0.1–66%) than that of the WT enzyme. It is noteworthy that R69K-, H240F- and D297N-catalysed transglycosylation reactions required relatively low amounts of catalyst (1.8−8 $\mu g \cdot mL^{-1}$, i.e. 32−140 nM; Fig. 6).

Previous work using random mutagenesis and screening already yielded mutants F26L and R69H[22] that, compared to F26H and R69K described herein, are slightly better catalysts for transglycosylation (1.2-fold). However, the current strategy is less labour-intensive, as it circumvents large library screening. In this respect, it is remarkable that the variants H240N and H240F display different yields and regioselectivities (Table S4), thus illustrating the value of further probing and fine-tuning of hotspots identified using our strategy.

Significantly, compared to all previously reported single-mutants that enhance the T/H ratio in *Tx*Abf[22,67–69], H240F/N display the highest overall transglycosylation yields combined with noticeably greater regioselectivity (Table S4). Combining this mutation with N216W, that favours higher regioselectivity towards the (1→2)-linkage of α-L-arabinofuranosyl moieties to the non-reducing terminal D-xylopyranosyl of xylotriose as acceptor (*i.e.* A^2XX)[22], revealed that in conjunction with both transglycosylation activity (SA_T, Table S5) and yield (62 compared to 82% for H240F) remaining relatively high, the reaction was almost completely regioselective (Table S4 and Fig. S10). Conveniently, available structural data and the considerable corpus of knowledge related to *Tx*Abf furnish hypotheses to explain how the different mutations enhance the T/H ratio (Fig. S11).

*Tx*Abf variant	Yield[a] (%)	Enzyme[b] (μg·mL^{-1}, nM)
WT	**9**	**0.2** (3.5)
F26H	**22**	**730** (12 800)
E28Q	**23**	**0.8** (14)
R69K	**37**	**2** (35)
N175T	**42**	**200** (3 500)
H240F	**82**	**8** (140)
H240N	**75**	**220** (3 900)
Y242F	**19**	**300** (5 300)
D297N	**47**	**1.8** (32)

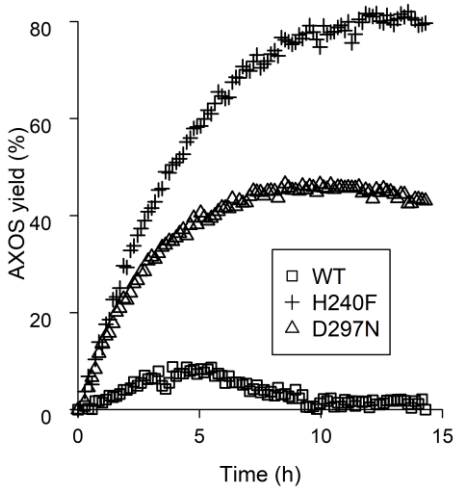


Fig. 6 | GH51 engineering. Left, [a]maximum overall yields of arabinoxylo-tetrasaccharides synthetized by *Tx*Abf-WT and mutants thereof using 5 mM α-L-Ara*f*O*p*NP as donor and 10 mM of xylotriose as acceptor. The yields of each regioisomers are indicated in Table S4. [b]Protein concentrations used to obtain the yields within 15 h given in μg·mL^{-1} and (nM). **Right,** NMR monitoring of transglycosylation yield progress with time for *Tx*Abf-WT and two of its more significant mutants.[4]

8. Discussion

Mutating enzymes from GH families 2, 10, 20, 29, and 51, which represent different clans (A, K and R), led to the successful enhancement of transglycosylation yields (2- to 9-fold compared to WT enzymes) in more than 50% of selected candidates. This powerful demonstration validates our conserved-residue approach and illustrates its applicability to retaining GHs irrespective of the structural fold or the precise mechanism. Moreover, a variety of sugars, D/L-configurations, pyranose/furanose conformations and α/β-stereochemistry are tolerated, thus new access to hitherto refractory syntheses is provided. In a rather fast and direct manner, the strategy procured the means to reach transglycosylation yields in the range 50−80%, and thus allowed the high yield synthesis of oligosaccharides such as α-L-Fuc*p*-(1→3)-D-GlcNAc, α-L-Fuc*p*-(1→6)-D-GlcNAc, lacto-*N*-triose II (LNT2), oligomannosides, oligoxylosides and arabinoxylo-oligosaccharides. Remarkably, in each of the five GHs families targeted, while using reasonable enzyme loadings, it proved possible to obtain at least one mutant displaying enhanced transglycosylation yield.

[4] The concentration of enzyme doesn't affect the transglycosylation yield.

Our approach is unusual in that it requires neither extensive screening, nor in-depth knowledge of the target enzyme. This is possible, because the method systematically targets conserved residues, which are generally omitted in enzyme engineering approaches in order to avoid loss of activity or stability. The power of the conserved sequence approach lies in the fact that it does not aim to improve a defining characteristic of an enzyme, but rather sets out to eliminate a property (*i.e.* the ability to perform hydrolysis). One obvious caveat is that some key determinants of the T/H ratio might not be conserved residues and will thus be undetectable using this method.

We anticipate that the conserved sequence approach can be applied to other enzymes displaying mechanistically-related activities (e.g. lipases performing hydrolysis and transesterification). However, it is necessary to stress that this strategy is unlikely to lead to the best possible solution. Instead it should be considered as a convenient and rather rapid first approach to pinpoint mutational hotspots. Mutants identified in this manner can then be further probed and recombined with other beneficial mutations in order to fine-tune potent biocatalysts for synthetic chemistry applications.

Converting hydrolytic GHs into potent glycosynthetic tools is an attractive approach to extend the synthetic chemist's toolbox, while introducing catalysts that obey green chemistry principles (e.g. use of non-toxic catalysts, aqueous solvents). However, so far, the success of this approach has been hampered by the time and effort necessary to develop appropriate biocatalysts for each target reaction. The strategy presented herein goes a long way to surmounting this obstacle, making it much simpler to obtain tailored biocatalysts that can then be used to operate straightforward, relatively inexpensive synthesis reactions that do not require difficult to obtain glycosyl sugar donors, lengthy protection/deprotection cycles or exorbitant quantities of enzyme. Moreover, the transferability implies that the reported mutations can be readily transferred to other GHs from the five described families, expanding the portfolio of available evolved transglycosylases and synthetic oligosaccharides.

9. Methods

9.1. Materials

Genes and genes variants, all codon-optimized for use in *E. coli* and inserted in pET24a, pET28b(+) or pET28a(+), were ordered from GenScript (Piscataway, USA) or Biomatik

(Ontario, Canada), respectively. Gene sequences can be found in the Uniprot database[70] under the codes Q9XCV4 (*Cf*Man2A), P96988 (Xyn10A), D4QAP4 (BbHI), A0A125UD88 (AlfB), A0A422MHI3 (AlfC) and O69262 (*Tx*Abf). Substrates were ordered from either Sigma-Aldrich, Carbosynth or Megazyme.

9.2. General procedures

9.2.1. Bioinformatics

Protein BLAST searches were performed on the NCBI server (https://blast.ncbi.nlm.nih.gov/Blast.cgi), from the non-redundant protein database[70], using default options, except in the case of the "Max target sequences" parameter, which was set at 20 000. Queries were made between January 31st, 2019 and May 3rd, 2019. Obtained sequences were clustered[26] to limit pairwise sequence identity at 80% by iterative cd-hit runs[27]. Iterative multiple sequence alignments were performed using ClustalΩ[28] to progressively increase minimum pairwise sequence identity to a predefined threshold (10–20%) and reach convergence using make_msa.sh, a homemade bash script, available at https://gitlab.univ-nantes.fr/sanejouand-yh/Sequences/tree/master. The same script was used to analyse sequence conservation.

9.2.2. Protein production and purification

Unless otherwise specified, pET24a, pET28a(+) or pET28b(+) plasmids bearing target genes were used to transform BL21(DE3) *E. coli* cells. Precultures of transformed cells were used to inoculate lysogeny broth media containing 30–50 $mg \cdot L^{-1}$ kanamycin. Cultures (0.5−2 L) were incubated at 37°C with shaking until OD_{600} reached ~0.5–1. Gene expression was induced with 200–500 µM isopropyl-β-D-1-thiogalactopyranoside and continued 3–16 h at 20–37°C. Cultures were subsequently centrifuged, pellets resuspended, lysed, and centrifuged. The enzymes were purified from the supernatant by Ni^{2+} (or Co^{2+})-affinity IMAC chromatography. Eluates were analysed by SDS-PAGE, and protein concentration was determined spectrophotometrically using UV_{280} absorption and molar extinction coefficients calculated in ExPASy (www.expasy.org).

9.2.3. GH Family-specific procedures

GH2, *Cf*Man2A

Activity of the *Cf*Man2A and its variants were determined by measuring the release of *p*-nitrophenol (*p*NP) from *p*-nitrophenyl-β-D-mannopyranoside (β-D-Man*p*O*p*NP), using a

variant of the method described by Zechel et al[34] adapted for microplate assay. Briefly, reactions mixtures containing 1 or 5 mM β-D-Man*p*O*p*NP, 35 mM sodium phosphate pH 7.0 and appropriately diluted enzyme were incubated at 35℃ for 10 min. The reaction was stopped by adding 1 M Na_2CO_3 and released *p*NP was measured on an Epoch Microplate spectrophotometer (BioTek Instruments, USA) at 405 nm. All assays and reactions (below) were in duplicates.

To evaluate the transglycosylation ability of the *Cf*Man2A forms, reactions containing 5 mM β-D-Man*p*O*p*NP and 3−34 μg·mL^{-1} enzyme in 35 mM sodium phosphate pH 7.0 were incubated up to 6 h at 35℃. Aliquots were collected at appropriate time intervals throughout the reaction, heat denatured (95°C, 10 min) and filtered (0.22 μm PTFE). Aliquots were separated at 40℃ on a Luna Omega SUGAR HPLC-column (Phenomenex, USA) using an UltiMate 3000 HPLC system (Thermo Fisher Scientific, USA) with 40:60 water:acetonitrile eluent (v/v) at 1 mL·min^{-1}. A VWD-3400RS detector was used (Thermo Fisher Scientific, USA), measuring absorbance at 300 nm. The presence of transglycosylation products were determined with mass spectrometry and analysed with NMR spectroscopy as described in supporting information, section 1.2.

Progress curves of *p*NP-Man_2 production for *Cf*Man2A forms were generated from HPLC analysis as above for reaction incubation times up to 6 h (7−10 sampling points, see Fig. 2). *p*NP-Man_2 was quantified using *p*NP-cellobioside as a standard in the HPLC analysis. The yield of *p*NP-Man_2 was calculated as the amount of β-D-Man*p*O*p*NP molecules used as either acceptor or donor in the production of *p*NP-Man_2 divided by the amount of loaded β-D-Man*p*O*p*NP.

GH10, *Rm*Xyn10A_CM

GH10 transglycosylation activity was characterized by incubating 20 mM X_4 with *Rm*Xyn10A_CM forms. Reaction mixtures containing 10–43 μg·mL^{-1} GH10 in 20 mM sodium phosphate pH 7.0 were incubated at 65℃ for 4 h. Aliquots (15 μL) were withdrawn at different time points, diluted in 0.5 mM NaOH, filtered, and analysed on high performance anion-exchange chromatography coupled with pulsed amperometric detection (HPAEC-PAD) using an ICS-5000 (Dionex) monitored by the software Chromeleon. Separation was carried out at 30℃ on a CarboPac PA-200 using 100 mM of NaOH at 0.5 mL·min^{-1} and a linear gradient of 0–120 mM sodium acetate.

The reaction mixtures for MALDI-TOF-MS analysis contained 10–43 μg·mL^{-1} GH10 in 20 mM sodium phosphate pH 7.0 and were incubated at 60℃ for 4 h. MALDI-TOF-MS spectra were obtained on a Bruker Daltonics Autoflex Speed MALDI-TOF (TOF) spectrometer in positive ion reflector mode and recorded in the mass range from 200 to 4000 or 5000 Da. Samples were diluted in MilliQ to a total salt concentration < 10 mM, and 1 μL of this dilution was mixed with 0.5 μL aqueous 10% dihydroxybenzoic acid (DHB) matrix solution on a stainless steel plate and left to dry at room temperature. Sample irradiation was done at 55% laser power by targeting the laser pulses at amorphous crystal regions, regularly shifting to remove heterogeneity in the sample. Calibration was done internally by addition of xylo-oligosaccharides X_1−X_6. The hydrolysis assay was performed in 0.1 mL reactions with 2 mM *p*-nitrophenyl-β-D-xylotrioside in 20 mM sodium phosphate pH 7.0, at 70℃ for 5 min and the reaction was stopped using one volume of 0.1 M NaOH. The absorbance was measured at 400 nm in Multiskan spectrophotometer from Thermo Scientific. The extinction coefficient (18250 M·cm^{-1}) of released *p*-nitrophenol (*p*NP) was determined using a standard of *p*NP (Sigma).

GH20, *Bb*HI

GH20 transglycosylation was monitored by HPAEC-PAD. Samples with 0.5–5 μM GH20, 10 mM β-D-Glc*p*NAcO*p*NP and 40 mM lactose in 50 mM sodium phosphate, 0.1% BSA pH 7.0 were incubated (2 h, 37°C), then heat denatured (2 min, 98°C), centrifuged, and the obtained supernatant was diluted five-fold in milliQ H_2O and filtered. Separation was carried out at 30°C on a CarboPac PA-1 (Dionex) using an ICS-5000 (Dionex) monitored by the software Chromeleon (Dionex). NaOH 0.1 M was used as eluent in 20 min separations at a flowrate of 250 μL·min^{-1}.

GH29, AlfB and AlfC

GH29 transfucosylation was assessed from NMR spectra recorded on an 800 MHz Bruker Avance III (799.75 MHz for ^{1}H) equipped with a 5 mm TCI cryoprobe using ^{1}H with presaturation. AlfB-catalysed reactions were carried out with 20 mM D-GlcNAc, 20 mM α-L-Fuc*p*OCNP, and 20–750 μg·mL^{-1} (0.4–16 nM) GH29 in 600 μL 40 mM sodium phosphate pD 7.0 (pD=pH$_{\text{meter reading}}$+0.4, 6.6 on reading)[71]. Time course experiments were obtained using pseudo-2D kinetics experiments, with ^{1}H NMR spectra recorded every 3 min. Integration of anomeric protons is inaccurate due to the closeness of the presaturated

HOD peak, therefore chemical shifts at 4.26 (H-2 of D-GlcNAc in α-L-Fuc*p*-(1→3)-D-GlcNAc) and 4.05 ppm (α-L-Fuc*p*OCNP) were used for integration (Fig. S6). AlfC-catalysed reactions were identical but for the use of 8 mM donor and acceptor, and the use of the fucose methyl group protons to assess concentrations (Figs. S7 and S8).

GH51, *Tx*Abf

The activities of *Tx*Abf and mutants thereof were determined using a discontinuous enzyme assay[67]. Reactions were performed in triplicate at 45°C in 50 mM sodium phosphate pH 7.0 containing 1 mg·mL^{-1} BSA, using 5 mM α-L-Ara*f*O*p*NP without/with 10 mM xylotriose as acceptor in hydrolysis and transglycosylation modes, respectively. The amount of *p*NP released, which was measured at 401 nm, was calculated using an appropriate standard curve of *p*NP. Negative controls containing all of the reactants except the enzyme were used to correct for spontaneous hydrolysis of the donor substrate.

To monitor the transglycosylation products profiles from ^{1}H NMR spectra at 500 MHz on a Bruker Avance II spectrometer equipped with a TCI probe, reactions were performed at 45°C in NMR tubes containing 600 µL of 5 mM α-L-Ara*f*O*p*NP, 10 mM xylotriose and *Tx*Abf enzymes (3.5−12 800 nM; enzyme solution correspond to 10% of total reaction volume) in 10 mM sodium phosphate pH 7.0/D_2O: 9/1, v/v, with 1 mg·mL^{-1} BSA. The quantity of enzyme (Fig. 6) was adjusted to suit the 13−19 h reaction time frame. Time course of NMR monitoring was obtained using pseudo-2D kinetics experiments based on a phase sensitive NOESY sequence with presaturation, with spectra collected every approximatively 9 min (i.e. twice 32 scans). The transglycosylation yields were determined by relative integration of anomeric proton signals from the α-L-Ara*f* unit of each AXOS (Table S4)[22,67]. Concentration of α-L-Ara*f*O*p*NP donor was quantified by integrating its relevant anomeric proton signals at 5.86 ppm. Molar balances, based on initial donor signal as internal reference, were used to convert the transglycosylation product signal integral into concentration.

10. Acknowledgements

This work was supported by a postdoctoral fellowship grant (NNF17OC0025660) to D. Teze from Novo Nordisk Foundation Synthesis and Production program. The PhD fellowship of J.Z. was supported by the China Scholarship Council (CSC). The GH29 NMR spectra were recorded at the NMR Center DTU, supported by the Villum Foundation. The GH51 NMR analyses were performed using facilities at MetaToul (Metabolomics &

Fluxomics Facilities, Toulouse, France, www.metatoul.fr), which is part of the national infrastructure MetaboHUB (The French National infrastructure for metabolomics and fluxomics, www.metabohub.fr) and is supported by grants from the Région Midi-Pyrénées, the European Regional Development Fund, SICOVAL, IBiSa-France, CNRS, and INRA. Guy Lippens (TBI) is gratefully acknowledged for insightful discussions and technical development of the NMR pseudo-2D kinetics experiments. H.S. thank the Swedish Foundation for Strategic Research (grant RBP14-0046) and FORMAS (grant **942-2016-117)** for financial support. We thank Julia Tanas Tanasi and Karina Jensen for technical support, and Teresa Mazarella and Haleema Saadia for protein production in GH 20 and 29.

Author contributions

The project was conceived by D.T. and overlooked and coordinated by D.T. and B.S.. Y.-H.S. conceived the conservation analysis approach, and D.T performed the bioinformatics. M.W. and H.S. conceived the GH2 experiments, which were performed by M.W. and G.C. (NMR analysis). K.Z.G.A. and E.N.-K. conceived the GH10 experiments, which were performed by K.Z.G.A.. B.S., D.T. and R.L. conceived the GH20 experiments, which were performed by R.L.. B.S., D.T. and J.D. conceived the GH29 experiments, which were performed by D.T.. J. Z. and R.F. conceived the GH51 experiments, which were performed by J.Z and supervised by M.O.D. and R.F.. D.T. wrote the first draft, and all authors contributed to write the paper.

Competing interests

The authors declare no competing financial interests.

References

1. Laine, R. A. Invited commentary: A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05×10^{12} structures for a reducing hexasaccharide: the Isomer Barrier to development of single-method saccharide sequencing or synthesis systems. *Glycobiology* **4**, 759–767 (1994).

2. Nielsen, M. M. & Pedersen, C. M. Catalytic glycosylations in oligosaccharide synthesis. *Chem. Rev.* **118**, 8285–8358 (2018).

3. Panza, M., Pistorio, S. G., Stine, K. J. & Demchenko, A. V. Automated chemical oligosaccharide synthesis: novel approach to traditional challenges. *Chem. Rev.* **118**,

8105–8150 (2018).

4. Krasnova, L. & Wong, C. H. Oligosaccharide synthesis and translational innovation. *J. Am. Chem. Soc.* **141**, 3735–3754 (2019).

5. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* **42**, 490–495 (2014).

6. Yonekura-Sakakibara, K. & Hanada, K. An evolutionary view of functional diversity in family 1 glycosyltransferases. *Plant J.* **66**, 182–193 (2011).

7. Faijes, M., Castejón-Vilatersana, M., Val-Cid, C. & Planas, A. Enzymatic and cell factory approaches to the production of human milk oligosaccharides. *Biotechnol. Adv.* **37**, 667–697 (2019).

8. Vasella, A., Davies, G. J. & Böhm, M. Glycosidase mechanisms. *Curr. Opin. Chem. Biol.* **6**, 619–29 (2002).

9. Koshland, D. E. Stereochemistry and the mechanism of enzymatic reactions. *Biol. Rev.* **28**, 416–436 (1953).

10. Teze, D. *et al.* The catalytic acid-base in GH109 resides in a conserved GGHGG loop and allows for comparable α-retaining and β-inverting activity in an N-acetylgalactosaminidase from *Akkermansia muciniphila*. *ChemRxiv* **Preprint**, (2019).

11. Lundemo, P., Karlsson, E. N. & Adlercreutz, P. Eliminating hydrolytic activity without affecting the transglycosylation of a GH1 β-glucosidase. *Appl. Microbiol. Biotechnol.* **101**, 1121–1131 (2017).

12. Bissaro, B., Monsan, P., Fauré, R. & O'Donohue, M. J. J. Glycosynthesis in a waterworld: new insight into the molecular basis of transglycosylation in retaining glycoside hydrolases. *Biochem. J.* **467**, 17–35 (2015).

13. Mackenzie, L. F., Wang, Q., Warren, R. A. J. & Withers, S. G. Glycosynthases: mutant glycosidases for oligosaccharide synthesis. *J. Am. Chem. Soc.* **120**, 5583–5584 (1998).

14. Hayes, M. R. & Pietruszka, J. Synthesis of glycosides by glycosynthases. *Molecules* **22**, (2017).

15. Malet, C. & Planas, A. From β-glucanase to β-glucansynthase: Glycosyl transfer to

α-glycosyl fluorides catalyzed by a mutant endoglucanase lacking its catalytic nucleophile. *FEBS Lett.* **440**, 208–212 (1998).

16. Moracci, M., Trincone, A., Perugino, G., Ciaramella, M. & Rossi, M. Restoration of the activity of active-site mutants of the hyperthermophilic β-glycosidase from *Sulfolobus solfataricus* dependence of the mechanism on the action of external nucleophiles. *Biochemistry* **37**, 17262–17270 (1998).

17. Mayer, C., Zechel, D. L., Reid, S. P., Warren, R. A. J. & Withers, S. G. The E358S mutant of *Agrobacterium sp.* β-glucosidase is a greatly improved glycosynthase. *FEBS Lett.* **466**, 40–44 (2000).

18. Yang, M., Davies, G. J. & Davis, B. G. A glycosynthase catalyst for the synthesis of flavonoid glycosides. *Angew. Chemie Int. Ed.* **46**, 3885–3888 (2007).

19. Packer, M. S. & Liu, D. R. Methods for the directed evolution of proteins. *Nat. Rev. Genet.* **16**, 379–394 (2015).

20. Teze, D., Daligault, F., Ferrières, V., Sanejouand, Y. H. & Tellier, C. Semi-rational approach for converting a GH36 α-glycosidase into an α-transglycosidase. *Glycobiology* **25**, 420–427 (2015).

21. Teze, D. *et al.* Semi-rational approach for converting a GH1 β-glycosidase into a β-transglycosidase. *Protein Eng. Des. Sel.* **27**, 13–19 (2014).

22. Bissaro, B. *et al.* Molecular design of non-Leloir furanose-transferring enzymes from an α-L-arabinofuranosidase: a rationale for the engineering of evolved transglycosylases. *ACS Catal.* **5**, 4598–4611 (2015).

23. Yang, J. *et al.* Engineering *T. naphthophila* β-glucosidase for enhanced synthesis of galactooligosaccharides by site-directed mutagenesis. *Biochem. Eng. J.* **127**, 1–8 (2017).

24. Wu, Y. *et al.* Enhancing the production of galacto-oligosaccharides by mutagenesis of *Sulfolobus solfataricus* β-galactosidase. *Food Chem.* **138**, 1588–1595 (2013).

25. Davies, G. J., Wilson, K. S. & Henrissat, B. Nomenclature for sugar-binding subsites in glycosyl hydrolases. *Biochem. J.* **321**, 557–559 (1997).

26. Li, W., Jaroszewski, L. & Godzik, A. Clustering of highly homologous sequences to reduce the size of large protein databases. *Bioinformatics* **17**, 282–283 (2001).

27. Huang, Y., Niu, B., Gao, Y., Fu, L. & Li, W. CD-HIT suite: a web server for

clustering and comparing biological sequences. *Bioinformatics* **26**, 680–682 (2010).

28. Sievers, F. & Higgins, D. G. Clustal Omega. *Curr. Protoc. Bioinforma.* **2014**, 3.13.1-3.13.16 (2014).

29. Davies, G. J., Planas, A. & Rovira, C. Conformational analyses of the reaction coordinate of glycosidases. *Acc. Chem. Res.* **45**, 308–316 (2011).

30. Gridley, J. J. & Osborn, H. M. I. Recent advances in the construction of β-D-mannose and β-D-mannosamine linkages. *J. Chem. Soc. Perkin Trans. 1* 1471–1491 (2000).

31. Nashiru, O. *et al.* β-mannosynthase: synthesis of β-mannosides with a mutant β-mannosidase. *Angew. Chemie Int. Ed.* **40**, 417–420 (2001).

32. Jahn, M. *et al.* Expansion of the glycosynthase repertoire to produce defined manno-oligosaccharides. *Chem. Commun.* **3**, 1327–1329 (2003).

33. Sasaki, A., Ishimizu, T., Geyer, R. & Hase, S. Synthesis of β-mannosides using the transglycosylation activity of endo-β-mannosidase from *Lilium longiflorum*. *FEBS J.* **272**, 1660–1668 (2005).

34. Zechel, D. L. *et al.* Mechanism, mutagenesis, and chemical rescue of a β-mannosidase from *Cellulomonas fimi*. *Biochemistry* **42**, 7195–7204 (2003).

35. Eneyskaya, E. V. *et al.* Transglycosylating and hydrolytic activities of the β-mannosidase from *Trichoderma reesei*. *Biochimie* **91**, 632–638 (2009).

36. Shi, P. *et al.* Cloning and characterization of a new β-mannosidase from *Streptomyces* sp. S27. *Enzyme Microb. Technol.* **49**, 277–283 (2011).

37. Stoll, D., Stålbrand, H. & Warren, R. A. J. Mannan-degrading enzymes from *Cellulomonas fimi*. *Appl. Environ. Microbiol.* **65**, 2598–2605 (1999).

38. Zechel, D. L. *et al.* Mechanism, mutagenesis, and chemical rescue of a β-mannosidase from Cellulomonas fimi. *Biochemistry* **42**, 7195–7204 (2003).

39. Morrill, J. *et al.* β-Mannanase-catalyzed synthesis of alkyl mannooligosides. *Appl. Microbiol. Biotechnol.* **102**, 5149–5163 (2018).

40. Rosengren, A. *et al.* Enzymatic synthesis and polymerisation of β-mannosyl acrylates produced from renewable hemicellulosic glycans. *Green Chem.* **21**, 2104–2118 (2019).

41. Iglesias-Fernández, J., Raich, L., Ardèvol, A. & Rovira, C. The complete conformational free energy landscape of β-xylose reveals a two-fold catalytic itinerary for β-xylanases. *Chem. Sci.* **6**, 1167–1177 (2015).

42. Evangelista, D. E., Kadowaki, M. A. S., Mello, B. L. & Polikarpov, I. Biochemical and biophysical characterization of novel GH10 xylanase prospected from a sugar cane bagasse compost-derived microbial consortia. *Int. J. Biol. Macromol.* **109**, 560–568 (2018).

43. Hu, J. & Saddler, J. N. Why does GH10 xylanase have better performance than GH11 xylanase for the deconstruction of pretreated biomass? *Biomass Bioenerg.* **110**, 13–16 (2018).

44. Nordberg Karlsson, E., Schmitz, E., Linares-Pastén, J. A. & Adlercreutz, P. Endo-xylanases as tools for production of substituted xylooligosaccharides with prebiotic properties. *Appl. Microbiol. Biotechnol.* **102**, 9081–9088 (2018).

45. Niderhaus, C., Garrido, M., Insani, M., Campos, E. & Wirth, S. Heterologous production and characterization of a thermostable GH10 family endo-xylanase from *Pycnoporus sanguineus* BAFC 2126. *Process Biochem.* **67**, 92–98 (2018).

46. Armand, S., Andrews, S. R., Charnock, S. J. & Gilbert, H. J. Influence of the aglycone region of the substrate binding cleft of *Pseudomonas* xylanase 10A on catalysis. *Biochemistry* **40**, 7404–7409 (2001).

47. Charnock, S. J. *et al.* Key residues in subsite F play a critical role in the activity of *Pseudomonas fluorescens* subspecies cellulosa xylanase A against xylooligosaccharides but not against highly polymeric substrates such as xylan. *J. Biol. Chem.* **272**, 2942–2951 (1997).

48. Charnock, S. J. *et al.* The topology of the substrate binding clefts of glycosyl hydrolase family 10 xylanases are not conserved. *J. Biol. Chem.* **273**, 32187–32199 (1999).

49. Moreau, A., Shareck, F., Kluepfel, D. & Morosoli, R. Alteration of the cleavage mode and of the transglycosylation reactions of the xylanase A of *Streptomyces lividans* 1326 by site-directed mutagenesis of the Asn173 residue. *Eur. J. Biochem.* **219**, 261–266 (1994).

50. Aronsson, A. *et al.* Structural insights of RmXyn10A – A prebiotic-producing GH10 xylanase with a non-conserved aglycone binding region. *Biochim. Biophys.*

Acta, Proteins Proteomics **1866**, 292–306 (2018).

51. Roth, C. *et al.* Structural and mechanistic insights into a *Bacteroides vulgatus* retaining N-acetyl-β-galactosaminidase that uses neighbouring group participation. *Chem. Commun.* **52**, 11096–11099 (2016).

52. Coines, J., Alfonso-Prieto, M., Biarnés, X., Planas, A. & Rovira, C. Oxazoline or oxazolinium ion? The protonation state and conformation of the reaction intermediate of chitinase enzymes revisited. *Chem. - A Eur. J.* **24**, 19258–19265 (2018).

53. Slámová, K., Bojarová, P., Petrásková, L. & Křen, V. β-N-Acetylhexosaminidase: what's in a name...? *Biotechnol. Adv.* **28**, 682–693 (2010).

54. Slámová, K. *et al.* Synthesis of derivatized chitooligomers using transglycosidases engineered from the fungal GH20 β-N-acetylhexosaminidase. *Adv. Synth. Catal.* **357**, 1941–1950 (2015).

55. Zeuner, B., Teze, D., Muschiol, J. & Meyer, A. S. Synthesis of human milk oligosaccharides: protein engineering strategies for improved enzymatic transglycosylation. *Molecules* **24**, 2033–2055 (2019).

56. Schneider, M., Al-Shareffi, E. & Haltiwanger, R. S. Biological functions of fucose in mammals. *Glycobiology* **27**, 601–618 (2017).

57. Zeuner, B., Vuillemin, M., Holck, J., Muschiol, J. & Meyer, A. S. A. S. Loop engineering of an α-1,3/4-L-fucosidase for improved synthesis of human milk oligosaccharides. *Enzyme Microb. Technol.* **115**, 37–44 (2018).

58. Saumonneau, A. *et al.* Design of an α-L-transfucosidase for the synthesis of fucosylated HMOs. *Glycobiology* **26**, 261–269 (2015).

59. Osanjo, G. *et al.* Directed evolution of the α-L-fucosidase from *Thermotoga maritima* into an α-L-transfucosidase. *Biochemistry* **46**, 1022–1033 (2007).

60. Rodríguez-Díaz, J., Carbajo, R. J., Pineda-Lucena, A., Monedero, V. & Yebra, M. J. Synthesis of fucosyl-N-acetylglucosamine disaccharides by transfucosylation using α-L-fucosidases from *Lactobacillus casei*. *Appl. Environ. Microbiol.* **79**, 3847–3850 (2013).

61. Rodríguez-Díaz, J., Monedero, V. & Yebra, M. J. Utilization of natural fucosylated oligosaccharides by three novel α-L-fucosidases from a probiotic *lactobacillus casei*

strain. *Appl. Environ. Microbiol.* **77**, 703–705 (2011).

62. Taha, H. A., Richards, M. R. & Lowary, T. L. Conformational analysis of furanoside-containing mono- and oligosaccharides. *Chem. Rev.* **113**, 1851–1876 (2013).

63. Wang, X. & Woods, R. J. Insights into furanose solution conformations: beyond the two-state model. *J. Biomol. NMR* **64**, 291–305 (2016).

64. Debeche, T., Cummings, N., Connerton, I., Debeire, P. & O'Donohue, M. J. Genetic and biochemical characterization of a highly thermostable α-L-arabinofuranosidase from *Thermobacillus xylanilyticus*. *Appl. Environ. Microbiol.* **66**, 1734–1736 (2000).

65. Rémond, C. *et al.* Synthesis of pentose-containing disaccharides using a thermostable α-L-arabinofuranosidase. *Carbohydr. Res.* **339**, 2019–2025 (2004).

66. Remond, C., Plantier-Royon, R., Aubrya, N. & O'Donohue, M. J. An original chemoenzymatic route for the synthesis of β-D-galactofuranosides using an α-L-arabinofuranosidase. *Carbohydr. Res.* **340**, 637–644 (2005).

67. Bissaro, B. *et al.* Mutation of a pH-modulating residue in a GH51 α-L-arabinofuranosidase leads to a severe reduction of the secondary hydrolysis of transfuranosylation products. *Biochim. Biophys. Acta* **1840**, 626–636 (2014).

68. Arab-jaziri, F. *et al.* Enhancing the chemoenzymatic synthesis of arabinosylated xylo-oligosaccharides by GH51 α-L-arabinofuranosidase. *Carbohydr. Res.* **401**, 64–72 (2015).

69. Arab-Jaziri, F. *et al.* Engineering transglycosidase activity into a GH51 α-L-arabinofuranosidase. *New Biotechnol.* **30**, 536–544 (2013).

70. Apweiler, R. *et al.* UniProt: the Universal Protein knowledgebase. *Nucleic Acids Res.* **32**, D115-9 (2004).

71. Covington, A. K., Paabo, M., Robinson, R. A. & Bates, R. G. Use of the glass electrode in deuterium oxide and the relation between the standardized pD (paD) scale and the operational pH in heavy water. *Anal. Chem.* **40**, 700–706 (1968).

Supplementary information

Rational enzyme design without structural knowledge: a sequence-based approach for efficient generation of glycosylation catalysts

David Teze[1,*], Jiao Zhao[2], Mathias Wiemann[3], Kazi Zubaida Gulshan Ara[4], Rossana Lupo[1], Mette Errebo Rønne[1], Göran Carlström[5], Jens Duus[6], Yves-Henri Sanejouand[7], Michael J. O'Donohue[2], Eva Nordberg-Karlsson[4], Régis Fauré[2], Henrik Stålbrand[3], Birte Svensson[1,*]

This SI section contains methods and a range of results from analytics (high-performance liquid chromatography, HPLC; mass spectrometry, MS; nuclear magnetic resonance spectroscopy, NMR) used to identify specific synthetic compounds and monitor their production during transglycosylation reactions.

1) GH2, β-*mannosidase Cf*Man2A

HPLC and MS analysis of *Cf*Man2A transglycosylation products

HPLC (LUNA Omega Sugar column) was used to monitor the formation of transglycosylation products arising from *Cf*Man2A-catalyzed reactions containing 5 mM *p*-nitrophenyl β-D-mannopyranoside (β-D-Man*p*O*p*NP) as both donor and acceptor substrate (Figure S1). For MS and NMR analyses, a 4 mL reaction mixture containing 5 mM β-D-Man*p*O*p*NP, 34 μg·mL^{-1} *Cf*Man2A-N428T, in 35 mM Na_2HPO_4 pH 7.0 was incubated for 2 h at 35℃. The reaction was stopped by heat denaturation (95℃, 10 min) and 100 μL was filtered and separated using HPLC. Elution volumes for β-D-Man*p*O*p*NP (peak 1) and *p*NP (peak 5) were determined by injection of chemical standards. Fractions eluting between 3.3 and 4.1 mL were collected (85 μL), covering peaks 2 through 5 (Fig. S1). The collected fractions were concentrated approximately 10-fold by evaporation in an RVC 2-18 vacuum concentrator (Martin Christ Freeze Dryers, Germany) and analysed using matrix-assisted laser desorption/ionisation time-of-flight MS (MALDI-TOF MS). Fractions (0.5 μL) were combined with 0.5 μL matrix solution (10 mg·mL^{-1} 2,5-dihydroxybenzoic acid in 5 mM sodium acetate pH 5.3) directly on a stainless steel sample plate and dried under warm air. The fractions were analysed on a 4700 Proteomics Analyzer (Applied Biosystems, USA) and data were analysed using DataExplorer software (Applied Biosystems, USA). MS analysis revealed that peaks 2 and 3 (eluted between 3.3 and 3.6 mL) of the chromatogram (Fig. S1) contained compounds with m/z corresponding to *p*-nitrophenyl-β-D-mannobioside (*p*NP-Man_2). Similarly, MS revealed that fractions collected between 3.8 and 4 mL (i.e. peaks 4 and 5) contain *p*-nitrophenyl-β-D-mannotrioside (*p*NP-Man_3).

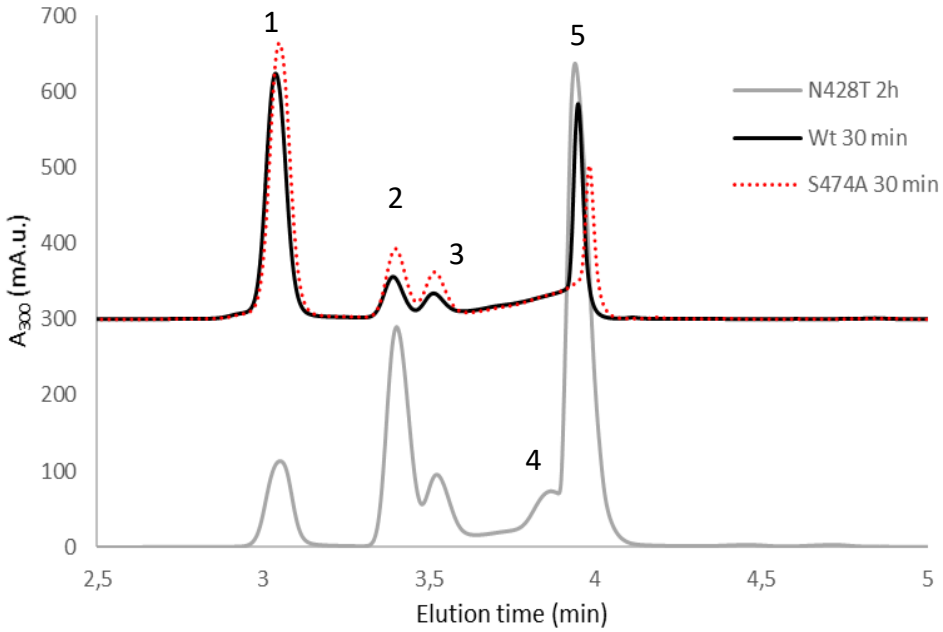


Figure S1 | HPLC analysis of transglycosylation reactions using *Cf*Man2A-WT (black line) and variants S474A (red dotted line; both reaction mixtures diluted four times) and N428T (grey line, undiluted reaction). The N428T reaction mixture was used for NMR analysis and peak identification by MS. Peak 1 is β-D-Man*p*O*p*NP as determined with a compound standard, peaks 2 and 3 were identified as *p*NP-Man_2 with MALDI-TOF MS (observed m/z: 486.10, theoretical $[M+Na]^+$ for *p*NP-Man_2: 486.12), peak 4 is *p*NP-Man_3 (observed m/z: 648.14, theoretical $[M+Na]^+$ for *p*NP-Man_3: 648.18) and peak 5 was identified as released *p*NP using a compound standard.

NMR analysis of the *Cf*Man2A-catalyzed transglycosylation reactions

NMR methods

The enzyme *Cf*Man2A-N428T was incubated with 5 mM *p*-nitrophenyl-β-D-mannopyranoside (β-D-Man*p*O*p*NP) for 2 h before removing 4 mL of the reaction mixture for NMR spectroscopic analysis. The sample was lyophilized, resuspended in 0.6 mL D_2O (99.9 atom % deuterium, Sigma-Aldrich), lyophilized and resuspended in 0.6 mL D_2O.

NMR spectra were acquired at 318 K using an Agilent Varian VNMRS 500 MHz spectrometer equipped with a 5 mm HCN probe. This temperature was chosen to prevent interference from the resonance of residual HDO. The NMR chemical shifts were referenced with respect to the resonances of the anomeric CH-group (H-1) of α-D-mannose at 5.17 (^{1}H) and 96.8 ppm (^{13}C), these values from the NMR chemical shift database BioMagResBank[1] (www.bmrb.wisc.edu) were measured using a solution of α-D-mannose (0.5 mM) in 50 mM phosphate buffer, equivalent to pH 7.4, in D_2O at 298 K, referenced to sodium trimethylsilylpropane-sulfonate (DSS, entry no. bmse000018). Standard 2D experiments (DQF-COSY, z-TOCSY with 120 ms mixing time, NOESY with 500 ms mixing time, ^{13}C-HSQC, and ^{13}C-HMBC, optimized for 8 Hz) were used for chemical shift assignments, using SPARKY[2] software. TOCSY and NOESY experiments were typically run using a spectral width of 4680 Hz in the indirect dimension, 512 increments, 32 scans, and a recycle delay of 2 s, resulting in roughly 22 h acquisition time. The 2D ^{13}C-HSQC experiment was acquired with improved resolution in the indirect dimension using non-uniform sampling (NUS) of 400 (41%) out of a total of 974 increments. The experiment used a recycle delay of 2 s, 32 scans, and an indirect spectral width of 12066 Hz, giving a total acquisition time of 16 h. The NUS sampled ^{13}C-HSQC experiment was reconstructed and processed using NMRPipe[3]. The $^1J_{\text{C-1,H-1}}$ scalar coupling constants were determined from the ^{13}C-satellites in a 1D ^{1}H spectrum.

An additional high resolution 1D ^{1}H spectrum at 318 K was obtained using a Bruker 500 MHz AVANCE III HD spectrometer equipped with a 5 mm BBOF probe. 1024 transients were acquired during 5 s and a recycle delay of 1 s. Using this spectrum, the $^3J_{\text{H-1,H-2}}$ scalar coupling constants were estimated from the doublet splitting of the anomeric H-1 signals in the spectrum, processed with a pure sine window function to obtain highest possible resolution. Estimates of the relative amounts of the different reaction products were obtained from integration of H-1 and resolved H-2 resonances in a 1D ^{1}H NMR spectrum.

NMR results

When using D-manno (D-Man) groups, samples were expected to contain the substrate β-D-Man*p*O*p*NP, transglycosylation products (Fig. S1) and D-mannose that results from hydrolysis of the substrate. Identification of the transglycosylation products was performed using ^{1}H- and ^{13}C-NMR spectroscopy. The 1D ^{1}H spectrum of the prepared sample revealed the presence of ten anomeric resonances, corresponding to ten D-Man*p* moieties (labelled A-J). Three of the D-Man units could be identified as α-D-Man*p*, β-D-Man*p*, and β-D-Manp^{B}O*p*NP, i.e. D-Man*p* unit B, from comparisons to ^{1}H- and ^{13}C- NMR spectra of the substrate β-D-Man*p*O*p*NP (NMR data acquired *in-house*), or using the NMR chemical shift database BioMagResBank[1], for α-, and β-D-Man*p* residues. Each of the spin systems with the remaining anomeric resonances were identified, and virtually complete ^{1}H and ^{13}C resonance assignments of the corresponding D-Man*p* residues were obtained, excluding the CH_2-groups in position 6 (Table S1). The chemical shifts of the ^{1}H anomeric (H-1) resonances are consistent with expected chemical shifts for substituted β-D-Man*p* units[4,5]. Three of the anomeric protons (from D-Man*p* moieties A–C) show correlations in the ^{13}C-HMBC spectrum to the *p*-nitrophenyl group, identifying the substituted D-Man*p* residues. Information on the connectivity between the different D-Man*p* residues were then obtained from elevated ^{13}C chemical shift of the substituted carbon[6], direct observation of heteronuclear three-bond correlations over the glycosidic linkage, and observation of NOE between protons close in space. The *p*NP-linked D-Man*p* units A and C, together with the remaining five D-Man*p* units (labelled D–G & I) could be shown to belong to D-Man*p* units of four reaction products, two dimannosides and two trimannosides, identified as β-D-Manp^{I}-(1⟶3)-β-D-Manp^{C}O*p*NP (18%), β-D-Manp^{E}-(1⟶4)-β-D-Manp^{A}O*p*NP (56%), β-D-Manp^{F}-(1⟶4)-β-D-Manp^{G}-(1⟶3)-β-D-Manp^{C}O*p*NP (13%), and β-D-Manp^{F}-(1⟶4)-β-D-Manp^{D}-(1⟶4)-β-D-Manp^{A}O*p*NP (13%), with the approximate relative yields given in parenthesis. In Table S2 we list the specific NMR observations used for the identifications of these different transglycosylation products. D-Man*p* residues termed either A, C, or F above have virtually identical chemical shifts as they have similar chemical environments, although they are each present in two different reaction products (Table S1). Many of the 2D resonances from these D-Man*p* groups show a slight deformation or skewness, indicating the presence of two signals. Distinct cross peaks and/or NOEs were observed from D-Man*p* residues A, C, and F to their respective adjacent D-Man*p* residues in the different reaction products (Table S2).

Our analysis of the two known dimannosides β-D-Man*p*-(1⟶3)-β-D-Man*p*O*p*NP and β-D-Man*p*-(1⟶4)-β-D-Man*p*O*p*NP is in good agreement with the literature (Table S1)[4], taking into account the different experimental temperatures and referencing procedures. However, it is noteworthy that comparing the previously reported[4] values with each other, ^{13}C-shifts for β-D-Man*p*-(1⟶4)-β-D-Man*p*O*p*NP are unusually small compared to their reported values for β-D-Man*p*-(1⟶3)-β-D-Man*p*O*p*NP, with a difference in the order of 3 ppm, possibly due to different referencing within the same study[4].

The anomeric configurations were in all cases determined to be of the form β-D-Man*p*, with an axial H-1. The value of the $^{1}J_{\text{C-1,H-1}}$ scalar coupling constant has been reported to be a reliable indicator of the anomeric configuration[7], and for D-Man*p* from groups A, C, and E the measured values for $^{1}J_{\text{C-1,H-1}}$ were all ~162 Hz, in contrast to a value of ~170 Hz for an α-configuration (i.e. with an equatorial H-1). The $^{1}J_{\text{C-1,H-1}}$ coupling constant could not be determined for the other D-mannosyl groups due to low sensitivity or spectral overlap. The $^{3}J_{\text{H-1,H-2}}$ scalar coupling constants could be determined for several of the D-Man*p* units, and were all ~1 Hz as expected for β-D-Man*p*-containing oligosaccharides[8]. Spectral overlap prevented determination of the coupling constants for units F and I. Additional confirmations for the β-configuration were obtained from strong NOE observed from H-1 to H-2, H-3, and H-5 protons[9], for all D-Man*p* moieties. D-Man*p* units D and G, which both are the internal unit in a trimannoside, have weaker H-1 to H-3 NOE.

Table S1 | ^{1}H and ^{13}C NMR chemical shifts (ppm) for identified dimannosides and trimannosides in D_2O at 318 K.

Carbon no.	**1**	**2**	**3**	**4**	**5**
β-D-Man*p*^{I}-(1⟶3)-β-D-Man*p*^{C}O*p*NP					
D-Man*p*^{C} – ^{1}H	5.47	4.43	4.06	3.84	3.64
D-Man*p*^{C} – ^{13}C	99.8	70.1	81.5[a]	67.7	79.0
D-Man*p*^{I} – ^{1}H	4.88	4.10	3.68	3.61	3.40
D-Man*p*^{I} – ^{13}C	99.7	73.5	75.7	69.6	79.2
β-D-Man*p*^{E}-(1⟶4)-β-D-Man*p*^{A}O*p*NP					
D-Man*p*^{A} – ^{1}H	5.50	4.29	3.92	3.95	3.73
D-Man*p*^{A} – ^{13}C	99.9	72.5	74.0	79.0[a]	77.9
D-Man*p*^{E} – ^{1}H	4.76	4.09	3.66	3.60	3.45
D-Man*p*^{E} – ^{13}C	102.9	73.3	75.6	69.4	79.2
β-D-Man*p*^{F}-(1⟶4)-β-D-Man*p*^{G}-(1⟶3)-β-D-Man*p*^{C}O*p*NP					
D-Man*p*^{C} – ^{1}H	5.47	4.43	4.06	3.84	3.64
D-Man*p*^{C} – ^{13}C	99.8	70.1	81.5[a]	67.7	79.0
D-Man*p*^{G} – ^{1}H	4.90	4.15	3.84	3.85	3.53
D-Man*p*^{G} – ^{13}C	99.6	73.0	74.3	79.4[a]	77.8
D-Man*p*^{F} – ^{1}H	4.73	4.06	3.65	3.58	3.44
D-Man*p*^{F} – ^{13}C	102.9	73.2	75.6	69.4	79.2
β-D-Man*p*^{F}-(1⟶4)-β-D-Man*p*^{D}-(1⟶4)-β-D-Man*p*^{A}O*p*NP					
D-Man*p*^{A} – ^{1}H	5.50	4.29	3.92	3.95	3.73
D-Man*p*^{A} – ^{13}C	99.9	72.5	74.0	79.0[a]	77.9
D-Man*p*^{D} – ^{1}H	4.78	4.14	3.82	3.84	3.57
D-Man*p*^{D} – ^{13}C	102.9	72.7	74.2	79.2[a]	77.8

D-Man*p*^{F} – ^{1}H	4.73	4.06	3.65	3.58	3.44
D-Man*p*^{F} – ^{13}C	102.9	73.2	75.6	69.4	79.2

[a]Elevated ^{13}C chemical shifts revealing site of substitutions are marked in red.

Table S2 | NMR information used for the identifications of transglycosylation products.

β-D-Man*p*^{I}-(1⟶3)-β-D-Man*p*^{C}O*p*NP: • ^{13}C chemical shift (δ = 81.5 ppm) for β-D-Man*p*^{C} C-3, • Direct cross peak in ^{13}C-HMBC between β-D-Man*p*^{I} H-1 and β-D-Man*p*^{C} C-3, • NOE between β-D-Man*p*^{I} H-1 and β-D-Man*p*^{C} H-2, • NOE between β-D-Man*p*^{I} H-1 and β-D-Man*p*^{C} H-3.
β-D-Man*p*^{E}-(1⟶4)-β-D-Man*p*^{A}O*p*NP: • ^{13}C chemical shift (δ = 79.0 ppm) for β-D-Man*p*^{A} C-4, • NOE between β-D-Man*p*^{E} H-1 and β-D-Man*p*^{A} H-4, • NOE between β-D-Man*p*^{E} H-1 and β-D-Man*p*^{A} H-3, • NOE between β-D-Man*p*^{E} H-1 and β-D-Man*p*^{A} H-6, Partial overlap of β-D-Man*p*^{A} C-4 and β-D-Man*p*^{E} C-5 prevents 3-bond correlation from β-D-Man*p*^{E} H-1 to be unambiguously observed.
β-D-Man*p*^{F}-(1⟶4)-β-D-Man*p*^{G}-(1⟶3)-β-D-Man*p*^{C}O*p*NP: • ^{13}C chemical shift (δ = 81.5 ppm) for β-D-Man*p*^{C} C-3, • ^{13}C chemical shift (δ = 79.4 ppm) for β-D-Man*p*^{G} C-4, • Direct cross peak in ^{13}C-HMBC between β-D-Man*p*^{G} H-1 and β-D-Man*p*^{C} C-3, • NOE between β-D-Man*p*^{G} H-1 and β-D-Man*p*^{C} H-2, • NOE between β-D-Man*p*^{G} H-1 and β-D-Man*p*^{C} H-1, • NOE between β-D-Man*p*^{F} H-1 and β-D-Man*p*^{G} H-4/3, • NOE between β-D-Man*p*^{F} H-1 and β-D-Man*p*^{D} H-4/3, Partial overlap of β-D-Man*p*^{G} C-4 and β-D-Man*p*^{F} C-5 prevents 3-bond correlation from β-D-Man*p*^{F} H-1 to be unambiguously observed.
β-D-Man*p*^{F}-(1⟶4)-β-D-Man*p*^{D}-(1⟶4)-β-D-Man*p*^{A}O*p*NP: • ^{13}C chemical shift (δ = 79.0 ppm) for β-D-Man*p*^{A} C-4, • ^{13}C chemical shift (δ = 79.2 ppm) for β-D-Man*p*^{D} C-4, • Direct cross peak in ^{13}C-HMBC between β-D-Man*p*^{D} H-1 and β-D-Man*p*^{A} C-4, • NOE between β-D-Man*p*^{F} H-1 and β-D-Man*p*^{D} H-4/3 • NOE between β-D-Man*p*^{D} H-1 and β-D-Man*p*^{A} H-4/3 Partial overlap of β-D-Man*p*^{D} C-4 and β-D-Man*p*^{F} C-5 prevents 3-bond correlation from β-D-Man*p*^{F} H-1 to be unambiguously observed.

2) GH10, β-xylanase *Rm*Xyn10A_CM

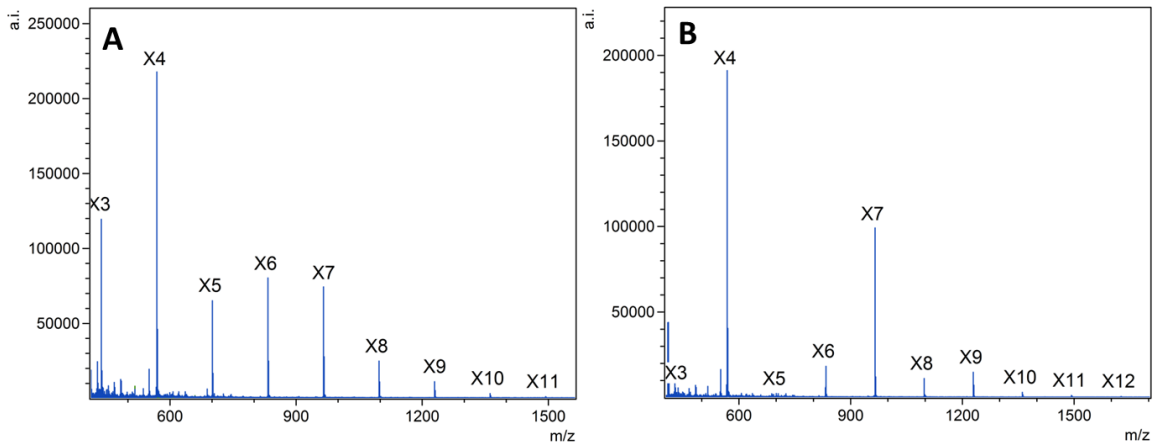


Figure S2 | Xylo-oligosaccharide analysis using mass spectrometry. MALDI-TOF-MS spectra of the sodium adducts of transglycosylation products generated using: **A.** *Rm*Xyn10A_CM-WT; **B.** *Rm*Xyn10A_CM-N118T. Reactions were run at 60°C and pH 7.0. Major transglycosylation product peaks were detected at m/z 833.4827, 965.5205 and 1097.5826 corresponding to the $[M+Na]^+$ ions of D-Xyl*p*-containing hexamer (X_6), heptamer (X_7) and octamer (X_8). Additional peaks were observed at m/z 1229.65, 1361.73 and 1493.83 corresponding to $[M+Na]^+$ ions of X_9, X_{10} and X_{11} (theoretical $[M+Na]^+$: 1229.38, 1361.42 and 1493.47).

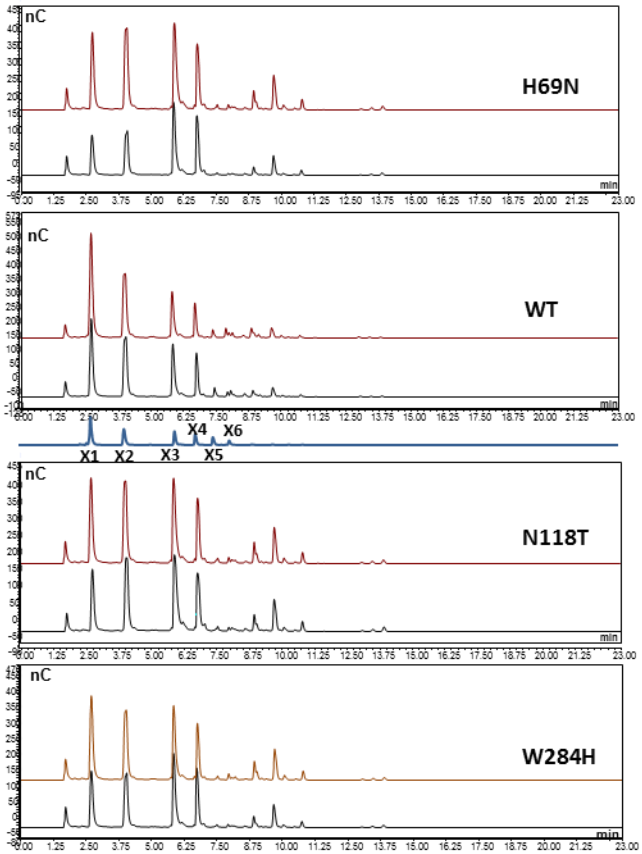


Figure S3 | HPAEC-PAD analysis of xylo-oligosaccharides. Chromatograms showing product profiles of transglycosylation reaction catalysed by *Rm*Xyn10A_CM-WT and its variants H69N, N118T and W284H. For each mutant, chromatograms display reaction profiles after 2 (black) and 4 h (brown), respectively. D-xylose (X_1) and xylo-oligosaccharides up to xylohexaose (X_6) were used as standards (blue) to facilitate product analysis.

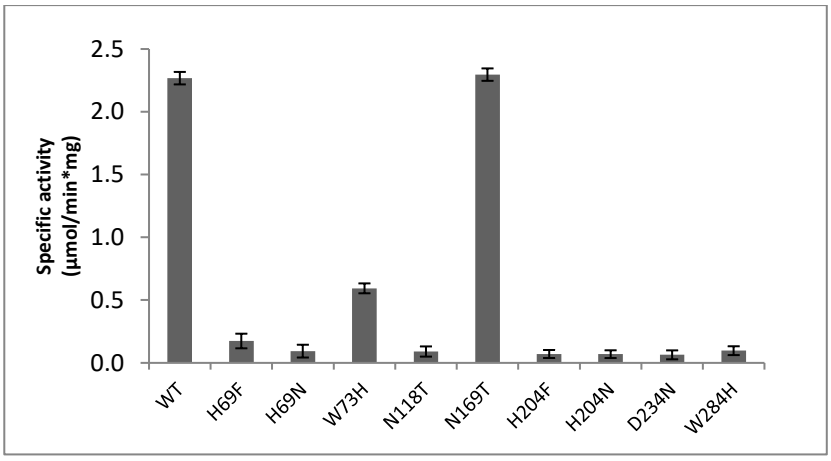


Figure S4 | *Rm*Xyn10A_CM forms hydrolytic activity. Hydrolytic activity of the catalytic module of *Rm*Xyn10A_CM forms (WT and mutants) on β-D-xylotrioside*O*p*NP*.

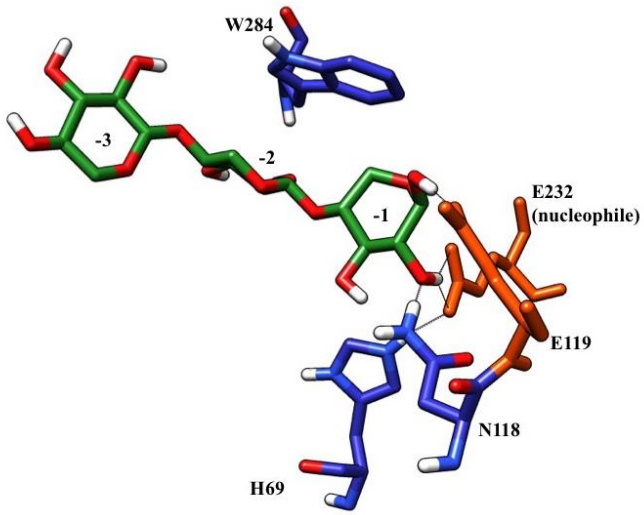


Figure S5 | Illustration of the binding region of modelled *Rm*Xyn10A_CM showing the position of H69, N118 and W284. Xylotriose (X3) coloured green is docked into –3 to –1 subsites. Catalytic residues are coloured orange and potential hydrogen bonds are shown as dotted black lines. This figure was prepared by using UCSF Chimera.

Table S3 | Areas (nC·min) of peaks corresponding to products identified in reactions catalysed by *Rm*Xyn10A_CM forms using 5 mM xylotetraose (X_4) as both donor and acceptor.

Time	**1 h**						**2 h**						**4 h**					
Peak	**X_5**	**X_6**	**X_7**	**X_8**	**X_9**	**X_{10}**	**X_5**	**X_6**	**X_7**	**X_8**	**X_9**	**X_{10}**	**X_5**	**X_6**	**X_7**	**X_8**	**X_9**	**X_{10}**
WT	**3.0** ±0.7	**2.7** ±0.5	**0.7** ±1.0	**0.9** ±1.3	**0.4** ±1.1	**0.1** ±1.4	**3.0** ±0.8	**1.2** ±0.4	**2.6** ±1.0	**3.8** ±2.0	**0.3** ±1.2	**0.6** ±1.0	**3.1** ±0.4	**2.4** ±0.5	**4.7** ±0.6	**4.9** ±2.1	**0.5** ±1.2	**0.6** ±1.0
H69F	**0.1** ±1.4	**0.3** ±2.1	**0.7** ±1.7	**1.4** ±0.4	**0.3** ±1.6	**0.2** ±1.1	**0.4** ±1.6	**0.3** ±1.1	**2.7** ±1.2	**6.8** ±1.1	**0.6** ±1.0	**1.5** ±1.2	**0.6** ±1.1	**0.6** ±1.3	**4.3** ±1.1	**8.4** ±2.3	**0.8** ±1.3	**1.6** ±0.5
H69N	**0.1** ±1.1	**0.3** ±1.1	**0.7** ±1.0	**1.5** ±0.1	**0.3** ±1.2	**0.3** ±1.3	**0.6** ±1.2	**0.4** ±1.3	**2.5** ±0.6	**6.4** ±2.7	**0.6** ±0.6	**1.5** ±1.0	**1.1** ±1.3	**0.8** ±1.3	**5.9** ±1.1	**11.6** ±1.9	**0.9** ±1.0	**1.8** ±0.4
W73H	**0.4** ±1.6	**0.7** ±1.2	**0.8** ±1.1	**1.7** **±0**.5	**0.3** ±2.5	**0.3** ±1.1	**0.9** ±2.5	**0.5** ±1.1	**2.8** **±**0.8	**6.6** ±1.2	**0.6** ±0.8	**1.5** ±1.0	**1.2** ±1.1	**0.6** ±1.5	**4.2** ±1.2	**8.3** ±2.6	**0.8** ±1.4	**1.6** ±0.6
N118T	**0.1** ±1.0	**0.4** ±2.0	**0.6** ±1.0	**1.5** ±1.3	**0.3** ±1.4	**0.3** ±2.9	**1.2** ±1.4	**0.9** ±2.9	**6.5** ±1.1	**13.4** ±2.2	**0.8** ±0.6	**2.0** ±0.9	**1.3** ±1.0	**1.0** ±1.4	**6.8** ±1.5	**12.1** ±3.3	**1.0** ±1.5	**1.7** ±1.4
N169T	**2.2** ±1.4	**2.6** ±1.1	**0.7** ±1.2	**1.4** ±3.0	**0.2** ±1.0	**0.2** ±1.6	**3.0** ±1.0	**1.9** ±1.6	**2.9** ±0.9	**5.5** ±1.5	**0.4** ±1.1	**1.2** ±0.9	**2.7** ±0.9	**1.1** ±1.0	**2.7** ±1.8	**4.8** ±1.9	**0.6** ±1.2	**1.2** ±1.0
H204F	**0.1** ±1.3	**0.1** ±1.6	**0.5** ±1.5	**1.3** ±1.1	**0.2** ±1.1	**0.2** ±1.4	**0.3** ±1.1	**0.4** ±1.4	**2.5** ±1.0	**5.8** ±1.8	**0.6** ±1.3	**1.4** ±1.0	**0.6** ±1.3	**0.7** ±1.1	**4.8** ±0.9	**8.9** ±2.4	**0.9** ±1.4	**1.5** ±1.2
H204N	**0.1** ±1.0	**0.1** ±1.2	**0.7** ±1.2	**1.5** ±0.3	**0.1** ±1.3	**0.2** ±1.2	**0.4** ±1.4	**0.4** ±1.2	**2.4** ±0.9	**6.0** ±0.5	**0.6** ±1.1	**1.3** ±0.9	**0.5** ±1.2	**0.6** ±1.3	**3.7** ±0.9	**6.9** ±2.1	**0.8** ±1.7	**1.4** ±0.6
D234N	**0.1** ±1.1	**0.4** ±2.5	**0.9** ±1.1	**1.7** ±1.0	**0.2** ±1.3	**0.3** ±1.2	**0.5** ±1.2	**0.3** ±1.6	**2.7** ±0.8	**5.7** ±0.4	**0.6** ±1.1	**1.5** ±0.8	**0.7** ±1.2	**0.6** ±1.1	**3.4** ±1.0	**6.6** ±1.5	**0.8** ±1.5	**1.5** ±1.1
W284H	**0.1** ±1.3	**0.3** ±1.3	**0.9** ±1.3	**1.7** ±1.7	**0.2** ±1.3	**0.2** ±1.2	**0.6** ±1.1	**0.5** ±1.0	**3.5** ±1.1	**7.4** ±1.1	**0.7** ±1.1	**1.5** ±1.1	**0.9** ±1.2	**1.0** ±1.1	**6.2** ±0.8	**10.7** ±1.5	**1.0** ±1.5	**1.7** ±1.6

3) GH29, α-fucosidase AlfB

Time course analysis of disaccharide formation catalysed by AlfB-WT reveals peculiar behaviour (Fig 5). Disaccharide formation is slow and once maximal disaccharide yield is reached; it appears to plunge rapidly. Plotting transglycosylation rate versus donor substrate concentration ([α-L-Fuc*p*OCNP]) reveals that this untypical behaviour can be partly explained by the fact that at high (>8 mM) concentration the donor inhibits synthesis of the disaccharide transglycosylation product (Fig. S6A) and that at low donor concentration secondary hydrolysis rapidly leads to the decomposition of the disaccharide (Fig. S6B). In the case of AlfB-H80F, and most other AlfB mutants, substrate inhibition appears to be abolished and secondary hydrolysis is reduced.

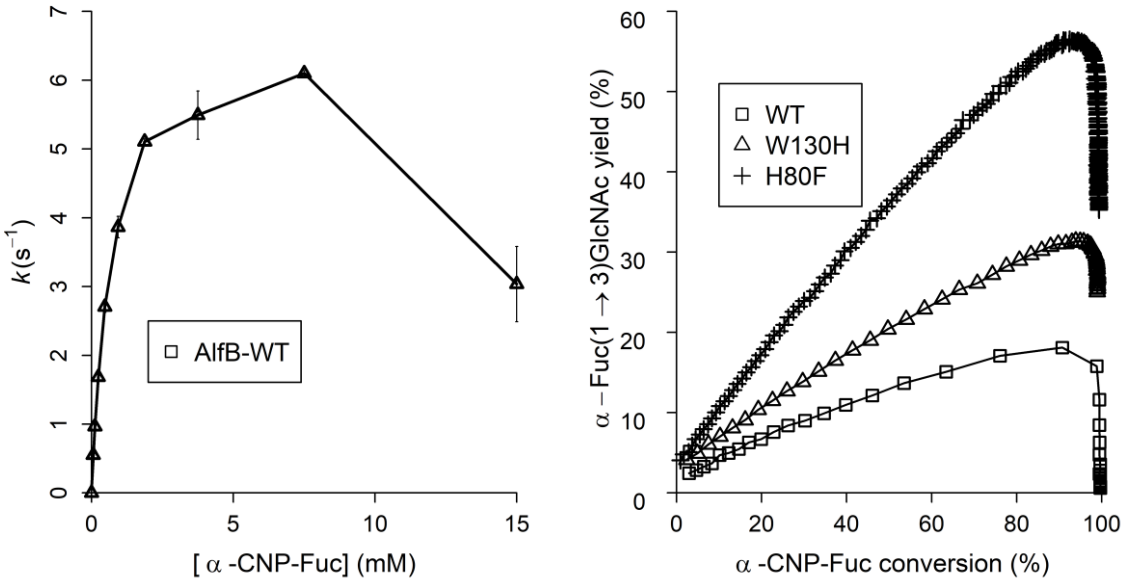


Figure S6 | AlfB behavior as a function of α-L-Fuc*p*OCNP concentration. Left, AlfB-WT hydrolysis kinetic profile assessed at 25 °C in 50 mM phosphate buffer pH 7, at 20 ng·mL^{-1}. Right, transglycosylation product evolution as a function of donor conversion.

NMR spectroscopy was used to assess the transfucosylation capabilities of AlfB forms (see Methods section). Not all anomeric signals could be followed due to the close vicinity with the HOD peak and its suppression, but signals for the donor, the acceptor and the disaccharide product could be monitored. Particularly, the latter was followed with 4 signals, one for both anomers of the disaccharide, two for its α-anomer and one for the β-anomer (δ = 5.06, 4.93, 4.9 and 4.06 ppm, respectively; Fig. S7).

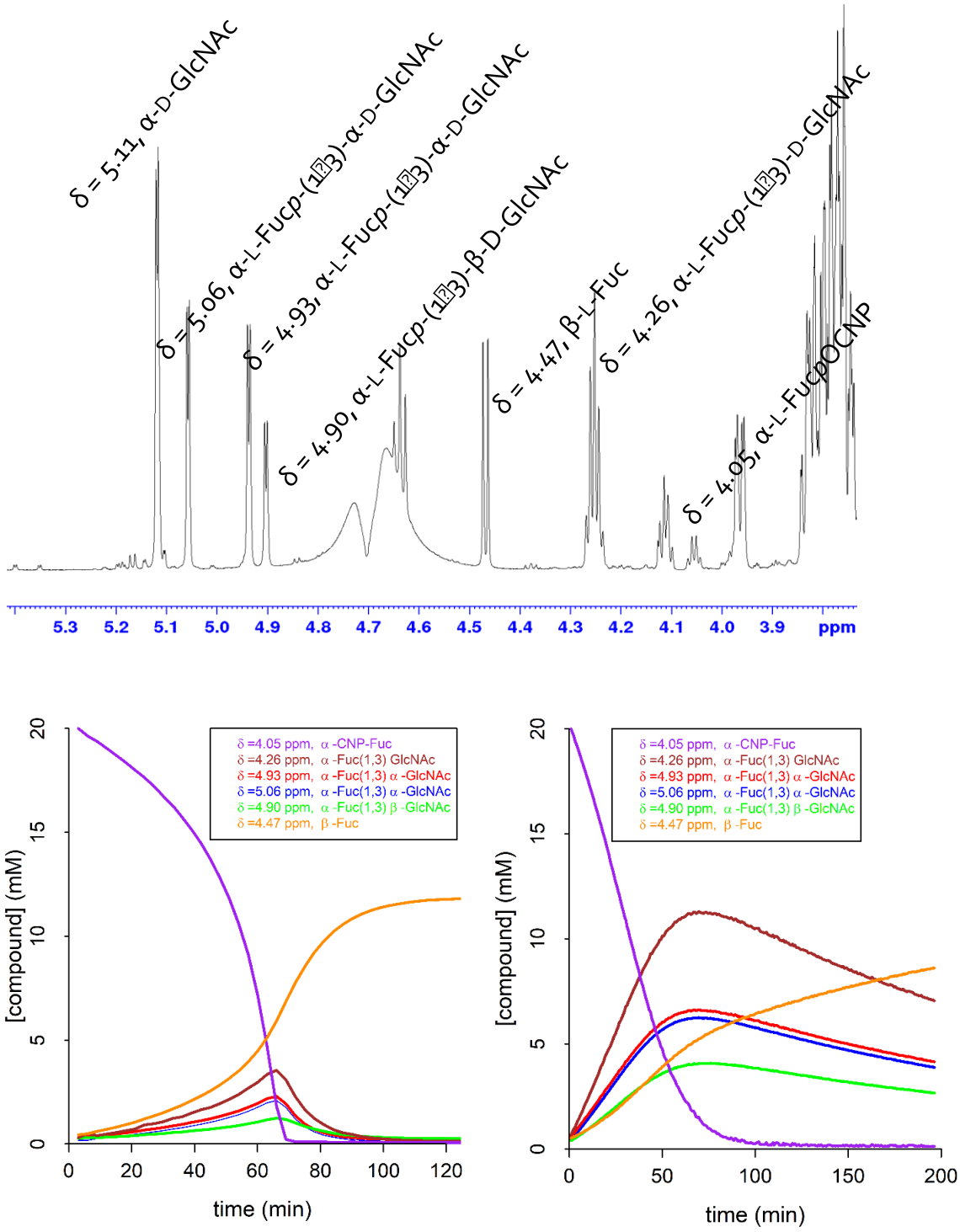


Figure S7 | NMR monitoring of transfucosylation by AlfB forms. Top, NMR signals used to monitor the reaction. Bottom Time-course reaction of AlfB-WT (left) and AlfB-H80F (right).

4) GH29, α-fucosidase AlfC

The feasibility of transposing a successful mutation to a related enzyme was assessed using the fucosidase AlfC from *Lactobacillus casei*. Signals from α-L-Fuc*p*-(1→6)-D-GlcNAc

were insufficiently resolved from those of α-L-Fuc*p*OCNP, L-Fuc and D-GlcNAc in the H-2 and anomeric region to allow accurate monitoring of yield (Fig. S7). Thus, compound formation and disappearance were monitored using the methyl group of the L-Fuc (Fig. S8). It is important to note that although progress curves established using the anomeric region were less precise, they were nevertheless consistent with those obtained when monitoring the methyl group (< 2% deviation for maximum yield assessment).

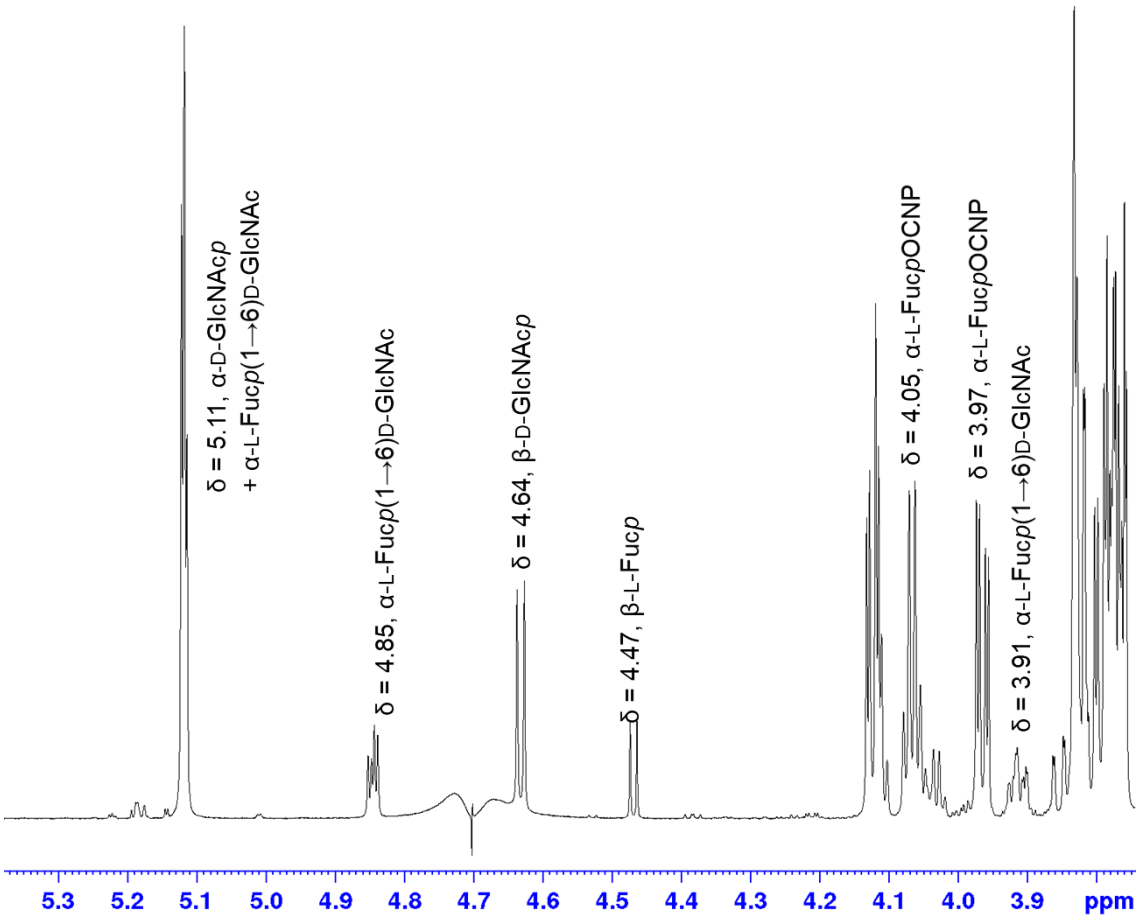


Figure S8 | NMR monitoring of transfucosylation by AlfC forms. NMR signals in the anomeric and H-2 regions.

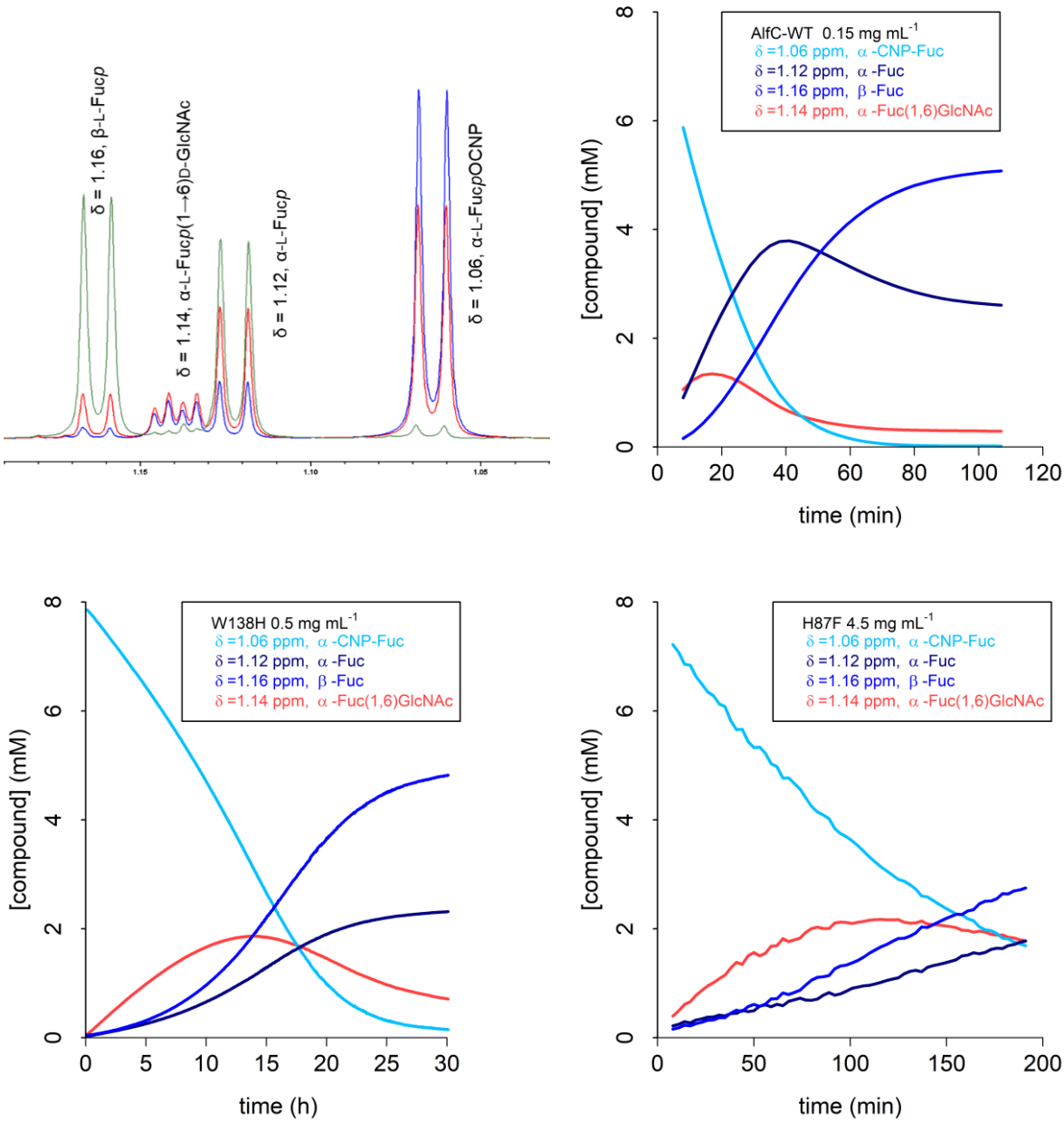


Figure S9 | NMR monitoring of transfucosylation by AlfC forms. Top left, NMR signals used to monitor the reaction (blue, after 8 min; red, 17 min; green, 53 min). Time-course reaction of AlfC-WT (top right), AlfC-W138H (bottom left) and AlfC-H87F (bottom right).

5) GH51, α-L-arabinofuranosidase *Tx*Abf

Table S4 | Maximum yields of arabinoxylo-tetrasaccharides synthetized by *Tx*Abf-WT and variants thereof.

***Tx*Abf form**	**Yield (%)**			
	XA^3X[a]	**A^3XX[a]**	**A^2XX (+ XA^2X)[a]**	**Total AXOS[b]**
	5.40 ppm[c]	5.32 ppm[c]	5.29-5.28 ppm[c]	
WT	2	4	4	9
F26H	8	5	11	22
F26L[d]	6	9	11	26
E28Q	5	8	10	23
R69K	10	13	16	37
R69H[d]	15	12	23	46
N175T	12	12	22	42
H240F	3	3	76	82
H240N	16	14	59	75
Y242F	4	4	12	19
D297N	12	15	23	47
N216W-H240F[d,e]	-	-	62[e]	62[e]
R69H-N216W-L352M[d,e]	-	-	70[e]	70[e]

[a]See Fauré et al. for a comprehensive description of AXOS nomenclature[10]. [b]Maximum yields of each AXOS were reached at different times explaining why total AXOS yield cannot be obtained by summing the individual maximum yields of each product. [c]NMR chemical shift of the anomeric proton of α-L-Ara*f* unit of AXOS reported in the literature at approximatively 25°C[11]. Displacement of ^{1}H chemical shifts for α-L-Ara*f* anomeric signal of each AXOS towards blinded region ($\Delta\delta$ = -0.08 ppm) is observed at 45°C. [d] Single mutants F26L[12] and R69H[11] were randomly generated and the triple mutant R69H-N216W-L352M[11] was obtained by recombination. [e]N216W procures highly regioselective synthesis of A^2XX tetrasaccharide[11].

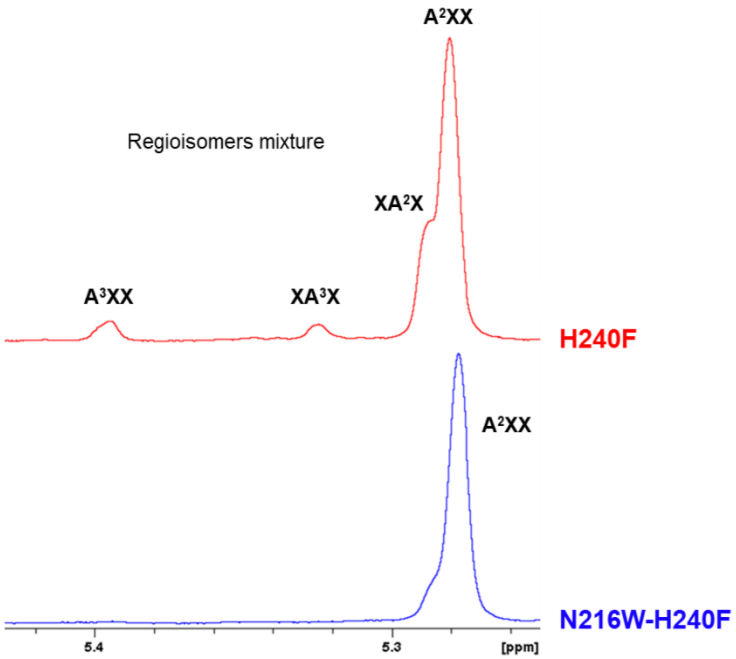


Figure S10 | NMR monitoring of transarabinofuranosylation by *Tx*Abf variants. ^{1}H NMR anomeric signals of α-L-Ara*f*-containing regioisomer(s) for H240F and regioselective N216W-H240F.

Table S5 | Specific activity (SA, in IU·mg^{-1}) of conserved residue mutants of *Tx*Abf determined by *p*NP release.

Enzyme	SA_H[a]	SA_T[b]	$R = SA_T/SA_H$
WT	261.79 ± 10.72	125.49 ± 1.27	0.5
F26H	0.02 ± 0.001	0.02 ± 0.001	1.0
F26L[c]	105.76 ± 2.72	83.26 ± 2.44	0.8
E28Q	52.95 ± 2.09	32.95 ± 0.44	0.6
R69K	14.12 ± 0.08	9.78 ± 0.36	0.7
R69H[c]	2.84 ± 0.07	7.41 ± 0.23	2.6
N175T	0.15 ± 0.003	0.14 ± 0.01	1.0
H240F	3.21 ± 0.09	3.66 ± 0.09	1.2
H240N	0.05 ± 0.001	0.19 ± 0.01	4.2
Y242F	0.07 ± 0.004	0.07 ± 0.01	1.0
D297N	7.71 ± 0.34	12.61 ± 0.22	1.7
N216W-H240F	5.49 ± 0.10	8.19 ± 0.21	1.5
R69H-N216W-L352M[c]	0.97 ± 0.11	1.82 ± 0.07	1.9

[a]SA_H was achieved in hydrolysis mode with 5 mM α-L-Ara*f*O*p*NP only. [b]SA_T was achieved in transglycosylation mode with 5 mM α-L-Ara*f*O*p*NP and in the presence of 10 mM xylotriose. [c]Single mutants F26L[12] and R69H[11] were randomly generated and triple mutant R69H-N216W-L352M[11] was obtained through recombination.

A B

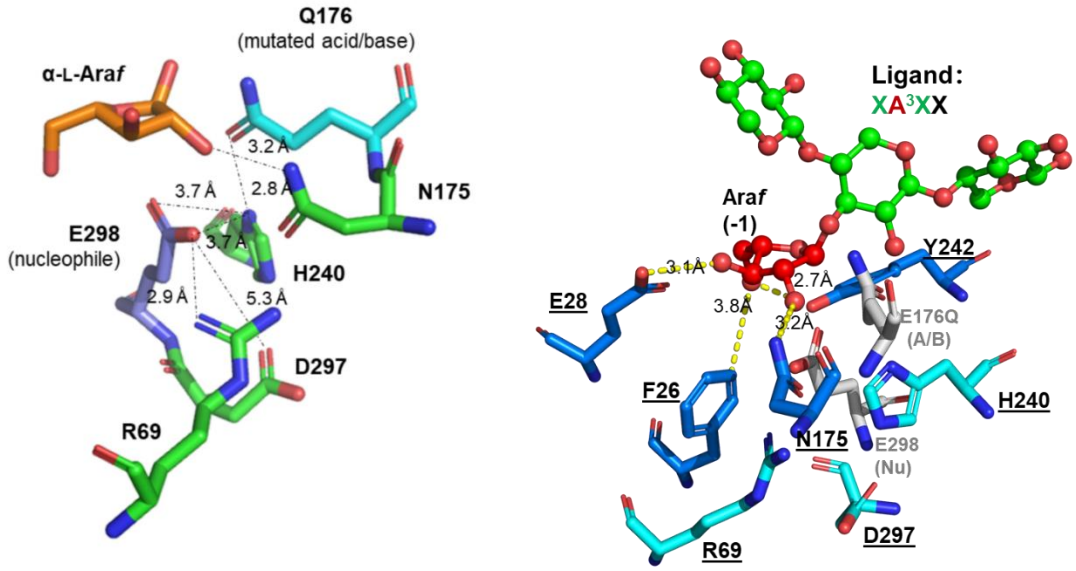


Figure S11 | (A) Location of R69, N175, H240 and D297 within *Tx*Abf-E176Q subsite –1 (PDB ID: 2VRQ). (B) Location of the 7 conserved residues (F26, E28, R69, N175, H240, Y242 and D297 within *Tx*Abf-E176Q subsite –1 (PDB ID: 2VRQ). The figure was prepared using PyMol Molecular Graphics System, v0.99 (Schrödinger).

Among the eight positions targeted by the conserved-sequence approach, the four best mutations are spatially (within approximately 5 Å) and/or sequentially close to the catalytic residues 176 (acid/base, in cyan) and E298 (nucleophile, in deep blue). Accordingly, mutations R69K, N175T, H240F/N and D297N (in green) are thought to impact the local H-bonding network and thus the pK_a cycling that occurs during catalysis.[13] Previously, the study of R69H revealed that R69 plays a vital role in the modulation of the ionization state and nucleophilic strength of E298.[11] Residue N175 is thought to be involved in transition state stabilization. Therefore, modification of the interaction N175···OH-2 of α-L-Ara*f* unit (in orange), for example in mutant N175T, might perturb the correct functioning of the two-step catalytic mechanism.[14] Additionally, H240 could be involved in a putative water channel affecting water dynamics.[15–17] Overall, the different mechanistic consequences of these four single-mutants all translate into reduced water-mediated deglycosylation and/or increased lifetime of the covalent glycosyl-enzyme intermediate, which in turn favours sugar-mediated deglycosylation. Therefore, all four single-mutants achieve the sought after result, namely tipping the T/H balance in favour of transglycosylation.

References

1. Ulrich, E. L. *et al.* BioMagResBank. *Nucleic Acids Res.* **36**, 402–408 (2008).

2. Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).

3. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).

4. Eneyskaya, E. V. *et al.* Transglycosylating and hydrolytic activities of the β-mannosidase from *Trichoderma reesei*. *Biochimie* **91**, 632–638 (2009).

5. Rosengren, A. *et al.* Enzymatic synthesis and polymerisation of β-mannosyl acrylates produced from renewable hemicellulosic glycans. *Green Chem.* **21**, 2104–2118 (2019).

6. Jansson, P.-E., Kenne, L. & Schweda, E. Nuclear magnetic resonance and conformational studies on monoacetylated methyl D-gluco- and D-galacto-pyranosides. *J. Chem. Soc. Perkins Trans.* **53**, 1689–1699 (1987).

7. Bubb, W. A. NMR spectroscopy in the study of carbohydrates: characterizing the structural complexity. *Concepts Magn. Reson. Part A Bridg. Educ. Res.* **19**, 1–19 (2003).

8. Harjunpää, V., Teleman, A., Siika-Aho, M. & Drakenberg, T. Kinetic and stereochemical studies of manno-oligosaccharide hydrolysis catalysed by β-mannanases from *Trichoderma Reesei*. *Eur. J. Biochem.* **234**, 278–283 (1995).

9. Agrawal, P. K. NMR Spectroscopy in the structural elucidation of oligosaccharides and glycosides. *Phytochemistry* **31**, 3307–3330 (1992).

10. Fauré, R. *et al.* A brief and informationally rich naming system for oligosaccharide motifs of heteroxylans found in plant cell walls. *Aust. J. Chem.* **62**, 533–537 (2009).

11. Bissaro, B. *et al.* Molecular design of non-Leloir furanose-transferring enzymes from an α-L-arabinofuranosidase: A rationale for the engineering of evolved transglycosylases. *ACS Catal.* **5**, 4598–4611 (2015).

12. Arab-Jaziri, F. *et al.* Engineering transglycosidase activity into a GH51 α-L-arabinofuranosidase. *New Biotechnol.* **30**, 536–544 (2013).

13. McIntosh, L. P. *et al.* The pK_a of the general acid/base carboxyl group of a glycosidase cycles during catalysis: a ^{13}C-NMR study of *Bacillus circulans* xylanase. *Biochemistry* **35**, 9958–9966 (1996).

14. Raich, L. *et al.* A trapped covalent intermediate of a glycoside hydrolase on the pathway to transglycosylation. Insights from experiments and quantum mechanics/molecular mechanics simulations. *J. Am. Chem. Soc.* **138**, 3325–3332 (2016).

15. Teze, D. *et al.* Conserved water molecules in family 1 glycosidases: a DXMS and molecular dynamics study. *Biochemistry* **52**, 5900–5910 (2013).

16. David, B. *et al.* Internal water dynamics control the transglycosylation/hydrolysis balance in the agarase (AgaD) of *Zobellia galactanivorans*. (2017).

17. David, B., Arnaud, P., Tellier, C. & Sanejouand, Y. Towards the design of efficient transglycosidases : the case of the GH1 of *Thermus thermophilus*. *Protein Eng. Des. Sel.* 1–8 (2019).

General conclusion, discussion and perspectives

Conclusion

This doctoral work aimed to contribute to the development of new tools for glycosynthetic purposes, specifically addressing the current paucity of viable solutions for the synthesis of furanosides. The choice of furanosides as targets is both fortuitous and meaningful. Fortuitous, because the key subject of this thesis is an α-L-arabinofuranosidase, an enzyme that over the years has proven to be an excellent model for the study of the T/H partition in retaining GHs. Meaningful, because furanosides are biologically-relevant being present in a variety of pathogenic microorganisms and of interest for the development of diagnostics and chemotherapeutics. Overall, the work described herein has achieved the aims that were initially set out and provides new knowledge that adds to our growing understanding of how rGHs can be converted into efficient TGs.

Developing *in vitro* chemo-enzymatic approaches using rGHs as the biocatalytic reservoir is a promising way to circumvent some of the shortcomings of conventional glycochemistry and the obstacles related to the use of Leloir glycosyltransferases. Nevertheless, to reach this goal it is vital to understand the underlying factors that determine the extent to which any given rGH is capable of performing transglycosylation and thus determine how the T/H partition can be tailored to suit requirements. In this regard, the comparison of mutants displaying different phenotypes, generated either through natural evolution or in the laboratory environment, is a convenient way to accumulate knowledge. Moreover, beyond the straightforward (but not simple) question of T/H modulation, converting an enzyme into a practical tool also requires answers to other questions related, for example, to ensuring maximum regioselective control. Accordingly, in this thesis work we have carefully studied and compared a series of mutants, some of which are the fruits of previous work, in order to extract as much information as possible on how specific mutations affect the T/H partition, how they affect the regioisomeric profile of the products and how they potentiate the use of the *Thermobacillus xylanilyticus* α-L-arabinofuranosidase (*Tx*Abf) for the catalysis of specific reactions. In this way, we throw light on some unexplored corners of the already extensively related *Tx*Abf story.

To further probe the molecular determinants that govern the T/H partition in *Tx*Abf, we targeted two subsite -1 residues, studying for the mutants F26L and L352M. For the study, these mutations were created in the background of R69H-N216W, a double mutant that already displays impressive transglycosylation capability. For our study, we deployed a series of techniques, including enzyme kinetics, 3D structure determination and molecular dynamic

simulations to clearly understand how L352M affects both the donor and the acceptor subsites, the latter effects being the result of a domino-like process that engenders increased loop flexibility and thus movements of two important acceptor binding residues. In contrast, we show that F26L, which is located in subsite -1 close to important residues such as E28, generates less modifications and actually results in a less efficient TG than the prototype R69H-N216W-L352M. Beyond this key difference, our results also reveal another putative alteration generated by both triple mutants that concerns the acid/base residue. Apparently, in the mutants this residue is more flexible, a property that appears to correlate with enhanced transglycosylation. Finally, this study once again stressed the determinant effect of R69H as a T/H modulator. In this regard, it is useful to note that our work reveals that R69 is part of a small group of 8 residues that are highly conserved in clan GH-A. Significantly, the mutation of any of the majority of these residues leads to enhanced transglycosylation.

Building on previous work that provided the excellent enzyme R69H-N216W, we investigated this mutant's ability to perform self-condensation, which is special form of transglycosylation. Accordingly, we confirm that this mutant's ability to procure oligo-L-arabino- and oligo-D-galactofuranosides is significantly better than that of the wild-type enzyme, and reveal that this is mainly due to significantly altered regioselectivity. Importantly, this work highlights the specific impact of the substitution N216W, since this procures an additional binding subsite that forms the basis for an alternative binding mode for acceptor substrate, which in turn yields three regioisomeric species including (1,2)-, (1,3)- and even (1,5)-linked difuranosides. An important finding in this study is that the creation of alternative acceptor binding does not abolish existing binding options. Therefore, this provides scope for future active site sculpting work.

Finally, the cumulative knowledge concerning the R69H-N216W mutant encouraged us to further explore the potentiality of this enzyme for the synthesis of biologically-relevant β-D-galactofuranosides. Significantly, this work revealed that it is possible to synthesize both β-D-Gal*f*-(1,3)-α-D-Glc*p* and Gal*f*-(1,m)-α-D-Glc*p*NAc. However, beyond the usefulness of this demonstration, the study also provided interesting data related to recognition of the D-Gal*f* moiety and how this is favorably affected by R69H-N216W that severely reduced K_M and K_d values on D-Gal*f* moiety. From a more fundamental standpoint this work thus reaffirmed the fact that destabilization of the transition state is important to move the T/H partition in favor of transglycosylation and thus TGs are generally catalytically-sluggish enzymes. However, it also underlines the fact that destabilization of the transition does not necessarily equate with

poor donor binding (*i.e* the case of the *Tx*Abf/β-D-Galf donor pair), but rather non-productive (with respect to generation of the transition state) binding.

Discussion and Perspectives

Perturbing the transition state is a prerequisite for enhanced transglycosylation

Most GHs are primarily hydrolases and some are extraordinarily potent catalysts that enhance the reaction rate up to 10^{17}-fold over the noncatalytic reaction rate.[1] In biological systems, GHs are often associated with catabolic or carbohydrate remodeling (involving bond breaking) functions. On the other hand, TGs are rarer and rather sluggish enzymes compared to GH counterparts, fulfilling important niche functions, such as the synthesis of extracellular polysaccharides in microorganisms. This observation has led to the suggestion that TGs might be evolutionary vestiges of ancestral enzymes that were less specialized and less potent than modern GHs. In this work we provide further evidence to support the postulate that the conversion of a rGH into a TG entails loss of catalytic efficiency (i.e. lower k_{cat}/K_M), which in turn is the consequence of transition state (TS) destabilization. According to current understanding, this is a necessary evil to establish a more even playing field for glycoside acceptors in a context where water holds an exceptionally high thermodynamic advantage.

Referring to the results presented herein, both F26 and R69 are sufficiently close to the nucleophile E298 to cause TS destabilization and, indeed, the mutation of either residue (F26L or R69H) procures enhanced transglycosylation activity. However, combining these two mutations is unproductive, since transglycosylation is lowered. Moreover, overall R69H is the better mutation regarding TS destabilization and combines well with mutations, such as L352M, which enhance further transglycosylation through the creation of favorable impacts in the acceptor subsites. Therefore, an obvious conclusion of our results is that TS destabilization can be achieved through different mutations. This conclusion in turn begs the question of what other mutants could be tested in the future? To answer this, it is relevant to recall that R69 is highly conserved in clan GH-A. Therefore, it would be interesting to search among other conserved residues. Previously, it was reported that Y242 plays a similar role to R69, participating in a hydrogen bond network with the nucleophile (E298) and stabilizing proton recycling during catalysis.[2,3] Encouragingly, we already know that Y242F improves transglycosylation, so it is mostly a matter of performing in depth kinetic studies to elucidate the exact nature of the impact of this mutant. Another target could be N175, since again this

residue is thought to be involved in a hydrogen network involving both catalytic residues and TS stabilization.[4]

Another interesting observation emerges from the comparison of the positional variants at amino acids 26 and 240 respectively, where mutations either globally maintain side-chain properties (*i.e.* F26L, which remains hydrophobic and H240N, which remains polar), or change them (*i.e.* F26H, and H240F, which represent hydrophobic/polar and polar/hydrophobic switches respectively). Changing the sidechain properties at these positions induces radical changes in activity, but actually has rather negligible effects on transglycosylation yields. Conversely, in the case of R69, its two variants, R69H and R69K, display relatively similar properties and transglycoslation yields. Presumably, this is because, unlike the previously cited positional variants, these mutations not only maintain sidechain polarity, but also the positive charge.

Globally, the mutation of conserved second-shell residues (e.g. R69, H240, D297) procures better results in terms of shifts in the T/H balance than mutations in the first-shell (e.g. F26, E28, N175 and Y242). This suggests that mutations in the first shell (*i.e.* residues that interact directly with the donor substrate) generate overly severe effects, whereas those in the second shell induce subtle, more effective impacts on the T/H partition. This conclusion is consistent with the results reported in Chapter 4 related to the mutation of first shell leucine residues (L30, L314 and L352) that are in close vicinity of D-Gal*f*? None of the substitutions at these positions improved global transgalactofuranosylation yield.

Overall, considering the findings of the current work, it will no doubt be useful in future research to focus on the deeper characterization of the single mutants described herein (*e.g.* E28Q, N175T, H240F/N Y242F and D297N). These can be also be combined with N216W, and possibly L352M, in order to assess transglycosylation performance.

Increasing transglycosylation when using donors that are structurally distant from a rGHs natural substrate

According to the peculiar saying “there is more than one way to skin a cat” there are often several ways to achieve something. Regarding TS destabilization, as we have seen this can be achieved using targeted mutagenesis to lower the value of k_{cat}/K_M,[3,5,6] which infers higher energy costs related to the glycosylation (and deglycosylation) transition state. However, as shown in the table below (Table 2), it is also possible to put more emphasis on raising the K_M

value, which can be achieved by using a donor substrate that is misaligned with the enzyme's natural specificity. The case of *Tx*Abf and β-D-Gal*f* is one example of this, while that of *Rhodococcus* sp. *endo*-glycoceramidase II (EGC) and CNP-β-cellobiose is another[7]. In these examples, the substrates are apparently poorly accommodated in the active sites of the respective enzymes. Accordingly, the previous results obtained with EGC and those presented here clearly reveal that further improvement of transglycosylation is actually achieved by improving substrate recognition (*i.e.* lowered K_d value), while moderately reducing k_{cat}. Indeed, such alterations appear to procure spectacular increases in transglycosylation. Therefore, as a general rule of thumb, we postulate that when one is dealing with a reaction involving a donor substrate that is structurally well-aligned with the natural specificity of a rGH, mutations to increase transglycosylation should target reductions in the value of k_{cat}/K_M.[5,8,9] However, in the case of poorly-recognized substrates, mutations to increase transglycosylation should first target reductions in the value of K_M, to improve substrate binding.

Table 2. Relationship of kinetics parameters in hydrolysis mode and T/H modulation

Enzyme	Donor	Acceptor	K_M (mM)	k_{cat} (s^{-1})	k_{cat}/K_M ($s^{-1}.mM^{-1}$)	Synthesis yield (%)
*Tx*Abf[3]	α-L-Ara*f*O*p*NP	Xylotriose	0.25	139	556	10[a]
*Tx*Abf R69H[3]			0.09	2.29	25.41	31[a]
*Tx*Abf R69H-L352M[3]			1.23	0.97	0.79	27[a]
*Tx*Abf R69H-N216W-L352M			0.48	0.58	1.21	80[b]
EGC[7]	CNP-β-Celb	5-hydroxypentan-2-one	22.4	3.7	0.17	11[b]
EGC D311Y[7]			7.0	2.9	0.42	75[b]
*Tx*Abf[10]	β-D-Gal*f*O*p*NP	α-D-Glc*p*O*p*NP	>50	15.02	0.13	20[c]
*Tx*Abf R69H-N216W			0.45	0.06	0.16	51[c]
*Tx*Abf R69H-N216W-L314N			2.82	0.27	0.10	60[c]
*Tx*Abf R69H-N216W-L352G			3.69	0.47	0.13	59[c]

[a]The yield refers to the main regioisomer, A^2XX.

[b]The yield refer to the synthesis of an unique cellobioside.

[c]The yield refers to the main regioisomer, β-D-Gal*f*-(1,3)-α-D-Glc*p*O*p*NP.

Alternative binding modes and docking to probe regioselectivity

Previous work (as yet unpublished) performed by our group led to the identification of an alternative binding mode in *Tx*Abf. This was observed for the reaction of β-D-Gal*f*O*p*NP (donor) with β-D-Xyl*p*-(1,4)-α-D-Xyl*p*OBn (acceptor), which procured the product benzyl β-D-Xyl*p*-(1,4)-[β-D-Gal*f*-(1,2)]-α-D-Xyl*p* (*i.e.* the β-D-Gal*f* is linked to the *O*-2 of α-D-Xyl*p* moiety) and involved occupation of a subsite +2' by the benzyl (Bn) group. Significantly, based on the results of our current work we predict that the reaction of β-D-Gal*f*O*p*NP and α-D-Glc*p*O*p*NP, which procures both β-D-Gal*f*-(1,2)-α-D-Glc*p*O*p*NP and β-D-Gal*f*-(1,3)-α-D-Glc*p*O*p*NP as products, involves two different binding modes (Fig. 1). Moreover, because the mutant R69H-N216W was more regioselective with respect to the (1,3)-linked β-Gal*f* product, we postulate that subsite +2 binding is improved and thus favored over subsite +2' binding. In future work, it will be very relevant to further investigate this using *in silico* docking,

constructing complexes of the different enzymes studied herein with different products. In this regard, it will be important to include the triple mutant R69H-N216W-L30G. So far, we did less work on this mutant because overall transglycosylation yields were lower than those of other mutants studied. However, it is significant that when catalyzing self-condensation of β-D-Gal*f*O*p*NP, R69H-N216W-L30G generated (1,2)-, (1,3)- and (1,5)-linked homo-disaccharides, with detection of the latter linkage being remarkable, because we have rarely observed this type of regioisomer in past studies using *Tx*Abf.

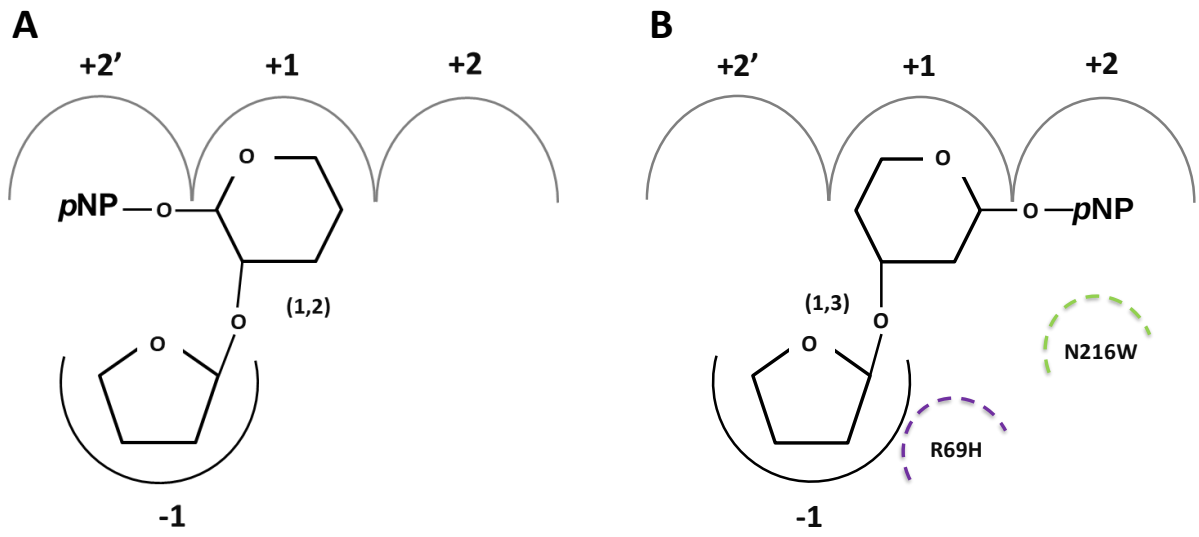


Fig. 1 Two proposed binding modes of α-D-Glc*p*O*p*NP. (A) Binding mode of β-D-Gal*f*-(1,2)-α-D-Glc*p*O*p*NP in wild-type *Tx*Abf; (B) Binding mode of β-D-Gal*f*-(1,3)-α-D-Glc*p*O*p*NP in R69H-N216W.

Towards the synthesis of other biologically-relevant glycoconjugates

We show that R69H-N216W is able to synthesize β-D-Gal*f*-(1,m)-α-D-Glc*p*NAcO*p*NP. Although uncertain at the time of writing, we presume that the linkage is a (1,4) bond. If so, this is gratifying, because this disaccharide glycomotif is present in *O*-linked oligosaccharide moieties found in a mucin-like structure of *Trypanosoma cruzi* (the causative agent of Chagas' disease). Despite, the potential impact of this result, it is important to note that the overall reaction yield with α-D-Glc*p*NAcO*p*NP (acceptor) is lower than that achieved when α-D-Glc*p*O*p*NP is the acceptor, and neither of the triple mutants R69H-N216W-L314N or R69H-N216W-L352G altered regioselectivity, nor improved transglycosylation yields. Moreover, these reactions were clearly more successful when the acceptors displayed α anomeric

configuration. Therefore, to complete this study and draw firmer conclusions, a docking study will be useful to investigate the binding mode of α-D-Glc*p*NAcO*p*NP in the acceptor subsites. Presumably the difference in reaction yields when using α-D-Glc*p*O*p*NP and α-D-Glc*p*NAcO*p*NP is explained by the presence of the bulky acetylated secondary amide at C-2, which logically alters accommodation of the sugar at subsite +1. Better understanding of how this occurs might provide clues that how to tailor subsite +1 using mutagenesis and thus improve binding of the α-D-Glc*p*NAcO*p*NP moiety and in turn increase yields of β-D-Gal*f*-(1,4)-α-D-Glc*p*NAcO*p*NP.

Scoping for and modifying water channels in the *Tx*Abf active site

Histidines have been pinpointed as possible determinants of an active site water channel in another clan GH-A enzyme. Interestingly, in our study we mutated H240 that is located in subsite -1. By homology, this histidine is also a putative water channel determinant that might affect water dynamics.[11–13] In *Tx*Abf, mutation of H240 (producing H240F and H240N) led to impressive increases in transglycosylation yield (>80% in the case of H240F compared to 9% for wt *Tx*Abf). In a recent study published by David *et al* water channels and internal water dynamics have been studied using in-house algorithms.[12] Moreover, an earlier study achieved this using Deuterium-Exchange Mass Spectroscopy (DEMS) and molecular dynamic simulations.[11] Therefore, it would appear to be relevant to our study to adopt some of these approaches in order to better characterize the role of H240 and perhaps discover other *Tx*Abf-water channel determinants.

References

1 R. Wolfenden, *Annu. Rev. Biochem.*, 2011, **80**, 645–667.

2 S. R. Marana, L. M. F. Mendonça, E. H. P. Andrade, W. R. Terra and C. Ferreira, *Eur. J. Biochem.*, 2003, **270**, 4866–4875.

3 B. Bissaro, J. Durand, A. Planas, P. Monsan, X. Biarnés, A. Planas, P. Monsan, M. J. O'Donohue, R. Fauré, *ACS Catal.*, 2015, **5**, 4598–4611.

4 L. Raich, V. Borodkin, W. Fang, J. Castro-López, D. M. F. van Aalten, R. Hurtado-Guerrero and C. Rovira, *J. Am. Chem. Soc.*, 2016, **138**, 3325–3332.

5 B. Bissaro, P. Monsan, R. Fauré and M. J. O'Donohue, *Biochem. J.*, 2015, **467**, 17–35.

6 D. Teze, F. Daligault, V. Ferrières, Y. H. Sanejouand and C. Tellier, *Glycobiology*, 2015, **25**, 420–427.

7 J. Durand, X. Biarnés, L. Watterlot, C. Bonzom, V. Borsenberger, A. Planas, S. Bozonnet, M. J. O'Donohue and R. Fauré, *ACS Catal.*, 2016, **6**, 8264–8275.

8 S. Luang, J. Il Cho, B. Mahong, R. Opassiri, T. Akiyama, K. Phasai, J. Komvongsa, N. Sasaki, Y. L. Hua, Y. Matsuba, Y. Ozeki, J. S. Jeon and J. R. Ketudat Cairns, *J. Biol. Chem.*, 2013, **288**, 10111–10123.

9 G. Potocki De Montalk, M. Remaud-Simeon, R. M. Willemot, P. Sarçabal, V. Planchot and P. Monsan, *FEBS Lett.*, 2000, **471**, 219–223.

10 R. Euzen, G. Lopez, C. Nugier-Chauvin, V. Ferrières, D. Plusquellec, C. Rémond and M. O'Donohue, *European J. Org. Chem.*, 2005, **2005**, 4860–4869.

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

12 B. David, R. Irague, D. Jouanneau, F. Daligault, M. Czjzek, Y.-H. Sanejouand and C. Tellier, *ACS Catal.*, 2017, **7**, 3357–3367.

13 B. David, P. Arnaud, C. Tellier and Y. Sanejouand, *Protein Eng. Des. Sel.*, 2019, 1–8.