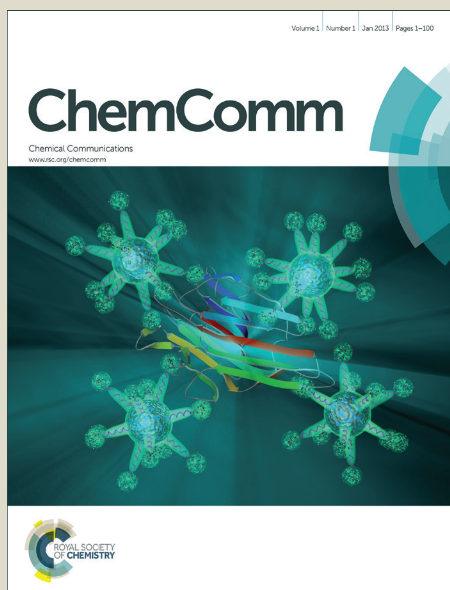
ChemComm

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

ChemComm RSCPublishing

COMMUNICATION

Engineering the Specificity of Trehalose Phosphorylase as a General Strategy for the Production of Glycosyl Phosphates

Cite this: DOI: 10.1039/x0xx00000x

Received 00th XX, 2014,
Accepted 00th XX, 2014

DOI: 10.1039/x0xx00000x

www.rsc.org/

Chao Chen,[a] Jef Van der Borght,[a] Rob De Vreese,[b] Matthias D'hooghe,[b] Wim Soetaert[a] and Tom Desmet*[a]

A two-step process is reported for the anomeric phosphorylation of galactose, using trehalose phosphorylase as biocatalyst. The monosaccharide enters this process as acceptor but can subsequently be released from the donor side, thanks to the non-reducing nature of the disaccharide intermediate. A key development was the creation of an optimized enzyme variant that displays a strict specificity (99%) for β-galactose 1-phosphate as product.

Glycosidic bonds can be found in a wide variety of biomolecules, ranging from food oligosaccharides and structural polysaccharides to medicinal glycosides and immunogenic glycoconjugates. Their synthesis is, however, far from trivial and typically requires the use of an activated glycosyl donor to maximize the yield.[1] Glycosyl phosphates, for example, have been successfully employed for the production of oligosaccharides and glycosides, comprising not only *O*- but also *C*- and *S*-glycosidic bonds.[2] Coupling of these donor substrates to a solid support has even enabled the development of automated procedures with increased throughput.[3] Unfortunately, chemical protocols for the supply of glycosyl phosphates suffer from low regio- and stereoselectivity,[4] explaining why few of them are commercially available.

Glycosyl phosphates can also be synthesized enzymatically, with either kinases or glycoside phosphorylases as biocatalysts. The use of kinases has been extensively explored,[5] but most of these enzymes phosphorylate carbohydrates at C6 rather than at C1. Furthermore, the need for ATP as an expensive phosphate source hampers large-scale applications, despite the development of recycling systems.[6] Alternatively, these products can be obtained by the phosphorolysis of cheap substrates like starch,[7] sucrose[8] or trehalose.[9] The main drawback of this approach is that glycoside phosphorylases only provide cost-effective access to glucose 1-phosphate and not to other glycosyl phosphates.[10] Although cellobiose phosphorylase has recently been engineered towards the production of α-galactose 1-phosphate from lactose,[11] there still is a strong need for a more broadly applicable technology.

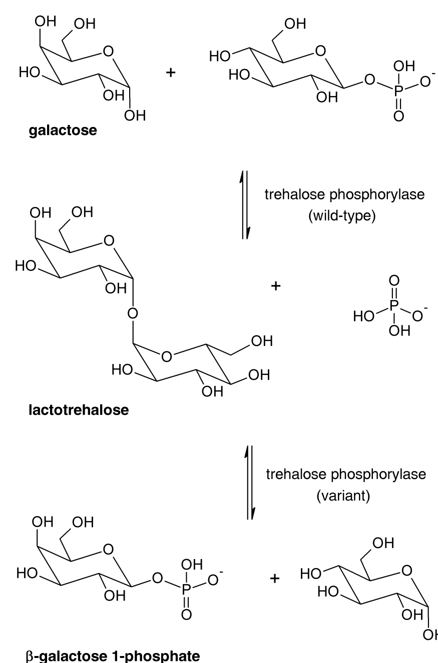


Fig. 1 Two-step conversion of galactose to β-galactose 1-phosphate with (variant) trehalose phosphorylases. Galactose can enter the process from the acceptor side but leave from the donor side, thanks to the non-reducing nature of the disaccharide intermediate.

Trehalose phosphorylase (TP) is a special biocatalyst because of the non-reducing nature of its substrate. Indeed, the disaccharide trehalose (α-D-glucopyranosyl α-D-glucopyranoside) has no directionality and thus still fits in the active site after being rotated 180°. Consequently, it should in principle be possible to convert its analogue lactotrehalose (α-D-galactopyranosyl α-D-glucopyranoside) to either D-glucose 1-phosphate (Glc-1P) or D-galactose 1-phosphate (Gal-1P), depending on the respective positioning of the carbohydrate moieties in the acceptor and donor subsites. Thanks to the rather broad acceptor specificity of inverting TP (EC 2.4.1.64),

ChemComm Accepted Manuscript

lactotrehalose can be synthesized from galactose and β-Glc-1P by exploiting the enzyme's reverse reaction.[12] The subsequent phosphorolysis of the product to β-Gal-1P would then amount to the net phosphorylation of galactose, with the phosphate group being transferred from β-Glc-1P in a two-step process (Fig. 1). For production purposes, however, the second enzyme will have to display a high preference for the release of β-Gal-1P over β-Glc-1P.

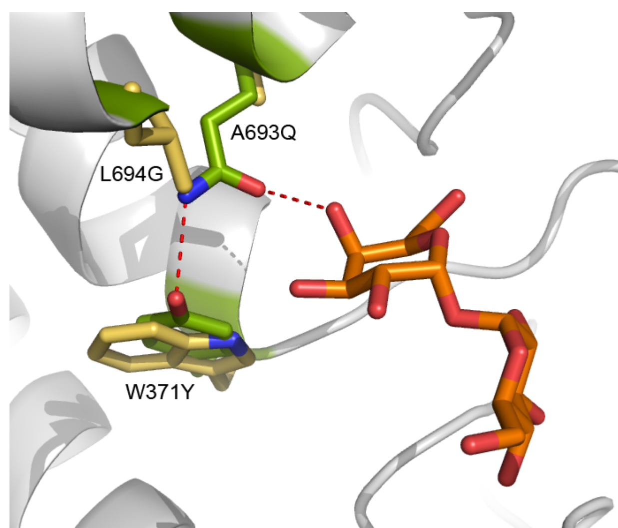


Fig. 2 View of the active site of TP. The structure is based on a homology model, in which the binding of lactotrehalose (orange) was simulated by automated docking. Positions that have been submitted to ISM are shown in yellow and green for the wild-type and variant enzymes, respectively.

To evaluate the proposed process, the reaction of the TP from *Caldanaerobacter subterraneus*[13] with 100 mM lactotrehalose and 100 mM P_i was analyzed by monitoring the release of glucose and galactose. These monosaccharides are formed in stoichiometric amounts to β-Gal-1P and β-Glc-1P, respectively, and can be conveniently quantified with colorimetric assays. Based on these measurements, the enzyme was found to have a 23-fold preference for the generation of galactose and β-Glc-1P over glucose and β-Gal-1P (Table 1). Although this ratio is unfavorable for the production of β-Gal-1P, the minor activity should provide a useful basis to optimize the specificity by means of enzyme engineering. To that end, iterative saturation mutagenesis (ISM) was applied, which is a very efficient yet simple strategy for the semi-rational design of enzyme properties.[14] Through homology modeling and ligand docking, three potential hotspots could be identified that are located close to the C4-OH of the carbohydrate moiety in the donor subsite (Fig. 2). These were used for the stepwise creation of single-site libraries, of which two 96-well microplates were always screened to ensure high coverage.[15] The ratio of β-Gal-1P formation over β-Glc-1P release was used as parameter to select the best variant, which was then purified and further characterized to confirm the screening results (Table 1).

In a first round, the substitution of L649 with glycine was found to improve the activity ratio with a factor 61. This is mainly caused by a decreased rate of β-Glc-1P release, which would imply that a specific interaction with glucose in the donor subsite has been weakened. The subsequent introduction of mutation A693Q further increased the activity ratio with a factor 9, this time by stimulating the release of β-Gal-1P. In our homology model, a new hydrogen bond is indeed predicted between the asparagine residue and the axial C4-OH of galactose (Fig. 2). Interestingly, this substitution seems to be possible only after space was created by mutation L649G, meaning that the order of the mutations was crucial in this case. Finally, the substitution of W371 with tyrosine resulted in a further 4-times improvement in preference for β-Gal-1P. This might be caused by a hydrogen bond between the newly introduced residues Y371 and Q693, which could stabilize the orientation of the latter (Fig. 2). In the end, a 2196-fold increase in the activity ratio of TP could thus be achieved, by screening of only about 600 variants. This result corresponds to a complete switch in specificity, with the production of β-Gal-1P being the minor activity (4%) in the wild-type enzyme but the major activity (99%) in the final variant.

The triple mutant was then applied for the production of β-Gal-1P starting from lactotrehalose and P_i. With equimolar concentrations of these substrates, the product yield was about 27% (at pH 7 and 60°C), which is consistent with what has been reported for the phosphorolysis of trehalose (K_{eq} = 0.122).[9] The thermodynamic equilibrium can, however, be shifted in the desired direction by the addition of excess amounts of inorganic phosphate. Using 100 mM lactotrehalose and 500 mM P_i, for example, 50% of the disaccharide substrate could be converted. After a simple purification procedure comprising precipitation and anion-exchange chromatography (ESI†), 3.8 g product was eventually obtained per liter reaction. Due to the enzyme's minor side activity, 1% of contaminating β-Glc-1P was still present in the product, as indicated by HPLC analysis (Fig. S1). The structure of the reaction product was also confirmed by NMR spectroscopy‡.

Table 1 Product preference of the TP variants obtained by ISM.[a]

Specific activity (U mg^{-1})	β-Gal-1P release	β-Glc-1P release	Ratio
wild-type	0.13 ± 0.01	3.01 ± 0.28	0.04
L649G	0.22 ± 0.05	0.09 ± 0.01	2.44
L649G/A693Q	1.92 ± 0.20	0.09 ± 0.01	21.3
L649G/A693Q/W371Y	4.71 ± 0.16	0.05 ± 0.01	93.8

[a] Purified enzymes were incubated with 0.1 M lactotrehalose and 0.1 M P_i at pH 7 and 60°C

With the new product in hand, a complete kinetic characterization of the enzyme could be performed. These data revealed that the variant's modified donor specificity is caused by changes in K_m as well as k_{cat} (Table 2). Indeed, a switch in affinity for the glycosyl phosphates is observed, while a significant decrease in activity on β-Glc-1P is also apparent. At the saturating concentrations that were used for the production process, however, β-Glc-1P still leads to a higher activity than β-Gal-1P. This would imply that effects in the acceptor subsite are also operative, which is indeed evident from our kinetic analysis (Table 2). In fact, the 5-times lower reaction rate with β-Gal-1P as donor is more than compensated by the 8-times higher reaction rate with glucose as acceptor, which could explain why these are the main products released from lactotrehalose. In the reverse reaction, 71 mM lactotrehalose could be generated from 100 mM β-Gal-1P and Glc, which is close to the theoretical maximum yield and demonstrates that the enzyme's synthetic capabilities are not hampered by the mutations. Finally, the stability of the variant enzyme was also evaluated, since mutations usually have a negative effect on a protein's structural integrity.[16] To that end, the temperature at

ChemComm Accepted Manuscript

which half of its activity is lost after one hour incubation (T_{50}) was determined and compared with that of the wild-type enzyme. That temperature was found to have decreased from 73°C to 70°C, which is only a modest change and does not hamper the applicability of the new biocatalyst.

Table 2 Donor and acceptor specificity of the wild-type and variant TP.[a]

	wild-type		triple mutant	
	k_{cat} (s^{-1})	K_m (mM)	k_{cat} (s^{-1})	K_m (mM)
β-Glc-1P	114 ± 8	0.7 ± 0.1	7.4 ± 0.6	3.1 ± 0.2
β-Gal-1P	1.2 ± 0.1	3.8 ± 0.3	1.6 ± 0.1	0.7 ± 0.1
Glc	86 ± 5	3.7 ± 0.2	4.2 ± 0.3	4.6 ± 0.4
Gal	113 ± 10	125 ± 12	0.5 ± 0.1	58 ± 3

[a] at pH 6 and 60°C, using 100 mM of Glc and 30 mM of β-Glc-1P as respective co-substrates

In conclusion, an optimized variant of the inverting TP from *C. subterraneus* has been created for the production of β-Gal-1P from lactotrehalose and inorganic phosphate. Interestingly, a similar strategy should be useful for the synthesis of other glycosyl phosphates, using trehalose analogues as intermediate products. Furthermore, access to the corresponding α-coupled glycosyl phosphates is also feasible, since TP enzymes with a retaining mechanism have been reported (EC 2.4.1.231).[17] We expect that the proposed concept will stimulate the production and use of glycosyl phosphates as cheap but efficient donor substrates for both chemical and enzymatic glycosylation reactions.

Financial support from UGent (MRP-project 'Ghent Bio-Economy'), the EC (FP7-project 'Novosides', grant 265854), the CSC (grant 2008679003) and the IWT (SB71279, SB73279) is gratefully acknowledged.

Notes and references

[a] Centre for Industrial Biotechnology and Biocatalysis Department of Biochemical and Microbial Technology, Ghent University, 9000 Ghent, Belgium. E-mail: tom.desmet@ugent.be
[b] SynBioC Research Group, Department of Sustainable Organic Chemistry and Technology, Ghent University, 9000 Ghent, Belgium

‡ The ^{1}H NMR data (chemical shifts and coupling constants) are consistent with the presence of a galactose moiety: ^{1}H NMR (300 MHz, D_2O): δ 3.38 (1H, dxd, J = 7.8, 10.1 Hz), 3.45-3.68 (4H, m), 3.72 (1H, d, J = 3.7 Hz); 4.68-4.71 (1H, m).

† Electronic supplementary information (ESI) available: Methods and supporting data. See DOI: 10.1039/x0xx00000x

1 (*a*) S. Hanessian and B. Lou, *Chem. Rev.*, 2000, **100**, 4443-4463; (*b*) K. C. Nicolaou and H. J. Mitchell, *Angew. Chem., Int. Ed.*, 2001, **40**, 1576-1624; (*c*) X. Zhu and R. R. Schmidt, *Angew. Chem., Int. Ed.*, 2009, **48**, 1900-1934; (*d*) T. Desmet, W. Soetaert, P. Bojarová, V. Kren, L. Dijkhuizen, V. Eastwick-Field and A. Schiller, *Chem. Eur. J.*, 2012, **18**, 10786-10801.
2 (*a*) K. R. Love, R. B. Andrade and P. H. Seeberger, *J. Org. Chem.*, 2001, **66**, 8165-8176; (*b*) E. R. Palmacci and P. H. Seeberger, *Org. Lett.*, 2001, **3**, 1547-1550; (*c*) O. J. Plante, E. R. Palmacci, R. B. Andrade and P. H. Seeberger, *J. Am. Chem. Soc.*, 2001, **123**, 9545-9554.
3 (*a*) D. K. Hunt and P. H. Seeberger, *Org. Lett.*, 2002, **4**, 2751-2754; (*b*) O. J. Plante, E. R. Palmacci and P. H. Seeberger, *Science*, 2001, **291**, 1523-1527.
4 (*a*) L. J. Edgar, S. Dasgupta and M. Nitz, *Org. Lett.*, 2012, **14**, 4226-4229; (*b*) O. J. Plante, R. B. Andrade and P. H. Seeberger, *Org. Lett.*, 1999, **1**, 211-214; (*c*) A. Ravida, X. Y. Liu, L. Kovacs and P. H. Seeberger, *Org. Lett.*, 2006, **8**, 1815-1818.
5 (*a*) D. Hoffmeister, J. Yang, L. Liu and J. S. Thorson, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 13184-13189; (*b*) L. Li, Y. Liu, W. Wang, J. Cheng, W. Zhao and P. Wang, *Carbohydr. Res.*, 2012, **355**, 35-39; (*c*) J. Yang, X. Fu, Q. Jia, J. Shen, J. B. Biggins, J. Jiang, J. Zhao, J. J. Schmidt, P. G. Wang and J. S. Thorson, *Org. Lett.*, 2003, **5**, 2223-2226.
6 (*a*) A. Kameda, T. Shiba, Y. Kawazoe, Y. Satoh, Y. Ihara, M. Munekata, K. Ishige and T. Noguchi, *J. Biosci. Bioeng.*, 2001, **91**, 557-563; (*b*) D. M. Kim and J. R. Swartz, *Biotechnol. Bioeng.*, 2001, **74**, 309-316.
7 R. S. Rathore, N. Garg, S. Garg and A. Kumar, *Crit. Rev. Biotechnol.*, 2009, **29**, 214-224.
8 K. De Winter, A. Cerdobbel, W. Soetaert and T. Desmet, *Process Biochem.*, 2011, **46**, 2074-2078.
9 J. Van der Borght, T. Desmet and W. Soetaert, *Biotechnol. J.*, 2010, **5**, 986-993.
10 (*a*) T. Desmet and W. Soetaert, *Process Biochem.*, 2012, **47**, 11-17; (*b*) C. Luley-Goedl and B. Nidetzky, *Biotechnol. J.*, 2010, **5**, 1324-1338.
11 (*a*) M. R. De Groeve, M. De Baere, L. Hoflack, T. Desmet, E. J. Vandamme and W. Soetaert, *Protein Eng. Des. Sel.*, 2009, **22**, 393-399; (*b*) M. R. De Groeve, V. Depreitere, T. Desmet and W. Soetaert, *Biotechnol. Lett.*, 2009, **31**, 1873-1877.
12 (*a*) H. Chaen, T. Nakada, N. Mukai and T. Nishimoto, *J. Appl. Glycosci.*, 2001, **48**, 135-137; (*b*) C. Chen, T. Desmet, J. Van der Borght, S. K. C. Lin and W. Soetaert, *Sep. Purif. Technol.*, 2012, **96**, 161-167; (*c*) J. Van der Borght, W. Soetaert and T. Desmet, *Biotechnol. Prog.*, 2012, **28**, 1257-1262.
13 J. Van der Borght, C. Chen, L. Hoflack, L. Van Renterghem, T. Desmet and W. Soetaert, *Appl. Environ. Microbiol.*, 2011, **77**, 6939-6944.
14 (*a*) A. S. Bommarius, J. K. Blum and M. J. Abrahamson, *Curr Opin Chem Biol*, 2011, **15**, 194-200; U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore and K. Robins, *Nature*, 2012, **485**, 185-194; (*b*) M. Reetz, *Angew. Chem., Int. Ed.*, 2011, **50**, 138-174.
15 M. T. Reetz, D. Kahakeaw and R. Lohmer, *ChemBioChem*, 2008, **9**, 1797-1804.
16 (*a*) J. D. Bloom, S. T. Labthavikul, C. R. Otey and F. H. Arnold, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 5869-5974; (*b*) N. Tokuriki, F. Stricher, L. Serrano and D. S. Tawfik, *PloS Comput. Biol.*, 2008, **4**, e1000002.
17 (*a*) C. Eis, M. Albert, K. Dax and B. Nidetzky, *FEBS Lett.*, 1998, **440**, 440-443; (*b*) C. Goedl, R. Griessler, A. Schwarz and B. Nidetzky, *Biochem. J.*, 2006, **397**, 491-500.

ChemComm Accepted Manuscript