TU Delft

Delft University of Technology

Transferases in Biocatalysis

Mestrom, L.

DOI
10.4233/uuid:e0fa631a-29a5-4aca-b245-dad04db03904

Publication date
2020

Document Version
Final published version

Citation (APA)
Mestrom, L. (2020). *Transferases in Biocatalysis*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:e0fa631a-29a5-4aca-b245-dad04db03904

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Transferases in Biocatalysis

Proefschrift

ter verkrijging van de graad van doctor

aan de Technische Universiteit Delft

op gezag van de Rector Magnificus Prof.dr.ir. T.H.J.J. van der Hagen,

voorzitter van het College voor Promoties,

in het openbaar te verdedigen op 18 september 2020 om 12:30 uur

door

Ir. Luuk MESTROM

Master of Science in Life Science and Technology, Technical University Delft, Delft, the Netherlands

Geboren in Sittard, the Netherlands

Dit proefschrift is goedgekeurd door de promotoren.

Samenstelling promotiecommissie bestaat uit:

Rector Magnificus,	voorzitter
Prof.dr. U. Hanefeld	Technische Universiteit Delft, promotor
Dr.ir. P.-L. Hagedoorn	Technische Universiteit Delft, copromotor
Dr. K. Szymańska	Silesian University of Technology

Onafhankelijke leden:

Prof.dr. F. Hollmann	Technische Universiteit Delft
Prof.dr. W.J.H. van Berkel	Wageningen University & Research
Prof.dr. G. van der Marel	Leiden University
Prof.dr. T. Desmet	Gent University
Prof.dr. W. R. Hagen	Delft University of Technology, reservelid

Sleutelwoorden:	biocatalysis, trehalose transferase, glycosyl transferase, immobilization, acyl transferase, carbohydrate
Drukkerij:	Gildeprint, Enschede
Boekomslag door:	Luuk Mestrom
ISBN:	9789464023886

http://repository.tudelft.nl/

All rights reserved. No parts of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photo-copying, recording, or otherwise, without the prior written permission of the author.

veur mien breurs

veur ôs Pap en Mam

)

Table of contents

Samenvatting

In de natuur zijn enzymen de katalysatoren voor de zeer efficiënte koppeling van koolhydraten. Enzymatische suikerkoppeling is een concurrerende technologie voor industriële glycosyleringsreacties, omdat chemische synthetische routes veelvuldig gebruik maken van bewerkelijke chemische modificaties, beschermende groepen vereisen en vaak regio- en stereoselectiviteit missen. In **hoofdstuk 2** worden de toepassingen van LeLoir glycosyltransferasen behandeld en hun uitstekende controle over de reactiviteit en selectiviteit van koolhydraten zonder beschermende groepen. De vooruitgang in de ontwikkeling van nucleotide-recyclingcascades heeft ervoor gezorgd dat de efficiënte productie en hergebruik van nucleotide suikerdonoren in robuuste een-pot multi-enzym glycosyleringscascades mogelijk zijn. Op deze manier kan de constructie van grote glycanen en glycoconjugaten met complexe stereochemie worden gerealiseerd. Met deze recente ontwikkelingen komt het gebruik van LeLoir glycosyltransferasen in multi-enzymatische, programmeerbare cascade suikerkoppelingsreacties voor industriële toepassingen dichterbij.

In **hoofdstuk 3** komt een LeLoir glycosyltransferase aan bod, genaamd trehalose transferase (TreT). TreT is van bijzonder belang omdat het de stereo- en enantioselectieve α,α-(1→1)-suikerkoppeling katalyseert van een nucleotide-suiker donor en een monosacharide acceptor voor de synthese van disacharide derivaten. Bij heterologe expressie van thermofiele trehalose transferasen bleken zij intrinsiek aggregatiegevoelig te zijn en komen ze hoofdzakelijk tot expressie als katalytisch actieve inclusion bodies in *Escherichia coli*. Om deze eiwitaggregatie tegen te gaan werd het thermostabiele eiwit mCherry uitgetest als een fluorescerend eiwitlabel. Na de fusie van trehalose transferase van *Thermoproteus uzoniensis* (*Tu*TreT) met mCherry toonde het enzym een verhoogde oplosbaarheid. Chaotrope componenten zoals guanidine of de tweewaardige kationen Mn(II), Ca(II), en Mg(II) verhoogden de enzymactiviteit van het fusie-eiwit.

In **hoofdstuk 4** vergelijken we de prestaties van een geïmmobiliseerd enzym met behulp van twee verschillende methoden: i) als drager-vrij katalytisch actieve inclusion bodies of ii) als drager-gebonden geïmmobiliseerd enzym. Om deze vergelijking te maken, gebruikten we een trehalose transferase van *Thermoproteus uzoniensis* gefuseerd met het fluorescerende zowel thermostabiele eiwit mCherry. Door de fusie van mCherry met trehalose transferase was de kwantificatie en visualisatie van het enzym in zowel natieve als genatureerde toestand mogelijk. De inclusion bodies presteerden beter dan het geïmmobiliseerde enzym door de eenvoud van productie van deze biokatalysator, wat resulteert in een hogere productiviteit van dit enzym. Enzym geïmmobiliseerd op dragermaterialen vertoonde een hogere katalytische activiteit en een robuustere prestatie onder batchproces condities.

Retaining LeLoir glycosyltransferasen katalyseren de vorming van glycosidische bindingen tussen nucleotidensuikerdonoren en koolhydraatacceptoren. De anomere selectiviteit van trehalose transferase van *Thermoproteus uzoniensis* werd onderzocht voor zowel D- als L-glycopyranose acceptoren in **hoofdstuk 5**. Het enzym koppelt een breed scala aan koolhydraten die trehalose analogen opleveren met uitstekende conversie en selectiviteit. De anomere selectiviteit keert om van α,α-(1→1)-glycosidebindingen voor D-glycopyranose acceptoren naar α,β-(1→1)-glycosidebindingen voor L-glycopyranose acceptoren, terwijl (*S*)-selectivteit behouden bleef voor beide soorten suikeracceptoren. Vergelijking van eiwitkristalstructuren van trehalose transferase in complex met α,α-trehalose en een onnatuurlijke α,ß-trehalose analoog benadrukte de mechanistische reden voor de waargenomen inversie van anomere selectiviteit.

Kinetische en thermodynamische controle voor transferase gekatalyseerde reacties is essentieel wanneer geactiveerde donoren worden gekoppeld aan acceptoren. Eén van de efficiënte acyltransferases, diegene van *Mycobacterium smegmatis,* katalyseert de verestering van een primaire alcohol in water als oplosmiddel. Bij gebruik van een zuur en alcohol als uitgangsmaterialen werden

lage opbrengsten waargenomen overeenkomstig met de wetten van de thermodynamica. Met geactiveerde esters, zoals ethylacetaat en vinylacetaat, kunnen zeer hoge opbrengsten van de gewenste ester worden bereikt in combinatie met de geschikte alcohol. In **hoofdstuk 6** hebben wij zowel de intrinsieke kinetische eigenschappen van MsAcT voor de hydrolyse en transesterificatie van esters in water onderzocht, als de thermodynamica van de reactie zelf. In vergelijking met de chemische of enzymatische estersynthese met behulp van een toxisch reagens en agressieve organische oplosmiddelen, kan de door MsAcT gekatalyseerde synthese van esters van primaire alcoholen efficiënt worden bereikt in water zonder neutralisatiestappen.

Summary

Enzymes are nature's catalyst of choice for the highly selective and efficient coupling of carbohydrates. Enzymatic sugar coupling is a competitive technology for industrial glycosylation reactions, since chemical synthetic routes require extensive use of laborious protection group manipulations and often lack regio- and stereoselectivity. In **chapter 2**, the application of Leloir glycosyltransferases and their excellent control over the reactivity and selectivity of glycosylation reactions with unprotected carbohydrates have been reviewed. The development of nucleotide recycling cascades has allowed for the efficient production and reuse of nucleotide sugar donors in robust one-pot multi-enzyme glycosylation cascades. In this way, the construction of large glycans and glycoconjugates with complex stereochemistry can be constructed. With recent advances, LeLoir glycosyltransferases are close to being applied industrially in multi-enzyme, programmable cascade glycosylations.

In **chapter 3**, the LeLoir glycosyltransferase trehalose transferase (TreT) has been described. It is of particular interest since it catalyzes the stereo- and enantioselective α,α-(1→1) coupling of a nucleotide sugar donor and monosaccharide acceptor for the synthesis of disaccharide derivatives. Heterologously expressed thermophilic trehalose transferases were found to be intrinsically aggregation prone and are mainly expressed as catalytically active inclusion bodies in *Escherichia coli*. To disfavor protein aggregation, the thermostable protein mCherry was explored as a fluorescent protein tag. The fusion of mCherry to trehalose transferase from *Thermoproteus uzoniensis* (*Tu*TreT) demonstrated increased protein solubility. Chaotropic agents like guanidine or the divalent cations Mn(II), Ca(II), and Mg(II) enhanced the enzyme activity of the fusion protein. The thermodynamic equilibrium constant, K_{eq}, for the reversible synthesis of trehalose from glucose and a nucleotide sugar was determined in both the synthesis and hydrolysis directions utilizing UDP-glucose and ADP-glucose, respectively.

In **chapter 4**, we compare the performance of an enzyme immobilised using two different methods: i) as carrier-free catalytically active inclusion bodies or ii) as carrier-attached immobilised enzyme. To make this comparison we used a trehalose transferase from *Thermoproteus uzoniensis* fused to the fluorescent thermostable protein mCherry. The fusion of mCherry to trehalose transferase allowed direct spectrophotometric quantification and visualisation of the enzyme in both native and denatured states. The catalytically active inclusion bodies outperformed the immobilised enzyme in their simplicity of biocatalyst production resulting in high enzyme productivity. Enzyme immobilised on carrier materials showed a higher catalytic activity and a more robust performance under batch process conditions.

Retaining LeLoir glycosyltransferases catalyze the formation of glycosidic bonds between nucleotide sugar donors and carbohydrate acceptors. The anomeric selectivity of trehalose transferase from *Thermoproteus uzoniensis* was investigated for both D- and L-glycopyranose acceptors in **chapter 5**. The enzyme couples a wide range of carbohydrates yielding trehalose analogues with excellent conversion and enantioselectivity. The anomeric selectivity inverts from α,α-(1→1)-glycosidic bonds for D-glycopyranose acceptors to α,ß-(1→1)-glycosidic bonds for L-glycopyranose acceptors, while (*S*)-selectivity was retained for both types of sugar acceptors. Comparison of protein crystal structures of trehalose transferase in complex with α,α-trehalose and an unnatural α,ß-trehalose analogue highlighted the mechanistic rationale for the observed inversion of anomeric selectivity.

Kinetic and thermodynamic control for transferase catalyzed reactions is essential when activated donors are coupled to acceptors. One efficient acyltransferase from *Mycobacterium smegmatis* catalyzes the esterification of primary alcohols in water. When utilizing acid and alcohol as starting materials low yields dictated by thermodynamics were observed. However, with activated esters such as ethyl acetate and vinyl acetate very high yields of the desired ester can be achieved in combination with the appropriate alcohol. In **chapter 6**, we investigated both the intrinsic kinetic properties of MsAcT for the hydrolysis and

transesterification of esters in water as well as the thermodynamics of the reaction. In comparison to the chemical or enzymatic ester synthesis using either toxic reagent, and harsh organic solvents, the MsAcT-catalyzed synthesis of esters of primary alcohols can be achieved efficiently in water without neutralization steps.

1

General introduction

“Niets is dwazer dan wijs te zijn op het verkeerde moment.”
Erasmus, *Het Lof der Sotheyt*, **1511**

1.1 Introduction

Catalysis is essential in all kingdoms of life for virtually every (bio)chemical reaction. One of the most prominent catalysts in nature are enzymes. The fundamental working principle of an enzyme is to reduce the amount of Gibbs free energy of activation of the catalyzed reaction in comparison the uncatalyzed chemical reaction by stabilizing the transition state complex(es).[1] Highly efficient rate enhancements of 10^8 – 10^{10}, even up to 10^{17}, make them exceed catalytic rates typically observed with (in)organic chemical catalysts.[2] For instance, the hydrolysis of the stable glycosidic bond has a half-live for spontaneous hydrolysis of nearly 5 million years. Enzymes hydrolyze such bonds with rate constants up to thousands per second.[3]

The high chemo-, regio-, and enantioselectivity of biocatalytic reactions renders the use of extensive protection groups unnecessary.[4] By reducing these additional reactions and intermediate purifications the *E-factor* can often be lowered, which is defined as the amount of waste generated per kilogram of product.[5] It is therefore not surprising that industrial biocatalytic processes are frequently awarded the Green Chemistry Award by the US Environmental Protection Agency.

Enzymes have been categorized in seven distinct Enzyme Commission (EC) classes. Transferases (EC 2) are one of the prominent enzyme classes which catalyze the transfer of one functional donor group to another acceptor compound. One of the hallmarks of transferases is their ability to perform highly selective reactions. These enzymes can be subclassified in ten distinctive classes (table 1) according to the group that is being transferred.

Table 1: Subclassification for transferases (EC 2)

EC number	Example	Transfer of group
2.1	Methyltransferase	methyl (one carbon)
2.2	Transketolase	aldehyde or ketone
2.3	Acyltransferase	acyl
2.4	Glycosyltransferase	glycosyl
2.5	Riboflavin synthase	alkyl or aryl (more than carbon)
2.6	Transaminase	nitrogenous
2.7	Kinase	phosphorus
2.8	Sulfotransferase	sulfurous
2.9	Selenotransferase	selenious
2.10	Molybdenumtransferase	molybdenum or tungsten

1.2 Transferases couple carbohydrates with high selectivity

Carbohydrates are the most abundant and diverse set of naturally occurring compounds in nature.[6] They are responsible for a wide range of biological functions, providing structure, enabling energy storage, and signaling. Their structural complexity requires the use of highly selective catalysts for transformations of functional groups. One of the most challenging aspects in the production of carbohydrates is the existence of multiple structural conformations for one single molecule. The Cremer-Pople (CP) puckering parameters describe the geometry of a pyranose ring for carbohydrates (Fig. 1a), being a chair (C), envelope (E), half-chair (H), skew (S), and boat (B).[7] The minimization of a number of stereo- and stereoelectronic interactions favor one of the configurations in solution, which is typically a $^4C_1$ chair configuration for D-glucopyranose (Fig. 1b).[8] Enzymes either alter the structural conformation to stabilize a transition state (i.e. oxocarbenium ion) or are highly selective for a single stereoconfiguration.

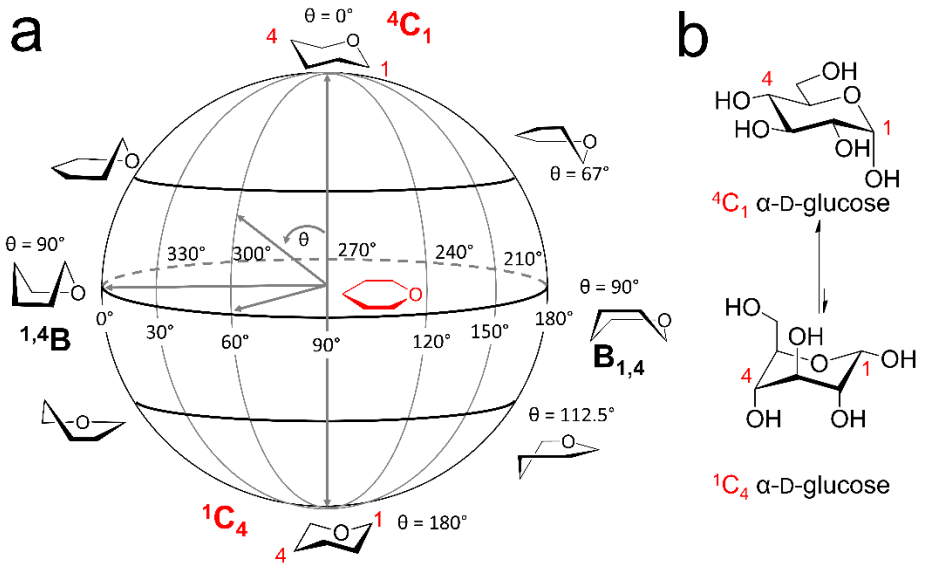
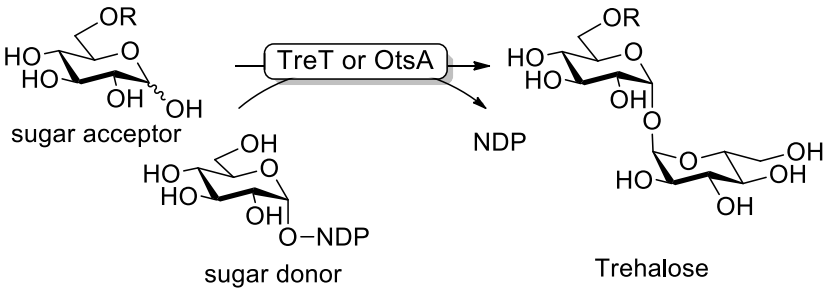


Figure 1: Cremer-Pople puckering of pyranose ring (**a**) and the preferred ring $^4C_1$ chair conformation of α-D-glucose (**b**).

One of the prominent classes of transferases for coupling carbohydrates are LeLoir glycosyltransferases (GT) (EC 2.4). These enzymes couple a sugar acceptor with an activated sugar nucleotide donor. LeLoir GTs are known for their broad substrate scope and high selectivity.[9] The main driving force for the reaction to go to completion is the exergonic release of activated nucleotide sugar donor, often resulting in high yields.[10]

Trehalose transferase (TreT) catalyse the formation of a (1→1)-glycosidic bond using a sugar acceptor and an activated nucleotide sugar donor (scheme 1).[11] Trehalose is a symmetrical non-reducing disaccharide containing a (1→1)-glycosidic bond of two α-D-glucopyranosides.[12] Its biological function ranges from energy storage to cellular protection as an osmolyte, and integral cell wall component of pathogenic bacteria like *Mycobacterium tuberculosis*.[13] Their wide natural function spurred the interest for synthesizing asymmetrical trehalose analogues.[14] Applications such as radio-imaging of pathogenic micro-organisms[15] or as food additives[16] have been explored. The main advantage of TreTs is their broad substrate scope for both the sugar donor and non-phosphorylated sugar acceptors. As phosphorylated sugar acceptors are often challenging to synthesise and expensive, we avoided to use of exclusively dependent 6-OH-phosphorylated sugar acceptors,[17] such as trehalose phosphate synthase (OtsA) from *E. coli*.[18]



Scheme 1: Enzymatic coupling of a sugar acceptor with OtsA (R = phosphate) or TreT (R=H) with a nucleotide diphosphate (NDP) sugar donor resulting in the formation of trehalose.

OtsA[18] and TreT[11d] have been proposed to proceed via a similar reaction mechanism. OtsA form *Escherichia coli* catalyses a 'same-face' nucleophilic substitution - S_Ni - (internal nucleophilic substitution) involving a transition state that is sufficiently open to allow the approach of the sugar acceptor and is guided by hydrogen bonding from the same face as the leaving group (Fig. 2).[18] The stable $^4C_1$ chair conformation in solution is not optimal for producing an oxocarbenium ion. The transition state that is stabilized by the transferase proceeds *via* a $^4H_3$ ring conformation, as the oxocarbenium ion-like character goes from sp^3 to sp^2 at C1 of the nucleotide sugar donor. It remains unknown if the reaction proceeds via a single transition state or *via* a concerted mechanism.[18]

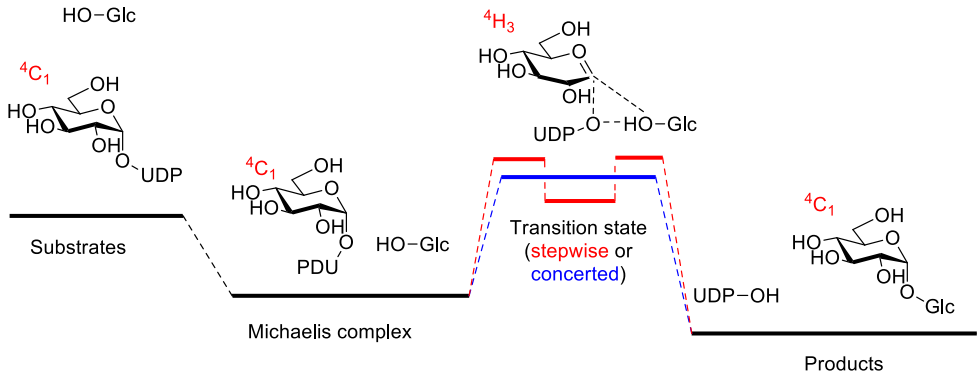


Figure 2: Thermodynamic scheme showing the reaction coordinate versus the energies of trehalose (phosphate) synthase catalyzed reactions. Binding of substrates often occurs in an ordered bi-bi fashion with the donor first, followed by the acceptor. The transition state is either concerted or a stepwise reaction is taking place with several separate transition states, both in a highly dissociative manner, where in both cases the nucleophile contributes to the stabilization of the intermediate. The transition state with the highest energy requirement is expected to form an oxocarbenium ion-like character altering pyranose conformation of $^4C_1$ to $^4H_3$ conformation.

1.3 What is the distinction between an (acyl)transferase and a hydrolase?

Transferases catalyze the transfer of an activated donor containing a sacrificial leaving group to an acceptor molecule (Fig. 3). As enzymes are dissolved in water abundant media a common side reaction is the undesired hydrolysis activity of an activated donor. Water acts as an acceptor hydrolyzing the activated donor molecule, which is the main activity of hydrolases. The distinction between a transferase and a hydrolase can be made by observing whether there is a preference for the transferase activity (Fig. 3). By measuring the initial transferase (k_1) to hydrolase (k_2) activity transferases can be classified, as they allow high maximum product yields under kinetic control to occur.[4]

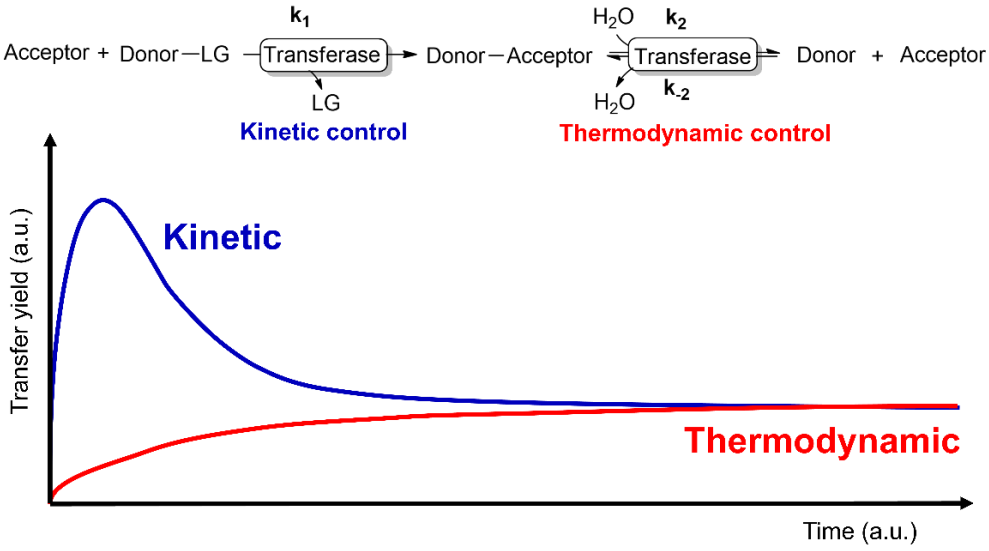
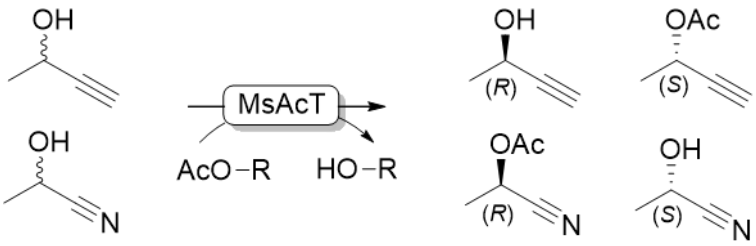


Figure 3: Separation between transferase and hydrolase activity using a donor with activated leaving group (LG) resulting in a maximum kinetic product yield or without an activated leaving group resulting in the thermodynamic product yield.

Enzyme engineering offers an opportunity to minimize hydrolysis and to increase the kinetic product yield, as it is an intrinsic property of the biocatalyst. If no activated leaving group is installed for the activated donor, the direct esterification of an acceptor and donor results in the thermodynamic product yield.

In such cases, hydrolases would result in the same observed yield and converges towards the thermodynamic equilibrium. Without activated donors, solely reaction engineering would offer a solution to increase the thermodynamic product yield. Lowering the thermodynamic activity of water for example can minimize reversible hydrolysis the donor.[19]

Acyltransferases (EC 2.3) are enzymes which catalyze the transfer of an activated acyl donor to a nucleophile, which is often an alcohol, amine, or thiol.[20] Acyltransferases are particularly useful for the (dynamic) kinetic resolution for the production of enantiopure amines and alcohols. According to the rule of Kazlauskas the preferential conversion of one of the enantiomers leads to a kinetic resolution, often designated as the *E ratio*.[21] In the case of acyltransferases, a common side reaction is the hydrolysis of the ester with water.[20] An acyltransferase from *Mycobacterium smegmatis* (MsAcT) has been applied for the synthesis of esters and amides in water abundant media.[22] It catalyzes the synthetic transesterification of activated acyl donors (i.e. ethyl acetate) with polar acceptor molecules (i.e. hydrogen peroxide) under kinetic control with high efficiencies.[23] When the kinetic resolution of (polar) acceptors in water with cyanohydrins and secondary alkynols with *Ms*AcT were evaluated, opposite (*R*)- and (S)-enantioselectivity was observed (scheme 2).[24] A computational study of MsAcT proposed that (*S*)-alkynols were bound *via* CH-π interactions, while cyanohydrins did not show this interaction.[25]



Scheme 2: Kinetic resolution of a (*S*)-cyanohydrin or (*R*)-alkynol via the transesterification of an activated acyl donor with MsAcT.

As hydrolases can also perform transesterification reactions it is the intrinsic properties of the catalyst which differentiate a hydrolase from an acyltransferase activity.[20] Within this thesis, we investigated the kinetic transient product yield of MsAcT and the thermodynamic equilibrium of the transesterification reaction. This study provides new insights in the efficiency of MsAcT and (trans)esterification reactions in water abundant media.

1.4 Scope of this thesis

The scope of this thesis is the application of transferases in applied biocatalysis. The first part is focused on the discovery, characterization, application of a trehalose transferase. The second part is focused on the acyltransferase from *Mycobacterium smegmatis.*

A review of recent literature using LeLoir glycosyltransferases (GTs) in applied biocatalysis is given in **chapter 2**. This chapter describes the fundamental insights combined with applied examples of LeLoir GTs for the development of nucleotide recycling cascades for the efficient production and reuse of nucleotide sugar donors in robust one-pot multi-enzyme glycosylation cascades. These advancements in combination with reaction and reactor engineering aspects highlights current knowledge and applications of LeLoir GTs.

In **chapter 3**, the biochemical characterization of a new trehalose transferases were investigated. Especially focus is directed on the fusion of a fluorescent protein mCherry to increase solubility and stability of TreT from *Thermoproteus uzoniensis*. The mCherry *Tu*TreT fusion protein allows the spectrophotometric quantification, while it also acts as a tool to evaluate protein denaturation and aggregation. The chapter also describes the discovery of the catalytic activity of inclusion bodies (IBs) of mCherry *Tu*TreT.

The successful application of enzymes in the production of complex food products and chemicals depends on the recyclability of the biocatalyst and its ease

of separation. We investigated the immobilization of mCherry *Tu*TreT in **chapter 4**, where the fluorescent protein mCherry is particular useful to observe the aggregation and denaturation processes during immobilisation. The aim of this study was to compare the performance with the carrier-attached mCherry *Tu*TreT to carrier-free catalytically active inclusion bodies (CatIBs) under batch process conditions.

Further investigation of applying mCherry *Tu*TreT for the production of trehalose analogues and its application in the enzymatic coupling of natural and unnatural sugar acceptors are described in **chapter 5**. The isolated trehalose analogues demonstrated opposite glycosidic linkages based on the stereochemistry of the sugar acceptor. The origin of the diastereoselectivity of mCherry *Tu*TreT was studied by evaluating the binding trehalose analogues within the protein crystal structure.

Chapter 6 investigates the fundamental difference between transferases and hydrolases for the enzymatic synthesis of esters in water. An extensive characterization of an acyltransferase from *Mycobacterium smegmatis* was performed to investigate both the intrinsic kinetic properties of *Ms*AcT for the hydrolysis and transesterification reactions as well as the thermodynamics of the reactions.

References

[1] K. Faber, in *Biotransformations in Organic Chemistry: A Textbook* (Ed.: K. Faber), Springer International Publishing, Cham, **2018**, pp. 1-30.

[2] R. Wolfenden, M. J. Snider, *Accounts of Chemical Research* **2001**, *34*, 938-945.

[3] D. L. Zechel, S. G. Withers, *Accounts of Chemical Research* **2000**, *33*, 11-18.

[4] S. R. Marsden, L. Mestrom, D. G. G. McMillan, U. Hanefeld, *ChemCatChem* **2020**, *12*, 426-437.

[5] I. A. R. A. Sheldon, Ulf Hanefeld, in *Green Chemistry and Catalysis*, WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim, **2007**, pp. 1-47.

[6] J. M. Berg, Tymoczko, J.L., Stryer, L., in *Biochemistry*, 6th ed., W. H. Freeman and Company, New York, **2006**, p. 303.

[7] D. Cremer, J. A. Pople, *Journal of the American Chemical Society* **1975**, *97*, 1354-1358.

[8] H. B. Mayes, L. J. Broadbelt, G. T. Beckham, *Journal of the American Chemical Society* **2014**, *136*, 1008-1022.

[9] B. Nidetzky, A. Gutmann, C. Zhong, *ACS Catalysis* **2018**, *8*, 6283-6300.

[10] A. Gutmann, B. Nidetzky, *Advanced Synthesis & Catalysis* **2016**, *358*, 3600-3609.

[11] a) Q. Qu, S.-J. Lee, W. Boos, *Journal of Biological Chemistry* **2004**, *279*, 47890-47897; b) T. Kouril, M. Zaparty, J. Marrero, H. Brinkmann, B. Siebers, *Archives of Microbiology* **2008**, *190*, 355; c) A. Nobre, S. Alarico, C. Fernandes, N. Empadinhas, M. S. da Costa, *Journal of Bacteriology* **2008**, *190*, 7939-7946; d) E.-J. Woo, S.-I. Ryu, H.-N. Song, T.-Y. Jung, S.-M. Yeon, H.-A. Lee, B. C. Park, K.-H. Park, S.-B. Lee, *Journal of Molecular Biology* **2010**, *404*, 247-259; e) S.-I. Ryu, J.-E. Kim, E.-J. Kim, S.-K. Chung, S.-B. Lee, *Process Biochemistry* **2011**, *46*, 128-134.

[12] L. Cordone, G. Cottone, S. Giuffrida, *Journal of Physics: Condensed Matter* **2007**, *19*, 205110.

[13] a) L. O. Martins, R. Huber, H. Huber, K. O. Stetter, M. S. Da Costa, H. Santos, *Applied and Environmental Microbiology* **1997**, *63*, 896-902; b) L. K. Conlin, H. C. M. Nelson, *Molecular and cellular biology* **2007**, *27*, 1505-1515; c) O. Fernandez, L. Béthencourt, A. Quero, R. S. Sangwan, C. Clément, *Trends in Plant Science* **2010**, *15*, 409-417.

[14] S. Ohtake, Y. J. Wang, *Journal of Pharmaceutical Sciences* **2011**, *100*, 2020-2053.

[15] a) B. L. Urbanek, D. C. Wing, K. S. Haislop, C. J. Hamel, R. Kalscheuer, P. J. Woodruff, B. M. Swarts, *ChemBioChem* **2014**, *15*, 2066-2070; b) S. Peña-Zalbidea, A. Y. T. Huang, H. W. Kavunja, B. Salinas, M. Desco, C. Drake, P. J. Woodruff, J. J. Vaquero, B. M. Swarts, *Carbohydrate Research* **2019**, *472*, 16-22; c) J. M. Groenevelt, L. M. Meints, A. I. Stothard, A. W. Poston, T. J. Fiolek, D. H. Finocchietti, V. M. Mulholand, P. J. Woodruff, B. M. Swarts, *The Journal of Organic Chemistry* **2018**, *83*, 8662-8667.

[16] N. D. Danielson, J. Collins, A. I. Stothard, Q. Q. Dong, K. Kalera, P. J. Woodruff, B. J. DeBosch, R. A. Britton, B. M. Swarts, *Chemical Communications* **2019**, *55*, 5009-5012.

[17] D. Tischler, S. Niescher, S. R. Kaschabek, M. Schlömann, *FEMS Microbiology Letters* **2013**, *342*, 113-122.

[18] S. S. Lee, S. Y. Hong, J. C. Errey, A. Izumi, G. J. Davies, B. G. Davis, *Nature Chemical Biology* **2011**, *7*, 631-638.

[19] a) I. Svensson, E. Wehtje, P. Adlercreutz, B. Mattiasson, *Biotechnology and Bioengineering* **1994**, *44*, 549-556; b) R. H. Valivety, P. J. Halling, A. R. Macrae, *FEBS Letters* **1992**, *301*, 258-260.

[20] M. Subileau, A. H. Jan, J. Drone, C. Rutyna, V. Perrier, E. Dubreucq, *Catalysis Science & Technology* **2017**, *7*, 2566-2578.

[21] R. J. Kazlauskas, A. N. E. Weissfloch, A. T. Rappaport, L. A. Cuccia, *The Journal of Organic Chemistry* **1991**, *56*, 2656-2665.

[22] a) I. Mathews, M. Soltis, M. Saldajeno, G. Ganshaw, R. Sala, W. Weyler, M. A. Cervin, G. Whited, R. Bott, *Biochemistry* **2007**, *46*, 8969-8979; b) I. Chiarelli Perdomo, M. Letizia Contente, A. Pinto, D. Romano, P. Fernandes, F. Molinari, *Flavour and Fragrance Journal*, *n/a*; c) M. L. Contente, A. Pinto, F. Molinari, F. Paradisi, *Advanced Synthesis & Catalysis* **2018**, *360*, 4814-4819.

[23] a) A. Drożdż, U. Hanefeld, K. Szymańska, A. Jarzębski, A. Chrobok, *Catalysis Communications* **2016**, *81*, 37-40; b) L. Wiermans, S. Hofzumahaus, C. Schotten, L. Weigand, M. Schallmey, A. Schallmey, P. Domínguez de María, *ChemCatChem* **2013**, *5*, 3719-3724.

[24] N. de Leeuw, G. Torrelo, C. Bisterfeld, V. Resch, L. Mestrom, E. Straulino, L. van der Weel, U. Hanefeld, *Advanced Synthesis & Catalysis* **2018**, *360*, 242-249.

[25] M. Kazemi, X. Sheng, F. Himo, *Chemistry – A European Journal* **2019**, *25*, 11945-11954.

(this page has been left intentionally blank)

2

Leloir Glycosyltransferases in Applied Biocatalysis: A Multidisciplinary Approach

This chapter is based on

Luuk Mestrom, Marta Musiol, Daria Kowalczykiewicz, André Pollender, Antje Kumpf, Stefan R. Marsden, Isabel Bento, Andrzej B. Jarzębski, Katarzyna Szymańska, Arkadiusz Chruściel, Dirk Tischler, Rob Schoevaart, Peter-Leon Hagedoorn, Ulf Hanefeld. *Int. J. Mol. Sci.*, **2019**, *20*(21), 5263, doi: *https://doi.org/10.3390/ijms20215263*

"Vanaf vandaag niks mier te vreeze
vanaf vandaag wet ik vort dat;
't neet erg is iets moeis te verleeze
beater verleeze dan dat ge 't noeit het gehad"
Rowwen Hèze*, "Heilige Antonius", Water, lucht en liefde,* **1997**

2.1 Introduction

Enzymes were already used for the conversion of glycosides even before all stereochemical details of the known carbohydrates were assigned[1]. In 1837, a crude formulation of almonds containing hydroxynitrile lyases catalyzed the enzymatic hydrolysis of the glycoside amygdalin[2]. Moving almost two centuries forward, the largest volumetric biocatalytic industrial process is the application of glucose isomerase for the production of high fructose syrup for food and drink applications, producing fructose from glucose at 10^7 tons per year[3]. The secret of the success of enzymes in the production or treatment of carbohydrates and glycosides is their exquisite stereo- and regioselectivity. The excellent selectivity of enzymes is required due to the diversity of structural features of carbohydrates[4], comprising D- and L-epimers, ring size, anomeric configuration, linkages, branching, and oxidation state(s). Since drug targets often exhibit specificity for all of these structural features, the production process should not contain any side-products to prevent undesired side-effects[5].

The challenge in the synthesis of carbohydrates is their wide variety of functionalities and stereochemistry (Figure 1). (Poly)hydroxyaldehydes containing a terminal aldehyde are referred to as aldoses and (poly)hydroxyketones are defined as ketoses. In aqueous solutions, monosaccharides form equilibrium mixtures of linear open-chain and ring-closed 5- or 6 membered furanoses or pyranoses, respectively. For aldoses, the asymmetric ring forms at C-1. For ketoses, it closes at C-2 with an axial (α) or equatorial (β) hemiacetal or hemiketal, respectively (commonly defined as the anomeric center). A glycosidic linkage is a covalent *O*-, *S*-, *N*-, or *C*-bond connecting a monosaccharide to another residue resulting in a glycoside, while glucoside is specific for a glucose moiety. The equatorial or axial position of the glycosidic bond is referred to as *α*- (axial) or *β*-linkage (equatorial). The number of carbohydrates linked via glycosidic bonds can be subdivided into oligosaccharides with two to ten linked carbohydrates, while polysaccharides (glycans) contain more than ten glycosidic bonds. A glycan either contains multiple different monosaccharides or more than ten glycosidic bonds. A glycoconjugate contains at least one or more monosaccharides or oligosaccharides covalently attached to a non-carbohydrate moiety (aglycon). If an oligosaccharide contains an aldose or ketose that is in equilibrium with its open-chain form, the aldehyde or

ketone can be oxidized with chemical reagents (e.g., with the Benedict reagent). This is referred to as the reducing end in oligosaccharides. If there is no possibility for the sugar to form the open chain-form, then this is called a non-reducing end. Non-reducing sugars are found in glycoconjugates (i.e. nucleotides) and oligosaccharides (i.e., raffinose).

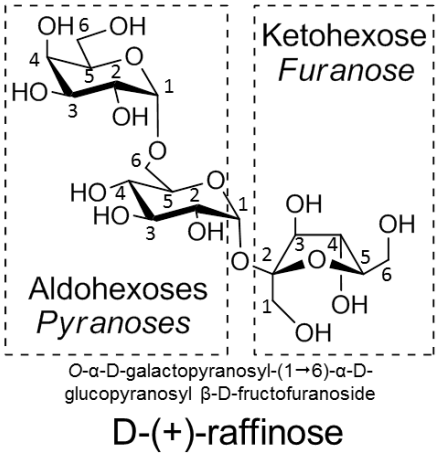
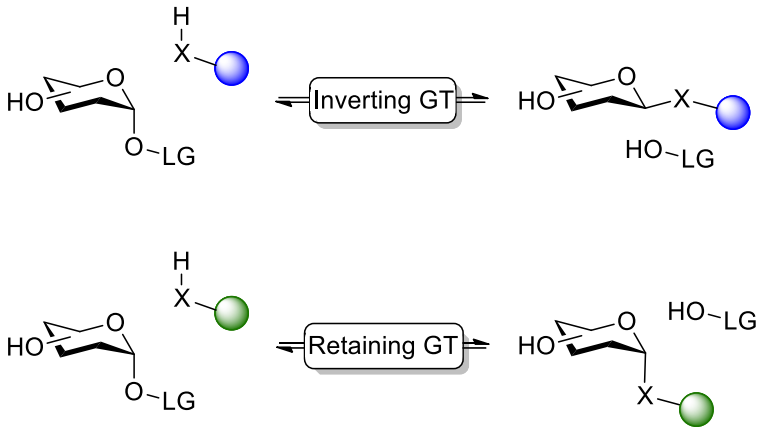


Figure 1. The nomenclature of glycosides and oligosaccharides.

Glycosyltransferases (GTs) catalyze the transfer of a carbohydrate acceptor from an activated sugar nucleotide donor with high selectivity and yield, enabling the stereo- and regioselective extension and branching of large glycans and glycoconjugates (Scheme 1). Upon formation of the glycosidic bond, the stereochemistry can either be retained or inverted by GTs with high selectivity for the α- or β-anomer. Leloir glycosyltransferases utilize carbohydrates linked to a nucleotide diphosphate (NDP) with an α-linked glycosidic bond, where non-Leloir glycosyl transferase utilize a phosphorylated sugar donor. For both types of glycosyltransferase, the main driving force for the reaction to go to completion is the exergonic release of either P_i or NDP from their respective sugar donors. The choice of nucleotide acceptor determines the (stereo) chemical outcome of the type of *O*-, *NH*-, *S*-, *C*-glycosidic bonds.



X = O, NH, S **LG** = NDP for LeLoir GTs, P_i for non-LeLoir GTs

Scheme 1. The overall scheme of an enzymatic glycosylation reaction for the biocatalytic synthesis of glycosides by retaining or inverting glycosyltransferases (GT) using NDP or P_i activated sugar donors for Leloir and non-Leloir GTs, respectively.

The enzymatic treatment of glycosides is mainly applied in the food industry using non-LeLoir GTs, enhancing flavors and functionality in complex food formulations, such as debittering[6], sweetening[7], or clarification[8]. The high costs of nucleotides, enzymes, and (enzymatic) regeneration systems for the treatment or production of low-value carbohydrate-containing products limited the application of nucleotide-dependent LeLoir GTs within the industry in the past. However, recent advancements in glycobiology have sparked interest in the (chemo) enzymatic production of high-value glycosides and glycoconjugates with high yield and selectivity for pharmaceutical applications[9]. As more LeLoir GTs are being reported with high protein expression, wide substrate scope, and high selectivity, industrial enzymatic glycosylation for the production of glycosides and glycoconjugates *in vitro* is becoming economically feasible. For instance, the expression of a large part of the human glycosyltransferases is a new hallmark for the production of human glycans or glycoconjugates[10], simplifying their chemoenzymatic synthesis. Besides these developments, the reaction methodologies are currently being further optimized. Multi-step enzymatic coupling with glycosyltransferases using non-natural sugar acceptors and nucleotide sugar

donors have been performed with automated synthesizers and are under development, as was recently reviewed [11].

The enzymatic synthesis of glycosides has received increasing attention in organic synthesis. However, the application of Leloir glycosyltransferases in a multi-enzymatic sugar coupling process is challenging from a process design point of view. The high costs, low stability, and difficult or limited availability of nucleotide sugar donors, in addition to the challenging protein production of Leloir glycosyltransferases hamper the development of enzymatic glycosylation. As a compromise, separate nucleotide sugar regeneration cascades and optimization of the protein production of industrial biocatalysts has been pursued[12]. Although there is a large body of scientific literature reporting on the biochemical properties and the reactions that glycosyl transferring enzymes catalyze, the performance of these biocatalytic processes has only sparingly been described. Due to their inherent complexity, kinetic and thermodynamic parameters have often not been analyzed in detail for the production of larger oligosaccharides using Leloir glycosyltransferases. In this review, the possibilities and limitations for industrial applications of Leloir glycosyltransferases are highlighted from the intersection of biochemical, chemical, thermodynamic, and reaction engineering perspectives, giving an overview of the requirements of industrial processes involving glycosyltransferases.

2.2 Glycosyltransferases in nature

Glycosyltransferases catalyze the formation of a glycosidic bond between an unactivated acceptor monosaccharide or aglycon and an activated sugar donor[13] to a di-, oligo-, polysaccharide[14], lipo(poly)saccharide[15] or peptidoglycan[16]. More than 484,620 glycosyltransferases in over 106 families have been identified according to the carbohydrate active enzymes (CAZy) database under the Enzyme Commission number E.C.2.4.x.y. (CAZy database, last updated 01/15/18[17]), representing an enormous number of metabolic pathways[18]. Glycosyltransferases can be sub-classified based on four different criteria: (i) the class of substrates[19]; (ii) the protein structure[13]; (iii) the preference in stereochemistry[14, 20]; (iv) the dependency on metals for catalytic activity[13]. Non-Leloir glycosyltransferases use phosphorylated donors (i.e., lipid polyprenol[21],

sugar 1-phosphates[19]) and can be described as phosphorylases. The second class are transglycosidases accepting non-activated di- or polysaccharides as carbohydrate donors. The largest class of glycosyltransferases are the nucleotide-dependent Leloir glycosyltransferases[13, 19, 22], named in honor of Luis Federico Leloir, who received a Nobel prize for the discovery of nucleotide sugar donors in 1970 (Figure 2).

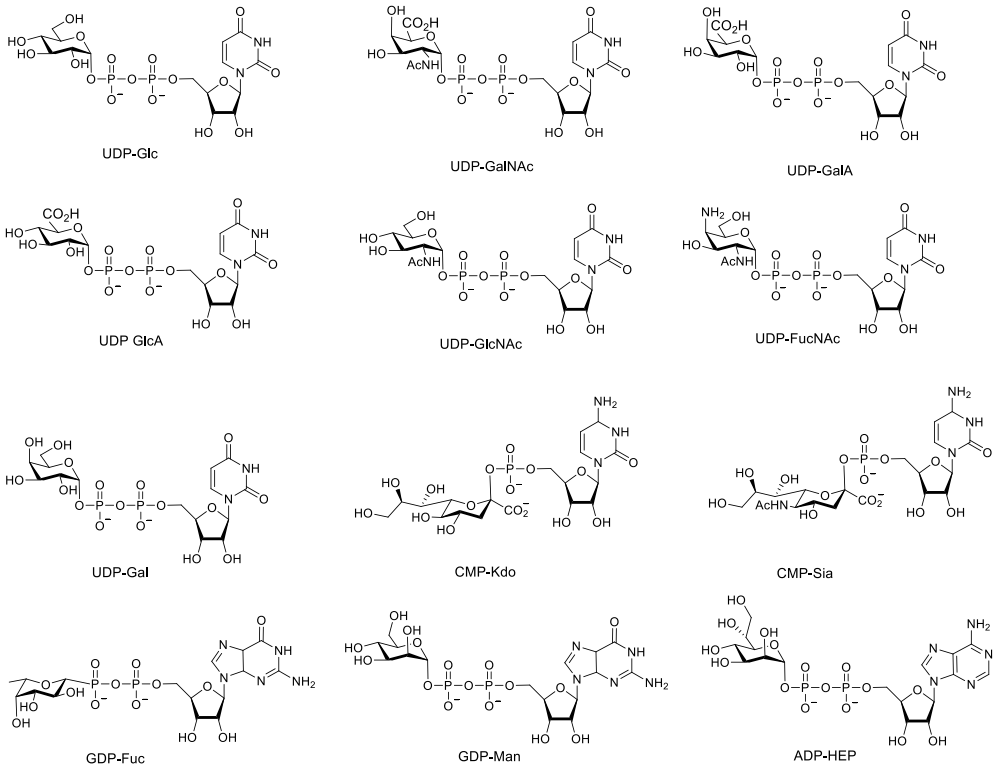


Figure 2. Common sugar nucleotides found in all kingdoms of life. Abbreviations: UDP-Glc, UDP-glucose; UDP-GalNAc, UDP-*N*-acetyl-2-deoxy-D-galactosamine; UDP-GalA, UDP-D-galacturonic acid; UDP-GlcA, UDP-D-glucuronic acid; UDP-GlcNAc, UDP-*N*-acetyl-2-deoxy-D-glucosamine; UDP-FucNAc, UDP-*N*-acetyl-L-fucosamine; UDP-Gal, UDP-D-galactose; CMP-KDO, CMP-3-deoxy-D-manno-octulosonate; CMP-Sia, CMP-*N*-acetylneuraminic acid; GDP-Fuc, GDP-L-fucose; GDP-Man, GDP-D-mannose; ADP-HEP, ADP-L-*glycero*-D-*manno*-heptose.

The protein sequence and crystallographic data demonstrate that glycosyltransferases are mainly comprise of five different protein folds (Figure 3)[13-14, 23]. Glycosyltransferases having a GT-A or GT-B fold consist of two β/α/β-Rossmann-like domains, abutting each other in case of the GT-A fold or facing each other for GT-B folds[13-14, 24]. Both folds contain separate donor and acceptor binding sites[13]. Gloster *et al*. reported that glycosyltransferases with a GT-A fold belong to the divalent metal ion dependent class of these enzymes, whereas GT-B folds are often metal ion independent[13-14]. Interestingly, glycosyltransferases having a GT-C fold are non-Leloir glycosyltransferases, utilizing membrane integrated or membrane linked proteins with lipid phosphate sugar donors, also known as non-Leloir donors[14, 19, 23]. The Leloir glycosyltransferases containing a GT-D fold catalyze the transfer of glucose to hexasaccharide *O*-linked to serine-rich repeats of bacterial adhesins[25]. The most recent addition, is the *N*-acetyl-D-mannose transferase utilizing non-Leloir undecaprenyl-linked glycosyl diphosphates with a unique GT-E fold[26].

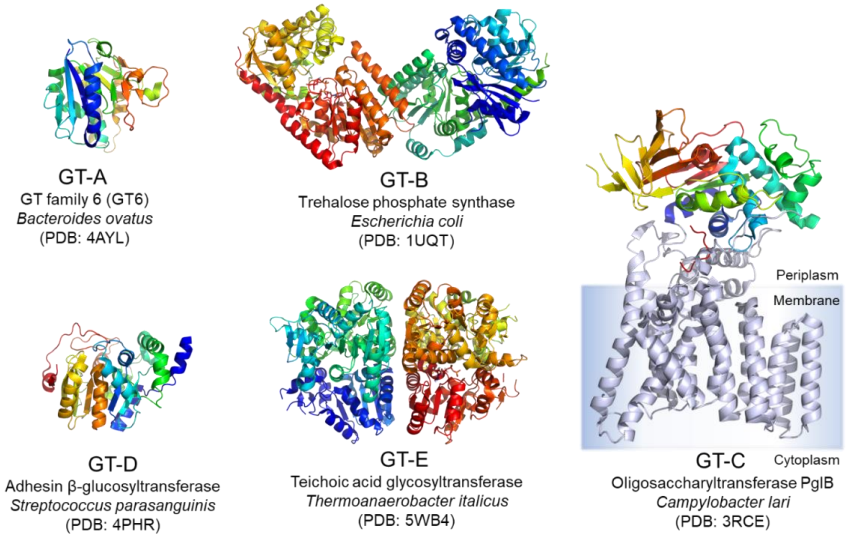


Figure 3. Protein folds of Leloir glycosyltransferases (GT-A, GT-B, GT-D) and non-Leloir glycosyltransferases (GT-C, GT-E).

For Leloir glycosyltransferases, the binding of the sugar donor nucleotide and acceptor follows a sequential ordered bi-bi catalytic mechanism via non-covalent interactions of the sugar donor nucleotide. The binding of the sugar or aglycone acceptor results in an enzyme-substrate ternary complex[27]. Hydrolysis of the sugar nucleotide donor is prevented by the tight binding in an unproductive state, where the high affinity of the enzyme for the sugar nucleotide donor is an indicator for product inhibition (K_i) by the released nucleotide[28]. For Leloir glycosyltransferases, a lower affinity or promiscuity towards the nucleotide donor results often in less product inhibition[29]. Upon binding of the sugar nucleotide donor, the enzyme undergoes a conformational change stabilizing the transition state, resulting in the formation of a glycosidic bond and the release of the nucleotide donor. Different reaction mechanisms of glycosyltransferases have been described and reviewed[13-14, 30]. The *inverting* transfer occurs via a S_N2 mechanism, while a *retaining* transfer can proceed via a concerted or ion-pair intermediate mechanism through a double displacement via a S_N2 mechanism. Also, a *transient covalent* intermediate via a S_Ni-type mechanism has been described for LeLoir GTs (Figure 4). *Inverting* glycosyltransferases use general base catalysis (i.e., aspartate or glutamate)[18, 31] to form an oxocarbenium ion-like transition state. They show a catalytic rate enhancement by utilizing divalent metals (i.e. Mn (II) or Mg (II)), which are often coordinated by the amino acid motif Asp-X-Asp.

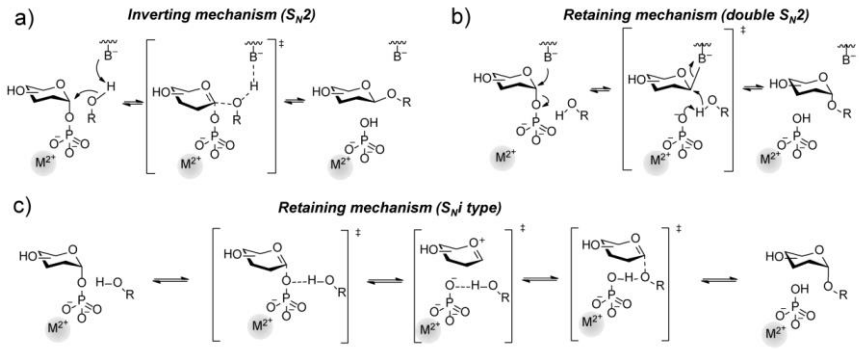


Figure 4. Reaction mechanism of glycosyltransferases upon inversion (**a**) or retention (**b**, **c**) of the anomeric glycosidic bond. The divalent metal (M^{2+}) is not necessarily a requirement for catalytic activity for GTs.

2.2.1. Distinguishing Glycosyl Transferases from Glycoside Hydrolases

Two main groups of enzymes can catalyze the regio-, stereo-, and enantioselective coupling of carbohydrates. *Glycoside hydrolases* and *glycosyltransferases* are often combined in biocatalytic retro-synthetic strategies for linear elongation and branching of oligosaccharides. Glycoside hydrolases are enzymes that condense a sugar donor with an aglycone acceptor. The broad substrate scope of glycoside hydrolases has resulted in numerous synthetic applications such as synthesis[32] or hydrolysis[33] of glycosidic bonds, and desymmetrization[34]. As a drawback, their broad substrate scope also leads to the formation of side-products. Glycosylations with glycoside hydrolases are under kinetic (transglycosylation) or thermodynamic control (direct glycosylation) using activated and non-activated sugars respectively (Figure 5). With *transglycosylation*, relatively high yields can be obtained in comparison with *direct glycosylation* due to a thermodynamically unfavorable reaction equilibrium (K_{eq}) in water. As a rule of thumb, transglycosylation should be faster than glycoside hydrolysis, as otherwise the activated sugars would hydrolyze before the glycosylation reaction is completed. Also, the rate of hydrolysis of the product should be slower than the activated glycosyl donor or the product yield decreases. As this is often not the case, an excess of the activated sugar donor is required under kinetic control. Similar to the coupling of protected glycosyl donors, the donors for transglycosylation, such as fluoro[35], -azido[36], *p*-nitrophenyl-[37] or *p*-nitropyridyl-[38], vinyl-[39], and allyl-glycosides[40] require their separate synthesis. The direct glycosylation is challenging due to the poor K_{eq} under aqueous reaction conditions, limiting the degree of conversion. The product yields with direct glycosylation can be improved by adding one substrate in excess, lowering the water activity[41], and *in situ* product removal[42].

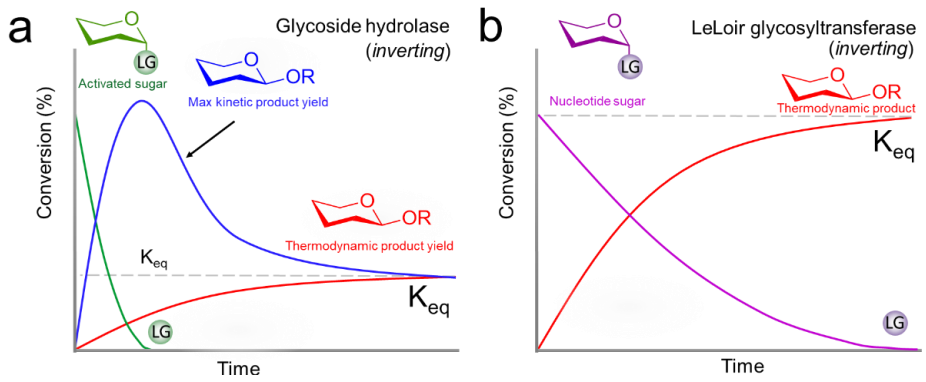


Figure 5. Exemplary enzymatic glycosylation of an activated sugar donor (green) and acceptor (R-group) to afford a maximum transient kinetic (blue) product yield catalyzed by a glycoside hydrolase, followed by reverse hydrolysis towards the thermodynamic product concentration. Direct esterification leads to the thermodynamic product yield K_{eq} without the requirement for an activated sugar (red) in (**a**). LeLoir GTs only catalyze the direct esterification of a nucleotide sugar donor (purple) to thermodynamic product (red) in (**b**).

Leloir glycosyltransferases couple NDP sugar donors with a wide range of sugar acceptors resulting in the formation of a glycosidic bond. The exclusion of hydrolysis activity of the nucleotide sugar donor separates glycoside hydrolases from glycosyltransferases. Nevertheless, hydrolysis of the nucleotide sugar donor in the absence of a sugar acceptor has been reported and is referred to as “error hydrolysis”[43]. Hence, the competition between water or a sugar acceptor as nucleophile is important for the efficiency of glycosylation. Only a handful of studies investigated the nature of the hydrolysis activity of Leloir glycosyltransferases with sugar nucleotide donors. For instance, the bacterial sialyltransferase from *Pasteurella dagmatis* hydrolyzed the rather hydrolysis-prone CMP-Neu5Ac in the absence of another substrate[44]. Directed evolution has been shown to be an effective tool to diminish the degree of hydrolysis of NDP sialyl donor[45]. In comparison to hemiketals, hemiacetals are more stable sugar donor nucleotides (i.e., GDP-L-fucose). Here, the Leloir glycosyltransferases catalyze hydrolysis to a lesser degree[46]. Interestingly, the affinity of water to the active site for the hydrolysis of sugar nucleotide donors has not been determined for Leloir glycosyltransferases.

2.2.2 Recombinant Expression of Glycosyl Transferases

Although protein structures and the reaction mechanism of Leloir glycosyltransferases are widely investigated, production of the enzyme is often challenging. Heterologous bacterial hosts such as *E. coli* often lead to poor expression or formation of inclusion bodies (IBs), in certain cases with retention of catalytic activity[47]. Besides the difficulties in recombinant protein production and isolation, the half-life of this class of enzymes is often less than a couple of hours[48]. Thermostable glycosyltransferases from thermophilic archaea show higher overall stability[29]. Leloir glycosyltransferases are often aggregation-prone *in vitro* [49]. As a solution to their aggregation, a large number of solubility tags have been successfully applied to increase the solubility of Leloir glycosyltransferases[10, 47b, 50]. The recent advance of using the fluorescent proteins mCherry[47b] or GFP[10] as tags allowed for both an increase in solubility as well as rapid protein quantification. For example, the fusion of GFP allowed for a modular expression approach of all human glycoenzymes in HEK293 cells enabling multi-milligram isolation from the culture media in 65% of all cases[10].

The optimization of protein expression, the number of enzymes discovered, and the characterization of a wide range of Leloir GTs has led to fundamental insights into their protein structures, reaction mechanism, and substrate spectrum. The result of this extensive biochemical knowledge is leading to the adoption of Leloir glycosyltransferases within the field of carbohydrate chemistry. Next, we will discuss how these biochemical insights have been developing alongside their application in chemoenzymatic glycosylations of glycoconjugates and oligosaccharides.

2.3 Application of Glycosyl Transferases in Organic Synthesis

The production of glycosides and glycans requires the use of highly selective catalysts to prevent the formation of side-products. The development of automated chemical methods such as the solid-phase production of oligosaccharides using the Seeberger method[51], the Demchenko synthesizer using HPLC-based platforms for automation[52], and the Yoshida procedure employing an electrochemical oxidation step[53], improved glycochemistry

significantly. The basic principle of elongating a sugar on a solid particle by performing a coupling-wash-deprotection-wash cycle under computer control allows for the rapid production of a wide variety of carbohydrates[54]. The mechanism of action is the assembly of an oligosaccharide using protection group manipulation of either an activated glycosyl acceptor or donor[55]. The purification of the intermediates produced in sequential reactions remains the largest hurdle for chemical synthesis of an oligosaccharide or glycan. In particular, the low orthogonality of activated glycosyl donors and acceptors limits multiple glycosylation reactions in one-pot reactions. Also, the inherently low chemical reactivity of certain glycosidic bond forming reactions, such as *α*-sialylation[56] and *β*-mannosylation[57], restrict different types of linkages. Enzymes which catalyze one-pot glycosylation reactions with unprotected sugars can produce different types of glycosidic linkages and have expanded the synthetic toolbox of glycochemistry considerably.

Leloir glycosyltransferases (GTs) transfer a nucleotide sugar donor to an aglycon acceptor, forming *O*-, *N*-[58] or the rare *C*-[59] and *S*-glycosidic bonds[60] under thermodynamic control. In comparison to chemical methods, the enzymatic coupling of carbohydrates occurs without the use of protecting groups in a highly selective manner, allowing for orthogonal one-pot multi-enzymatic (OPME) reactions. With a few robust GTs, complete libraries of glycans can be constructed[61], which is particularly interesting since most of the human GTs are accessible in heterologous expression systems[10]. The advantages of employing GTs are their mild reaction conditions, short reaction times, pH tolerance, high specific activity, and high yields allowing for the (poly)glycosylation of a wide array of glycans.

2.3.1 Catalytic Reversibility of Glycosyltransferases

One of the notable discoveries on glycosyltransferases was the recognition that glycosyltransferases do not catalyze unidirectional reactions[62]. Alternatively, synthetic sugar donors and/or (chemo)enzymatic regeneration systems either alter the overall K_{eq} or regenerate the nucleotide *in situ*[63]. Such regeneration systems are not always a requirement; the glycosylation with nucleotide sugar donors

allows for repeated glycosylation on a single aglycon (i.e., flavonol-*O*-diglycoside[64]) or elongation of a (poly)saccharide, such as glycogen with a molecular weight of up to 10^7 kDa[65]. The high glycosylation efficiency with Leloir GTs arises from a favorable thermodynamic equilibrium K_{eq} in these examples, determined by the sugar nucleotide donor and carbohydrate or aglycone acceptor, pH, and ionic strength. As mentioned earlier, in a few examples the hydrolysis of the NDP-sugar donor has been reported for Leloir GTs[43-46]. In these particular cases, it is important to emphasize that the glycosylation with Leloir GTs is under kinetic control, and the sugar acceptor and water are competing nucleophiles throughout the entire course of reaction [43-46].

A large impact on the field of glycobiology is the improved group estimation method[66] for the determination of the change in Gibbs free energy of formation of glycosylation reactions with increased accuracy, named eQuilibrator 2.0[67]. In comparison to empiric thermodynamic data (i.e., Thermodynamics of Enzyme-Catalyzed reactions Database[68]), prediction tools allow for a much higher coverage of Gibbs free energies of formation for different compounds. As a drawback, such prediction methods can lead to contradictory observations due to either experimental uncertainties[69] or incorrect analysis of given data[70]. Using Equilibrator 2.0, the synthesis of naturally occurring glycosides with nucleotide diphosphates (NDPs) were shown to be thermodynamically favorable, as is known for the glycosylation of phenolic[71], amino[72], or alcoholic[73] aglycones (Figure 6), and has been reviewed recently[30a]. Interestingly, the importance of the pH has been reported for the glycosylation of acids[71i, 74] resulting in a low $K_{eq} < 1$ at a neutral pH. The K_{eq} depends on the pK_a of the aglycone- or saccharide acceptor, as well as the terminal phosphate of the sugar nucleotide donor.

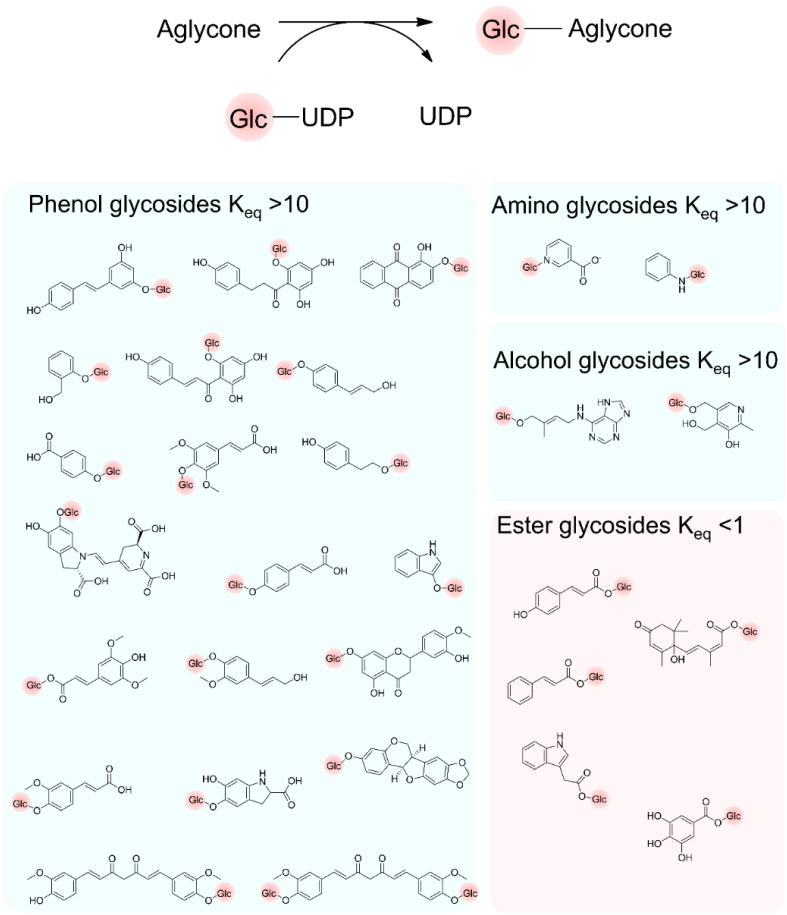


Figure 6. Glycosylation of aglycones producing phenolic glycosides, amino glycosides, alcohol glycosides, ester glycosides, and disaccharides with their estimated K_{eq}. The K_{eq} was calculated from the Gibbs free energy $\Delta G_r'^0$ using the eQuilibrator web interface (http://equilibrator.weizmann.ac.il)[67b] assuming the following conditions: ionic strength 0.1 M, pH 7.0, aglycon (1 mM), UDP (1 mM), UDP-D-glucose (1 mM), glycosylated product (1 mM), and 298 K.

2.3.2 Sugar Donors and Acceptors and Their Glycosylation Efficiency

The thermodynamic constraints of enzymatic glycosylations of sugar acceptors with nucleotide donors for the synthesis of di-, oligo-, or polysaccharides has been explored to a lesser extent. Sucrose synthase has been employed for the regeneration of nucleotide sugars[12, 28, 48a, 75]. The equilibrium constant (K_{eq}) of the

reaction of sucrose with UDP to afford the sugar donor UDP-glucose was determined[76]. The pH influences the K_{eq} for the synthesis of UDP-glucose with *Acidithiobacillus caldus* sucrose synthase (*Ac*SuSy) due to the (de)protonation of the phosphate group of the NDP: going from a pH of 5.0 to 7.0 lowered the K_{eq} of 1.14 to less than 0.1[76]. Enzymatic regeneration of NDP-glucose can also be achieved using trehalose as substrate[77]. For the regeneration of nucleotide sugars, sucrose has been described as a more attractive D-glucopyranosyl donor than α, α-D-trehalose due to the lower free energy of the glycosidic bond, resulting in a more favorable thermodynamic equilibrium[30a, 78].

While the type of carbohydrate donor and acceptor determines the glycosylation product, the respective choice of nucleotide used for activation of the donor glycoside is important from a thermodynamic point of view. Similar enzyme activities and affinities were observed for the coupling of UDP-, GDP-, and ADP-glucose with α-D-glucose by trehalose transferase (TreT) from *Pyrococcous horikoshii*. However, different K_{eq} were observed for the enzymatic production of D-trehalose.[29] In line with these observations, a trehalose transferase from *Thermoproteus uzoniensis* fused to a mCherry solubility tag also reported different K_{eq} for ADP- and UDP-glucose for the production of D-trehalose[47b]. Hence, the overall extent of conversion for the synthesis of disaccharides were determined by the thermodynamics of the nucleotide. Although a thorough examination of the Gibbs free of formation of NMP, NDP, or NTP salt or metals pairs in aqueous solution is beyond the scope of this review, it should be noted that pKa of nucleotides differ affecting the Gibbs free energy of formation. Indeed, the ADP/ADP-glucose couple shows the largest Gibbs free energy change for a transfer of α-glucopyranosyl moiety to a nucleotide, followed by UDP, CDP, and dTDP according to Equilibrator 2.0[67b]. Nature might evolve enzymes to catalyze either the synthesis of nucleotide sugar donors or reactions based on the K_{eq} of nucleotides, as TreT of *Thermococcus litoralis* solely accepts ADP for the transfer of an α-glucopyranosyl moiety from trehalose to produce ADP-glucose[79]. Oppositely, TreT from *Thermoproteus tenax* utilizes UDP-glucose for the synthesis of trehalose since UDP favors synthesis[80]. Further work regarding this is required to elucidate the nature of nucleotides on the K_{eq} in a more comprehensive manner.

Under thermodynamic control, Leloir glycosyltransferases produce oligosaccharides if the overall glycosylation reaction is exergonic (Figure 7). A one-pot procedure using five enzymes allowed for the production of raffinose and stachyose from sucrose[81], using unpurified cell-free extract formulations and supplementation of UDP with a total-turnover number (TTN) of 337. Thermodynamic constraints were observed in the endergonic nucleotide sugar donor production, while coupling of the galactinol, raffinose, and stachyose were exergonic, thereby driving the overall reaction toward oligosaccharide synthesis. The estimation of the Gibbs free energy of individual components gives insights into energetic constraints of one-pot multi-enzyme Leloir glycosyltransferase catalyzed glycosylation reactions. An understanding of these limitations is essential for the optimization of industrial process conditions and reactor design (i.e., product removal) for a biocatalytic process.

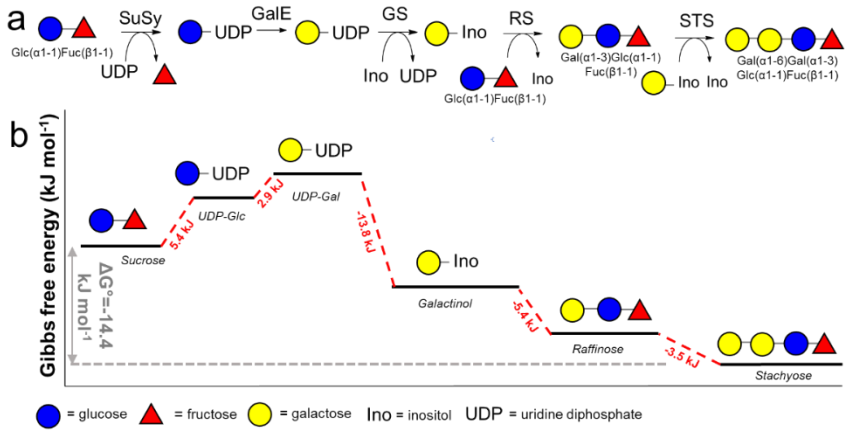


Figure 7. Enzymatic cascade for the production of stachyose from sucrose with glycosyltransferases (**a**). The standard Gibbs free energy changes of the individual reactions (ΔG°s, red) and the total reaction (ΔG°, grey) shown in (**b**)[81]. The $\Delta_rG'°$ represents the change of Gibbs free energy and was calculated using the eQuilibrator web interface (http://equilibrator.weizmann.ac.il)[67b] using the following conditions: ionic strength 0.1 M, pH 7.0, 1 mM of component, 298 K. Abbreviations: UDP-D-glc, UDP-D-glucose; UDP-D-gal, UDP-D-galactose, SuSy, Sucrose synthase; GalE, UDP-D-glucose-4-epimerase; GS, galactinol synthase; RS, raffinose synthase; STS, stachyose synthase.

2.3.3 NTP Regeneration for NDP-Sugar Donor Production

Recycling of the nucleotide sugar donor is considered essential for the application of Leloir glycosyltransferases in large scale applications by preventing product inhibition from the released nucleotide and reducing costs of expensive nucleotides. The use of purified enzymes in comparison to whole-cell systems is often preferred, due to undesired side-reactions of endogenous enzymes of the recombinant hosts during glycosylation of complex oligosaccharides. Most of the glycosyltransferases and enzymes involved in the regeneration of nucleotides operate under neutral conditions and often require the presence of divalent metals, such as Mg^{2+} or Mn^{2+}. As Leloir glycosyltransferases use the elimination of the nucleotide as a driving force for the glycosylation reaction, the high energy gain poses a problem during the regeneration of NDP-sugar donors. For the production of NTP, the driving force then has to be derived from even more energy-rich donors.

Four of the most widely applied enzymatic methods for the regeneration of the nucleotide triphosphates (NTPs) are (see Figure 8): (1) pyruvate kinase using phospho(enol)pyruvate (PEP), (2) acetate kinase using acetyl phosphate, (3) creatine kinase using creatine phosphate, and (4) polyphosphate kinase using polyphosphate. The reaction equilibrium for PEP is highly favorable and the phosphate donor is stable in solution[82]. However, commercial phosphoenolpyruvate is expensive. Creatine phosphate is an alternative donor which is more affordable, but has considerably lower energetic advantages than PEP. A cheap energy-rich phosphate donor is acetyl phosphate, which can be synthesized directly from acetic anhydride and phosphate in excellent yields[83]. The disadvantage of using acetyl phosphate is the rapid spontaneous hydrolysis in water, requiring either continuous supplementation or an excess of acetyl phosphate. The inexpensive (poly)phosphate is a linear polymer that contains from ten to hundreds of energy-rich phosphate linkages[84]. (Poly)phosphate can drive the glycosylation reaction towards completion by the exergonic cleavage of the phosphoanhydride bond ($\Delta G° = 30–32\ kJ \cdot mol^{-1}$ [70]) upon phosphorylation of nucleosides with polyphosphate kinase (PPK). Mono- or diphosphorylation with PPK have been reported for ATP[82, 85], UTP[85b, 86], CTP[87], tTMP[88], often showing broad promiscuity towards different nucleotides[84].

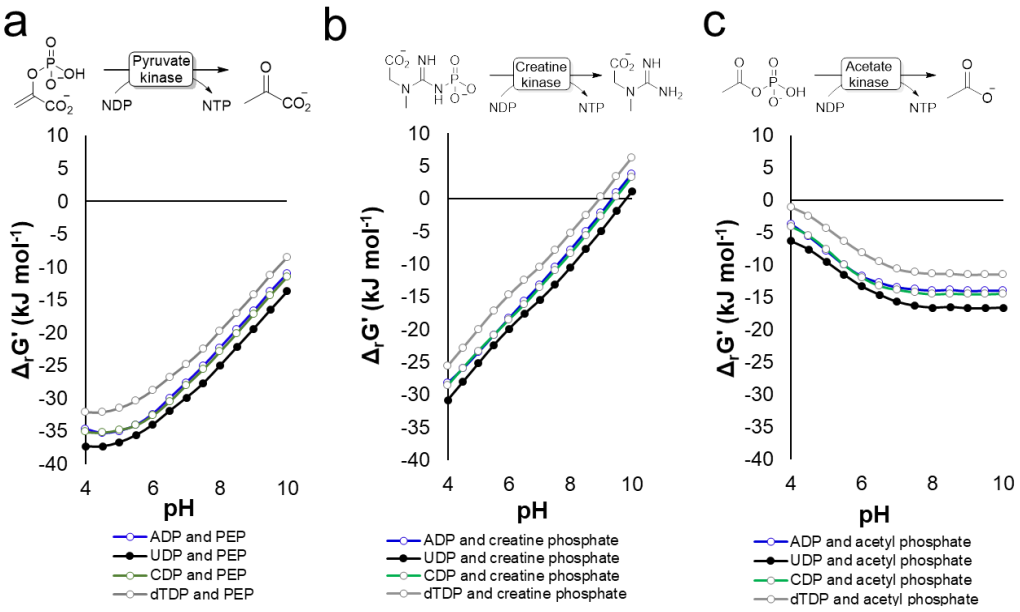


Figure 8. The use of different energy-rich phosphate donors to regenerate NTP using either pyruvate- (**a**), creatine- (**b**), or acetate kinase (**c**). The $\Delta_rG'°$ represents the standard change of Gibbs free energy and was calculated using the eQuilibrator web interface (http://equilibrator.weizmann.ac.il)[67b] using the following conditions: ionic strength 0.1 M, pH 7.0, 1 mM of component, 298 K. Abbreviations: NDP, nucleotide diphosphate; NTP, nucleotide triphosphate; ADP, adenosine diphosphate; UDP, uridine diphosphate; CDP, cytidine diphosphate; dTDP; deoxythymidine diphosphate; PEP, phosphoenolpyruvate.

Different (re)generation schemes for the *in-situ* production of nucleotide sugars for the transfer of a galactosylpyranoside moiety with either stoichiometric amounts of NTP[89], PEP[90], poly(phosphate)[85b], or acetyl phosphate[91] are shown in Figure 9. The main driving force for the glycosylation reaction is the exergonic hydrolysis of pyrophosphate to phosphate by pyrophosphatases or alkaline phosphatases. Although it has been suggested that the sacrificial hydrolysis of NTPs with alkaline phosphatases is beneficial due to the removal of the nucleotide mono-, di-, or triphosphate inhibitors[30a, 92], experimental evidence separating thermodynamics (additional hydrolysis of pyrophosphate) from kinetics (product inhibition) is often not investigated in detail. It is evident that under thermodynamic control nucleotide regeneration and enzymatic glycosylation can only occur with

highly exergonic sacrificial substrates (i.e., hydrolysis of pyrophosphate), as was proposed by Hirschbein *et al*. who compared the energy of hydrolysis of the sacrificial donors as a rationale for glycosylation efficiency[93]. Besides, for the common nucleotide glycosylation donors UDP-Glc, UDP-GlcNAc[94], UDP-GlcA[95], UDP-Gal[94], UDP-GalA, UDP-Xyl, GDP-Man, GDP-Fuc[94], CMP-Neu5Ac[90, 94] the (re)generation systems for the production have been employed for rare or synthetic nucleotide sugar donors, such as CMP-MAnNGc[90], CMP-Man[90], CMP-ManNac5OMe[90], CMP-KDO[90], ADP-Hep[96], and dTDP-Rha[97].

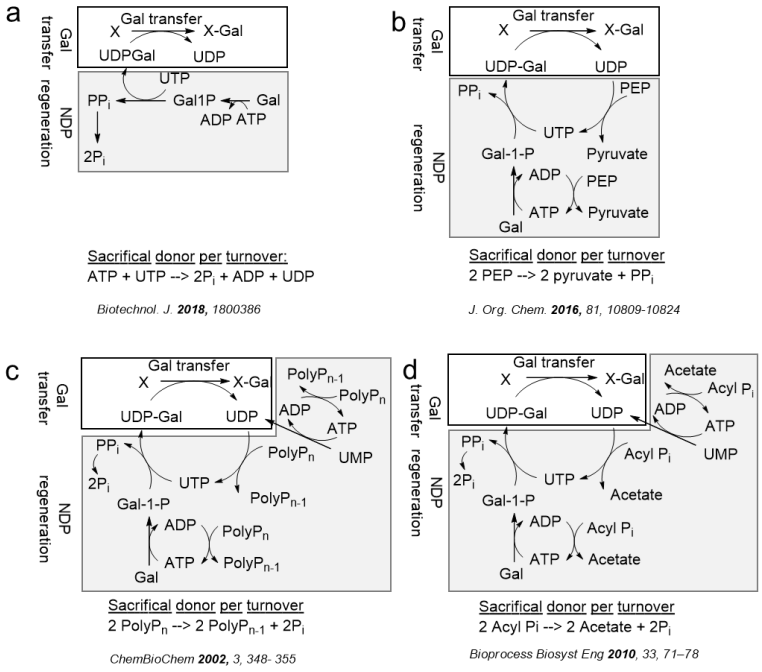


Figure 9. Several sacrificial phosphate donors for sugar nucleotide (re)generation systems of galactosyltransferases using a stoichiometric amount of NTPs (**a**)[89], PEP **(b)**[90], $PolyP_n$ (c)[85b], and acyl P_i (**d**)[91]. Abbreviations: PP_i, pyrophosphate; P_i, orthophosphate; UMP, uridine monophosphate; UDP, uridine diphosphate; UTP, uridine triphosphate; UDP-Gal, UDP-D-galactose; ADP, adenosine diphosphate; ATP, adenosine triphosphate; Gal, D-galactose; Gal1P, D-galactose-1-phosphate; PEP, (phospho)enol pyruvate; $PolyP_n$, (poly)phosphate; acyl P_i, acetyl phosphate.

Often, one-pot multienzyme (OPME) cascade reactions do not go to completion without the thermodynamic driving force from *in-situ* regeneration systems of nucleotide sugar donors. For instance, the gram-scale OPME cascade of the glycoconjugate *N*-acetyl-D-lactosamine resulted in 85% isolated yields and TTN of 80 for UTP (Figure 10a)[98]. The PEP/UDP-regeneration system produces UTP at the expense of PEP while the hydrolysis of pyrophosphate provides the thermodynamic driving force to complete the glycosylation cycle (Figure 10b). Upon replacement of the PEP/UDP-regeneration system with (poly)phosphate/UDP for the enzymatic production of *N*-acetyl-D-lactosamine, no additional pyrophosphatases are required[86b]. Here, the required driving force is generated by the hydrolysis of the energy rich phosphoanhydride bond in (poly) phosphate instead of pyrophosphate hydrolysis.

Besides enzymatic regeneration, the enzymatic NTP synthesis can be performed directly from nucleosides in the presence of an excess of a phosphate donor reducing the overall costs of reagents (i.e., less than US$ 10 per gram UTP). A mutant of uridine kinase from *Thermus thermophilus* phosphorylates a broad range of nucleosides[99]. The addition of an excess of acetyl phosphate allows for the phosphorylation of nucleosides to NMPs with lysates from recombinant *E. coli* containing the overexpressed and promiscuous uridine kinase[97]. Advantageously, cell-free extracts from *E. coli* contain naturally occurring kinases that catalyze sequential phosphorylations to NTP in high yields. Recently, a recombinant *E. coli* strain containing an enzymatic cascade of eight enzymatic steps showed promising production titers of 1.4 g UTP per liter within 2.5 h starting from uracil[100]. Such regeneration systems have been extended to non-natural nucleosides[97, 101].

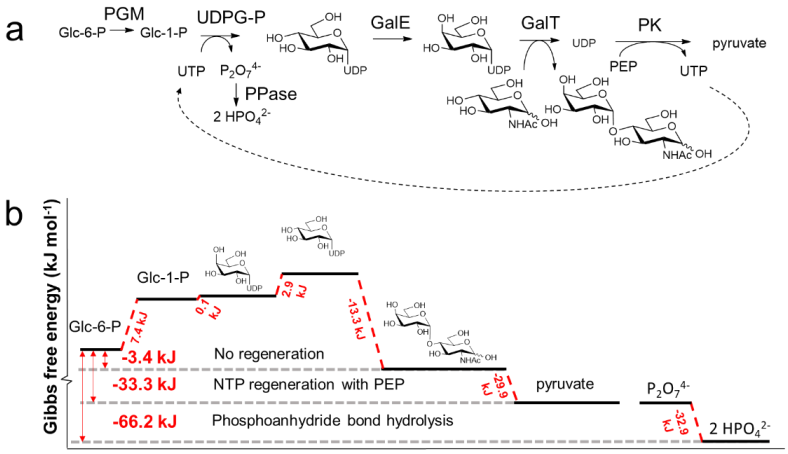


Figure 10. Enzymatic glycosylation for the production of *N*-acetyl-D-lactosamine from glucose-6-phosphate and *N*-acetyl-D-glucosamine (a)[98]. The $\Delta_rG'°$ represents the standard change in Gibbs free energy (b) was calculated using the eQuilibrator web interface (http://equilibrator.weizmann.ac.il)[67b] using the following conditions: ionic strength 0.1 M, pH 7.0, 1 mM of component, 298 K. Abbreviations: Glc-6-P, D-glucose-6-phosphate; Glc-1-P, D-glucose-1-phosphate; $P_2O_7^{4-}$, pyrophosphate; HPO_4^{2-}, orthophosphate; UDP, uridine diphosphate; PGM, Phosphoglucomutase; UDPG-P, UDP-glucose pyrophosphorylase; GalE, UDP-galactose epimerase; GalT, galactosyltransferase; PK, pyruvate kinase.

2.3.4. Chemoenzymatic NTP Regeneration Cascades

The chemoenzymatic synthesis of NDP-sugar donors has been investigated with activated glucose donors under kinetic control. Although the realization that glycosyltransferases catalyze the reverse reaction dates back to 1957[102], the first application for the glycosylation of nucleotides with an activated glycosyl fluoride was reported much later in 1999[103]. The use of β-glucosyl fluoride for the production of UDP-α-glucose using a flavonoid *O*- and *C*-β-glycosyltransferases has been successful for the production of 3′-β-*C*-glucosylated phloretin under kinetic control[104]. Disadvantageously, fast hydrolysis of β-glucosyl fluoride in water limits its practical application. The pioneering work using nitrophenol glycosides demonstrated the broad adaptability of activated sugar acceptors for the glycosylation of nucleotide donors by alternating the thermodynamics of the reaction[105]. The engineered inverting macrolide-

inactivating glycosyltransferase (OleD) from *Streptomyces antibioticus* accepts a wide range of conveniently synthesized aromatic β-D-glucopyranoside donors for the production of UDP-α-D-glucose in the presence of UDP[63a]. The directionality of the reaction is dependent on the nitrophenol β-D-glucopyranoside donor, ranging from exergonic favoring UDP-sugar formation to endergonic favoring the production of the aromatic sugar donor[63a]. Alternatively, by coupling the 2-chloro-4-nitrophenol glycosides to catalytic amounts of nucleotide diphosphate, the glycosylation of a wide variety of substrates has been demonstrated[106]. However, the undesired hydrolysis of 2-chloro-4-nitrophenol glycosides by Leloir glycosyltransferases was observed as well[63e]. Hence, separating glycosyltransferase from glycoside hydrolase activity is not always evident in Leloir glycosyltransferases.

2.3.5. One-Pot Multi Enzyme Cascades

The use of a wide variety of NDP sugar donor regeneration systems coupled to exergonic sacrificial P_i donors inspired the extension of OPME systems towards oligosaccharides and glycans. Key to the success of Leloir glycosyltransferases is the selection of glycosyltransferases with high selectivities towards their substrates, and avoiding the formation of side-products. The human cancer antigen Globo H is a neutral hexasaccharide glycosphingolipid, which has been synthesized chemically by a linear sequence of 11 synthesis steps with predesigned building blocks resulting in a 2.6% overall yield[72a]. The optimization of chemical glycosylation using the OptiMer program with custom-synthesized carbohydrate building blocks constructed Globo H in three consecutive steps, with an isolated yield of 41%[107]. The OptiMer program was improved to 83% isolated yield by a one-pot approach using a complex carbohydrate building block containing a Galα1-4Gal bond from a multi-step synthetic route[108], as is shown in Figure 11. A one-pot biocatalytic coupling of readily available nucleotide sugar donors UDP-galactose, UDP-*N*-acetyl-D-glucosamine, and GDP-fucose with three glycosyltransferases resulted in 54% isolated yield without any nucleotide sugar donor regeneration cycles[109]. Additional regeneration of the nucleotide sugar donor improved the overall yield to 94% at large-scale (i.e. 4.5 g allyl Globo-H)[59e]. The efficiency of enzymatic glycosylation, the availability of the nucleotide sugar

donors, the simplicity of the one-pot reaction, and the mild reaction conditions demonstrate the effectiveness of Leloir glycosyltransferases as catalysts for the production of complex saccharides.

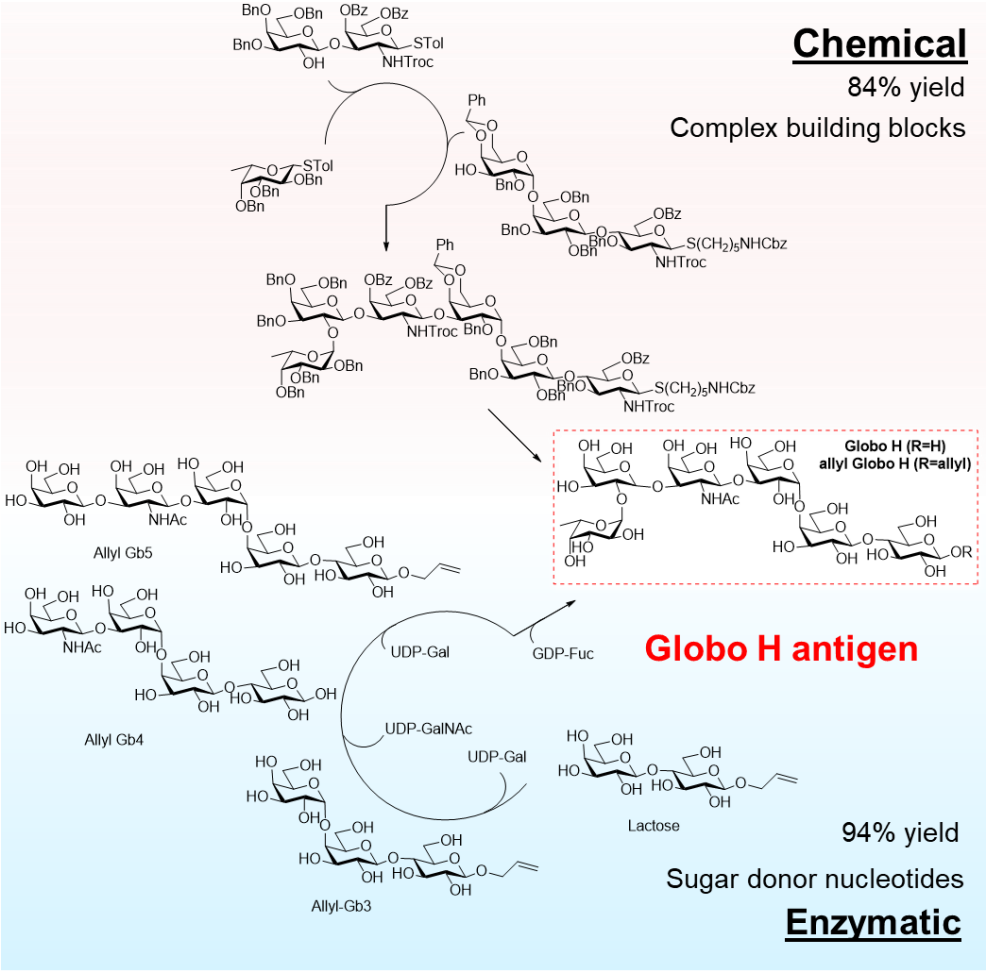


Figure 11. Reaction scheme of both the multistep chemical glycosylation utilizing previously synthesized carbohydrate building blocks[108] and enzymatic glycosylation utilizing nucleotide sugar donors for linear saccharide elongation capped with a terminal vinyl group[59e].

Sequential OPME synthesis allows for the coupling of glycosidic linkages which are synthetically challenging, such as sialylation. The production of

disialoganglioside cancer antigens GD1b and its derivatives by two sequential α-sialylation reactions has been performed using an engineered Leloir glycosyltransferase[90] with limited nucleotide donor hydrolysis activity (Figure 12). [110] Lactose was converted to the trisaccharide GM3 using *α*2-3 sialyltransferase 1 (M144D) from *Pasteuralla multocida*, followed by a second *α*3,8-sialylation in 85% yield to GD3 fusing α2-3/8-sialyltransferase from *Campylobacter jejuni*. Subsequently, the quantitative enzymatic *β*1-4-GalNAc coupling to GD2 with *β*1-4-GalNAc transferase from *Campylobacter jejuni* followed by *β*1-3-Gal transfer with *β*1-3-galactosyltransferase resulted in GD1b with an overall isolated yield of 73%[111].

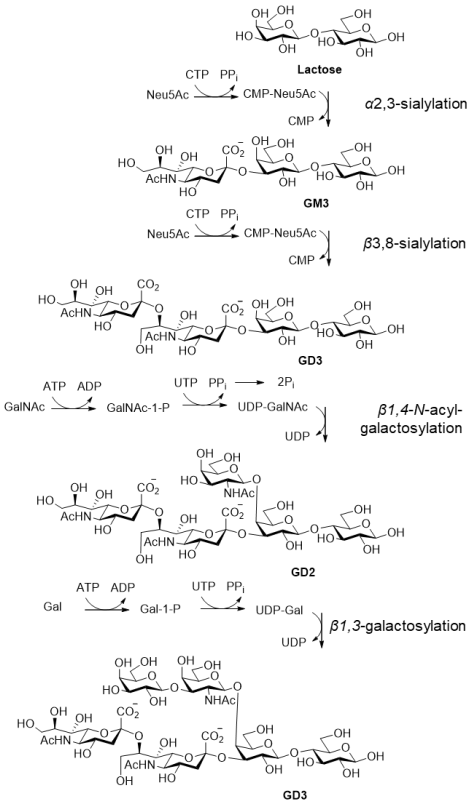


Figure 12. The enzymatic synthesis of GD1B glycan using OPME *α*2,3-sialylation, *α*3,8-sialylation, *β*1,4-*N*-acyl-galactosylation, and *β*1,3-galactosylation with a sacrificial (re)generation system for *N*-acetylneuraminic acid (Neu5Ac), *N*-acetylgalactosamine (GalNAc), and D-galactose (Gal).

2.4 Reactor Engineering for (Non)-LeLoir Glycosyltransferases

One of the advancements in integrated biocatalytic processes using glycosyltransferases is the development of automated enzymatic synthesis, using either immobilized substrates or enzymes. Immobilized substrates allow for the spatiotemporal control of the produced glycoconjugate or oligosaccharide in a reactor. Two prominent approaches using immobilized substrates exist: (i) enzymatic *solution-phase synthesis* with tagged products allowing for rapid purification and (ii) enzymatic *solid-phase synthesis* on the surface of an insoluble carrier with soluble substrates and products.

Solution-phase synthesis with a substrate bound to water-soluble[112] or thermo-responsive polymers[113], fluorous-[106a, 114], ion exchange[115], or lipid-like tags[116] has attracted much interest since it can bypass compatibility issues between enzymes and solid carriers. The main disadvantage of solution-phase assembly of oligosaccharides catalyzed by glycosyltransferases is the low catalytic efficiency and affinity for the substrates due to different steric and stereoelectronic properties induced by the substrate-bound tag. Automated enzymatic synthesis of oligosaccharides with Leloir glycosyltransferases has emerged as a promising approach with the thermoresponsive polymer poly(*N*-isopropylacrylamide) (PNIPAM) as a soluble or insoluble support of the sugars, allowing for the synthesis of the antigen of blood groups A, B, and O, as well as the production of the ganglioside GM1 in microchannel reactors (Figure 13)[117]. The Wong group reported a variety of water-soluble polymers of PNIPAM attached to carbohydrates with different linkers to minimize deleterious effects of the presence of the support on the activity of enzymes[11]. As a disadvantage, the covalent attachment of oligosaccharides to PNIPAM requires cleavage of the oligosaccharide with hydrogen peroxide (1M, pH 10), conditions which are incompatible with oxidative labile carbohydrates (i.e. thiosugars).

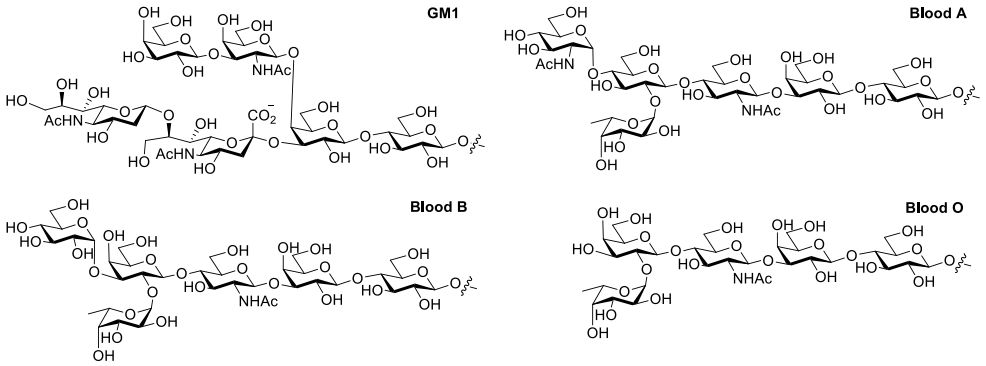


Figure 13. Examples of oligosaccharides synthesized with automated enzymatic synthesis where GM1 is a well-known ganglioside, and the antigens of blood types A, B, and O.

Alternatively, substrate bound *solid-phase synthesis* strategies also received attention[118]. The sugar is immobilized on the solid carrier, while the enzyme and sugar nucleotide donor are dissolved in the mobile phase[119]. Requirements for the full automatization of immobilized substrates are dependent on the development of (i) efficient enzymes; (ii) availability of glycosylation donors; (iii) the use of carrier and support material and (iv) linkers, spacers, and tags[120]. Continuous biocatalytic processes using *immobilized enzymes* are from an engineering perspective highly attractive due to the ease of reuse of the Leloir glycosyltransferase[85b, 121]. The immobilization of Leloir GTs enables a simplification of the reactor's structure and allows for precise control of the enzymatic glycosylation process[121j, 122]. The immobilization of glycosyltransferases has been achieved by attachment onto *solid supports*[47a, 59d, 123], *entrapment* inside a porous carrier[121g, 124], or *cross-linking* in larger aggregates (CLEA)[125], as is shown in Figure 14. Furthermore, after immobilization the reusability, thermal, pH, and operational stability of the enzymes was often increased[59d, 125-126]. In particular cases, enzyme immobilization even created a more favorable micro-environment for enzyme activity[127] and selectivity[128].

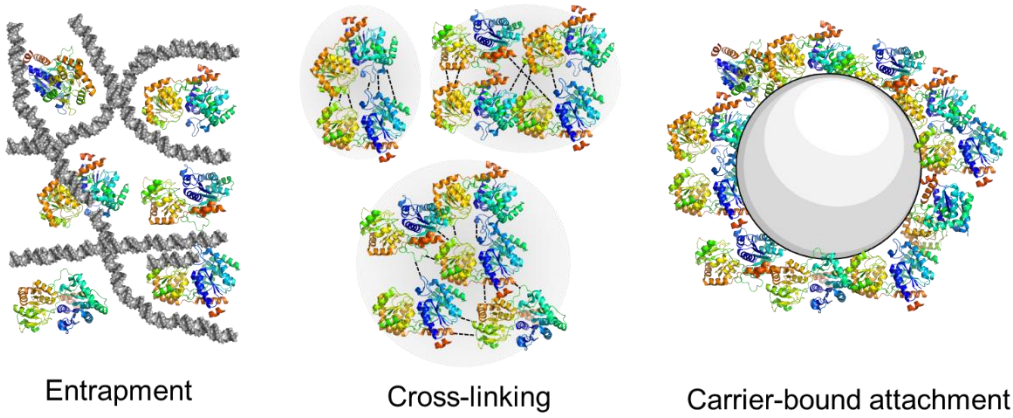


Figure 14. Different modes of immobilization of Leloir glycosyltransferases with entrapment, cross-linking, or carrier-bound attachment.

On the other hand, the reaction conditions in a continuous biocatalytic process with immobilized enzymes can be harsh from an engineering point of view, due to vigorous mixing, high pressures, and flow rates. High enzyme stability of Leloir glycosyltransferases is required to tolerate shear stress[129]. Also, due to the lack of an universal enzyme immobilization technique many factors must be considered[122a, 129b, 130], including mode of *interactions* (i.e., enzyme-substrate/product, enzyme-carrier, substrate/product-carrier), *compatibility* of the carrier to reaction conditions (solvent, temperature, pH, etc.), and the *type of reactor or process* (i.e. batch reactor, packed-bed reactor, basket reactor, microfluidic reactor).

2.4.1 Reactor Design for Glycosyltransferases

For industrial applications, besides the appropriate immobilization of the enzymes, the chemical character of the carrier's surface is of high importance as are the particle size and pore structure which need to match the type of reactor: i.e. tank vs. tube/column and the mode of operation: periodic/batch vs. continuous flow. Conventional stirred tank reactors (STR)[123b, 128b, 131] are still frequently applied, but basket and in particular rotating bed reactors[132] are increasingly used in batch operations. In continuous flow applications, packed-bed (micro)reactors[133] have been most popular, but structural (micro)reactors[134], lab-on-a-chip[135], and

capillary microreactors are on the rise[136]. The STRs are predominantly used in biotechnology owing to cost efficiency and versatility. However, vigorous mixing results in frequent collisions of fine biocatalyst particles resulting in tensile and shearing forces which enhance abrasion or disintegration of the enzyme or its carrier[132]. Moreover, the mixing on a medium and large scale can be insufficient to prevent “hot spot” formation, resulting in enzyme denaturation. But even more importantly, as mass transfer of reactants to fine particles of the biocatalysts or freely suspended enzymes is low, the mass transfer does not keep up the pace of intrinsic activity of the highly active enzymes expressed by TOF values of about or over 10^4 s^{-1}. In effect, it is the mass transfer that strongly impacts, or even fully controls the apparent rate of the enzyme catalyzed reactions carried out in STRs, and that has unfortunately often been overlooked[136-137]. The recovery of the freely suspended biocatalyst particles can be challenging, requiring filtration or centrifugation[138]. The biocatalyst recovery can be simplified using a basket-rotating bed reactor (RBR), which enables simultaneous mixing and effective percolation of liquid through the bed of the catalyst packed in a cylindrical basket, thus avoiding the catalysts destruction, enhancing mass transport, and facilitating separation of the catalyst[139]. However, the size of the applied catalyst particles/enzyme carriers needs to be larger than 0.1-0.2 mm[139a, 139b].

The scale-up of RBRs is challenging due to the large size of the rotor as well as the power required for stirring, although a 750-L scale has been successfully demonstrated (Chiralvision). Low-volume continuous flow reactors with flow through channels typically in the range of diameters 0.1-0.5 mm (capillary microreactors) are better scalable[136, 140]. The problems related to scaling up are resolved by numbering up the (micro) reactors in a parallel process, designated as ‘scaling out’. In particular, if thermal effects are not too strong, as is often the case, for the same type of packing the scalability is not a problem. Average linear velocity of flowing reactants and a mean residence time need to be identical on different scales to obtain the same conversion. In addition, owing to narrow channels the time needed for substrate transfer from the center of the liquid channel to the wall-attached catalysts is significantly reduced and the transfer of heat and mass increased, with a positive effect on the apparent reaction rate. Also, the ratio of the active surface of the reactor to its volume may be increased by a factor of 50 or

even much more, (e.g., from about 10^3 m^{-1} in industrial reactor to ~$5 \cdot 10^4$ m^{-1} for 0.1 mm capillaries) which results in an increase of volumetric productivity (space-time-yield) and hence more profit and lower investment costs[136-137, 140c]. Moreover, the low reaction volume favors the reduction of potential hazards, particularly important in the case of highly exothermic reactions or when hazardous substances are involved[137, 140b]. A simplified approach may be used to evaluate process boundaries for the capillary microreactors with enzymes attached to the wall surface[136]. But more in-depth analyses of design and modelling issues for various continuous flow microreactors have also been reported[140c].

Different examples of reactors that have been used with glycosyltransferases are shown in Figure 15. *Stirred tank reactors* are flexible in design and operation conditions, but often require high operation costs and vary in the product quality per batch[59d, 123b, 128b, 131]. *Microchannel reactors* feature flow-through channels of micrometric sizes that contain the enzyme immobilized on their wall surface[136]. *Packed-bed (micro) reactors* contain fine particles with immobilized enzymes in a flow-through channel, allowing for a higher volumetric activity than microchannel reactors. The heterogenous biocatalyst should not be able to compact to avoid high pressure drops, while mass transport between liquid reactants and catalyst surface is enhanced owing to a more chaotic flow which facilitates mixing. The large pressure drop, even at low flow rates may, however, be a problem if fine catalyst particles are applied. Structured microreactors contain a reactor core made of a porous monolithic structure with open, usually curved pores/channels connected with each other offering excellent mixing and mechanical stability. Moreover, the pressure drop can be significantly reduced and flow rate increased, compared to the packed-bed reactors, and this boosts productivity. The enzymes are immobilized either on the external surface of the monolith or in its pores[133].

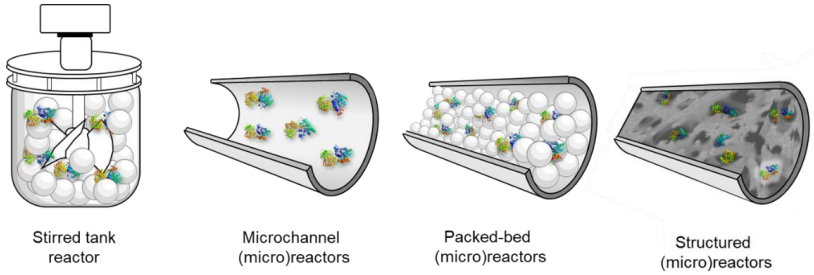


Figure 15. Different reactor types for the immobilization of Leloir glycosyltransferases.

A packed-bed reactor is a commonly used system for continuous production with a heterogeneous biocatalyst, especially because one can immobilize simultaneously different enzymes. Schöffer, et al. used glutaraldehyde-activated chitosan spheres [123b] and amino- or thiol- functionalized silica particles[123c] as a support for cyclodextrin glucosyltransferase immobilization. The silica-based biocatalyst was successfully applied in a packed-bed reactor for continuous cyclodextrin production and maintained 100% of its initial activity after 200 h, whereas activity of the chitosan-based catalyst decreased to 78% of its initial value already after 50 h. This was ascribed to the super packing of spheres, resulting in the reduction of bed height by 45%, and thus in a decrease in the residence/reaction time. However, after washing and re-packing, the spheres recovered 100% of their initial activity.

An interesting effect was observed by Cho, et al.[131b], who compared the performance of batch and continuous packed-bed reactors using Eupergit C250L as an enzyme support. The batch reaction was performed for trehalose production from maltose using trehalose synthase for more than 20 h. They found that the product composition was almost the same after 10 h and a maximum trehalose production yield of 25% was established. Trehalose production was improved using a packed-bed bioreactor, wherein the yield reached 42% with a retention time of 100 min. The authors claimed that continuous feeding of fresh substrate into the packed-bed reactor might have eliminated and removed inhibitory compounds from the solution such as by-products formed and accumulated during the batch reaction and thus increased the production yield. A combination of continuous-flow stirred tank reactor (CSTR) with a packed-bed reactor (PBR) was also studied[131a].

A highly concentrated starch solution was first partially converted to β-cyclodextrins in a CSTR, which resulted in a decrease of starch viscosity. After that, the reaction mixture was pumped through the PBR. The integrated reactor offered much higher concentration of the final product than each of the reactors separately.

Integration of two or more reactors attracts increasing attention, especially if two or more enzymes are applied. For continuous flow nucleoside synthesis Cattaneo, et al.[133] combined a PBR, filled with purine nucleoside phosphorylase immobilized on silica particles, with uridine phosphorylase immobilized on a silica monolith. In the first approach, co-immobilization of both enzymes on a slightly longer silica-filled PBR was tested, and a high immobilization yield was obtained. However, a very high backpressure of the system (>10 MPa) was registered even at a low flow rate value of 0.1 mL/min, thus hampering the full characterization of the reactor system and resulting in a dramatic drop in conversion. Nonetheless, the application of a monolithic reactor, which exhibited only 6 MPa of pressure drop at a flow rate of 0.5 mL/min, combined with a shorter PBR, showed good activity and stability.[133] The additional advantage of this set-up would be the availability of a single bioreactor that could be used independently, either for “one-enzyme” synthesis or in a different sequence.

In the most recent work Nidetzky, et al.[136], presented an elegant exemplary glycosylation process with sucrose phosphorylase immobilized on the internal surface of a microchannel (Figure 16). Its mathematical model clearly demonstrated that microreactors with the lower hydrolytic channel diameter (d_h) exhibit enhanced performance in terms of conversion and space-time-yield (STY). As the enzyme was attached on the microchannel’s wall only the external mass transfer had to be considered[140c], and the enzymatic transformations appeared to experience a shift from diffusion to reaction control with miniaturization of d_h (second Damköhler number—$Da_{II} < 1$). Thus, the microreactors, in consequence of their small d_h, emerge as an effective means of gaining full control of the reaction[136]. However, the practical limits to the decrease in d_h, due to high pressure drop, and increased tendency of microchannel clogging have to be kept in mind[134a, 136]. Therefore, to boost both STY and microreactor performance, a combination of d_h decrease and enzyme activity increase appears to be the rational solution[136].

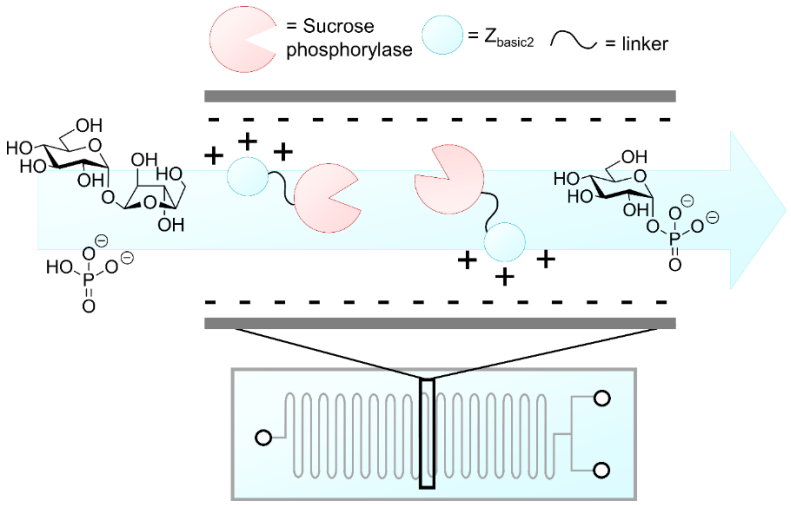


Figure 16. Glass microchannel reactor with immobilized enzyme sucrose phosphorylase attached to the siliceous wall[136]. Attachment occurs via the highly positively charged Z_{basic2} binding module, which binds to the negatively charged silica surface.

To summarize, the application of miniaturized synthetic systems and flow microreactors with immobilized enzymes in particular, attracts attention for the synthesis of more complex carbohydrates. The reported studies clearly demonstrate important advantages of microreactor-based synthetic processes: good stability and high activity that allow for very effective/highly productive syntheses of targeted carbohydrates. While the application of capillary or packed-bed microreactors has been best characterized, structured microreactors are emerging as a class of miniaturized devices that offer additional advantages.

2.5 Summary and Outlook

The discovery, characterization, and engineering of Leloir glycosyltransferases has expanded the synthetic toolbox to couple, elongate, or branch glycoconjugates, oligosaccharides, and glycans with high regio- and stereoselectivity. Efficient regeneration systems and large-scale production of nucleotide sugar donors under either kinetic or thermodynamic control increased the efficiency of enzymatic glycosylation, reducing overall process costs and the

use of stoichiometric amounts of nucleotide phosphates. In this review, the importance of the thermodynamics of glycosylation reactions has been given attention, separating kinetics from thermodynamics for the coupling of a wide variety of NDP sugar donors with acceptors, including sacrificial phosphate donors for NDP (re)generation.

The wide range of different Leloir glycosyltransferases has allowed for the implementation of enzymes in glycochemistry, and consequently, industrially applicable glycosylation methodologies are now in progress. The protein production improved significantly by using (fluorescent) solubility tags, allowing for the production of the biocatalyst in high(er) titers. Despite of the production of enzymes becoming a routine, no industrial application of Leloir glycosyltransferases for the glycosylation of large oligosaccharides has yet been scaled to large volumes, in contrast to non-Leloir glycosyltransferase (i.e. cyclodextrin glycosyltransferases). Until now, one of the main limitations has been the cost-efficient production of NDP sugar donors, which has been an important topic of interest in the last decade. Indeed, due to the advance of many chemical and biocatalytic NDP-sugar production processes, their commercial cost-price has been rapidly declining over the last few years.

As the demand for high-value antigens is increasing, the number of biocatalytic glycosylation processes applied for the synthesis of these complex oligosaccharides can be anticipated to rise. One trend is the embracement of automated enzymatic synthesizers for the computer-controlled synthesis of large oligosaccharides using Leloir GTs. In comparison to non-enzymatic coupling strategies, which mostly rely on protection group chemistry, Leloir GTs have now been shown to couple a wide spectrum of unprotected sugar acceptors with excellent regio- and enantioselectivity. Due to the availability of the nucleotide sugar donors and well-established NDP (re)generation systems, Leloir GTs matured as a competitive glycosylation strategy for the enzymatic synthesis of oligosaccharides.

Future developments for industrial enzymatic glycosylation are expected to mostly be focused on optimizing the overall glycosylation process conditions. The main drivers for selecting the most optimal process can be attributed to economic (i.e., revenue of products), development (i.e., time), and process (i.e., performance)

parameters. Multiple aspects influence these important aspects, such as the selection of the most optimal reactor design, separating batch versus continuous process operation, the choice of either immobilized substrates to immobilized enzymes, or the use of NDP-regeneration system or stoichiometric use of NDPs. Different enzymatic solution-phase and solid-phase glycosylation strategies have been developed for the automated enzymatic synthesis of carbohydrates. The enzymatic synthesis of an immobilized substrate allows for the purification to become more straight-forward, but requires the stoichiometric use of enzymes. On the other hand, immobilized Leloir GTs in continuous operations have been described sparingly in cascade glycosylation reactions. Immobilized enzymes, ensuring process flexibility and the purification of the produced oligosaccharide are engineering design challenges which have yet to be met.

References

[1] a) E. Fischer, *Berichte der Deutschen Chemischen Gesellschaft* **1891**, *24*, 1836-1845; b) E. Fischer, *Berichte der Deutschen Chemischen Gesellschaft* **1891**, *24*, 2683-2687.

[2] F. Wöhler, J. Liebig, *Annalen der Pharmacie* **1837**, *22*, 1-24.

[3] R. DiCosimo, J. McAuliffe, A. J. Poulose, G. Bohlmann, *Chemical Society Reviews* **2013**, *42*, 6437-6474.

[4] R. A. Laine, *Glycobiology* **1994**, *4*, 759-767.

[5] A. Varki, R. D. Cummings, M. Aebi, N. H. Packer, P. H. Seeberger, J. D. Esko, P. Stanley, G. Hart, A. Darvill, T. Kinoshita, J. J. Prestegard, R. L. Schnaar, H. H. Freeze, J. D. Marth, C. R. Bertozzi, M. E. Etzler, M. Frank, J. F. G. Vliegenthart, T. Lütteke, S. Perez, E. Bolton, P. Rudd, J. Paulson, M. Kanehisa, P. Toukach, K. F. Aoki-Kinoshita, A. Dell, H. Narimatsu, W. York, N. Taniguchi, S. Kornfeld, *Glycobiology* **2015**, *25*, 1323-1324.

[6] a) M. Puri, A. Kaur, R. S. Singh, J. R. Kanwar, Transworld Research Network, **2008**, pp. 91-103; b) H. J. Vila Real, A. J. Alfaia, A. R. T. Calado, M. H. L. Ribeiro, *Food Chemistry* **2007**, *102*, 565-570; c) M. Puri, U. C. Banerjee, *Biotechnology Advances* **2000**, *18*, 207-217.

[7] a) M. Kubota, *New Food Industry* **2005**, *47*, 17-29; b) M. Walmagh, R. Zhao, T. Desmet, *International Journal of Molecular Sciences* **2015**, *16*, 13729.

[8] a) A. R. Szaniawski, H. G. Spencer, *Journal of Membrane Science* **1997**, *127*, 69-76; b) I. Alkorta, C. Garbisu, M. J. Llama, J. L. Serra, *Process Biochemistry* **1998**, *33*, 21-28; c) P. Lozano, A. Manjón, J. Iborra, M. Cánovas, F. Romojaro, *Enzyme and Microbial Technology* **1990**, *12*, 499-505; d) P. Lozano, A. Manjón, F. Romojaro, M. Canovas, J. L. Iborra, *Biotechnology Letters* **1987**, *9*, 875-880; e) I. Alkorta, C. Garbisu, M. J. Llama, J. L. Serra, *Biotechnology Techniques* **1995**, *9*, 95-100.

[9] a) M. P. Lisboa, N. Khan, C. Martin, F.-F. Xu, K. Reppe, A. Geissner, S. Govindan, M. Witzenrath, C. L. Pereira, P. H. Seeberger, *Proceedings of the National Academy of Sciences* **2017**, 201706875; b) M. Emmadi, N. Khan, L. Lykke, K.

Reppe, S. G. Parameswarappa, M. P. Lisboa, S.-M. Wienhold, M. Witzenrath, C. L. Pereira, P. H. Seeberger, *Journal of the American Chemical Society* **2017**, *139*, 14783-14791; c) M. Cavallari, P. Stallforth, A. Kalinichenko, D. C. K. Rathwell, T. M. A. Gronewold, A. Adibekian, L. Mori, R. Landmann, P. H. Seeberger, G. De Libero, *Nature Chemical Biology* **2014**, *10*, 950; d) D. Oldrini, T. Fiebig, M. R. Romano, D. Proietti, M. Berger, M. Tontini, R. De Ricco, L. Santini, L. Morelli, L. Lay, R. Gerardy-Schahn, F. Berti, R. Adamo, *ACS Chemical Biology* **2018**, *13*, 984-994; e) D. J. Marciani, J. B. Press, R. C. Reynolds, A. K. Pathak, V. Pathak, L. E. Gundy, J. T. Farmer, M. S. Koratich, R. D. May, *Vaccine* **2000**, *18*, 3141-3151; f) T. Fiebig, M. R. Romano, D. Oldrini, R. Adamo, M. Tontini, B. Brogioni, L. Santini, M. Berger, P. Costantino, F. Berti, R. Gerardy-Schahn, *Npj Vaccines* **2016**, *1*, 16017.

[10] K. W. Moremen, A. Ramiah, M. Stuart, J. Steel, L. Meng, F. Forouhar, H. A. Moniz, G. Gahlay, Z. Gao, D. Chapla, S. Wang, J.-Y. Yang, P. K. Prabhakar, R. Johnson, M. d. Rosa, C. Geisler, A. V. Nairn, J. Seetharaman, S.-C. Wu, L. Tong, H. J. Gilbert, J. LaBaer, D. L. Jarvis, *Nature Chemical Biology* **2017**, *14*, 156.

[11] L. Wen, G. Edmunds, C. Gibbons, J. Zhang, M. R. Gadi, H. Zhu, J. Fang, X. Liu, Y. Kong, P. G. Wang, *Chemical Reviews* **2018**, *118*, 8151-8187.

[12] F. De Bruyn, J. Maertens, J. Beauprez, W. Soetaert, M. De Mey, *Biotechnology Advances* **2015**, *33*, 288-302.

[13] L. L. Lairson, B. Henrissat, G. J. Davies, S. G. Withers, *Annual Review of Biochemistry* **2008**, *77*, 521–555.

[14] T. M. Gloster, *Current Opinion in Structural Biology* **2014**, *28*, 131–141.

[15] C. Whitfield, M. S. Trent, *Annual Review of Biochemistry* **2014**, *83*, 99–128.

[16] A. Typas, M. Banzhaf, C. A. Gross, W. Vollmer, *Nature Reviews Microbiology* **2011**, *10*, 123–136.

[17] B. L. Cantarel, P. M. Coutinho, C. Rancurel, T. Bernard, V. Lombard, B. Henrissat, *Nucleic Acids Research* **2009**, *37*, D233-238.

[18] J. A. Campbell, G. J. Davies, V. Bulone, B. Henrissat, *The Biochemical Jjournal* **1997**, *326 (Pt 3)*, 929–939.

[19] M. Sinnott, *Carbohydrate Chemistry and Biochemistry: Structure and Mechanism*, Royal Society of Chemistry, Cambridge, **2013**.

[20] I. Tvaroška, *Carbohydrate Research* **2015**, *403*, 38–47.

[21] a) P. Welzel, *Chemical Reviews* **2005**, *105*, 4610-4660; b) A. L. Lovering, L. H. de Castro, D. Lim, N. C. J. Strynadka, *Science* **2007**, *315*, 1402-1405.

[22] M. Timm, J. Görl, M. Kraus, S. Kralj, H. Hellmuth, R. Beine, K. Buchholz, L. Dijkhuizen, J. Seibel, *ChemBioChem* **2013**, *14*, 2423-2426.

[23] Y. Bi, C. Hubbard, P. Purushotham, J. Zimmer, *Current Opinion in Structural Biology* **2015**, *34*, 78–86.

[24] C. Breton, L. Snajdrová, C. Jeanneau, J. Koca, A. Imberty, *Glycobiology* **2006**, *16*, 29R–37R.

[25] H. Zhang, F. Zhu, T. Yang, L. Ding, M. Zhou, J. Li, S. M. Haslam, A. Dell, H. Erlandsen, H. Wu, *Nature Communications* **2014**, *5*, 4339.

[26] M. D. Kattke, J. E. Gosschalk, O. E. Martinez, G. Kumar, R. T. Gale, D. Cascio, M. R. Sawaya, M. Philips, E. D. Brown, R. T. Clubb, *PLOS Pathogens* **2019**, *15*, e1007723.

[27] A. Ardèvol, C. Rovira, *Journal of the American Chemical Society* **2015**, *137*, 7528-7547.

[28] A. Gutmann, A. Lepak, M. Diricks, T. Desmet, B. Nidetzky, *Biotechnology Journal* **2017**, *12*, 1600557.

[29] S.-I. Ryu, J.-E. Kim, E.-J. Kim, S.-K. Chung, S.-B. Lee, *Process Biochemistry* **2011**, *46*, 128-134.

[30] a) B. Nidetzky, A. Gutmann, C. Zhong, *ACS Catalysis* **2018**, *8*, 6283-6300; b) A. Chang, S. Singh, G. N. Phillips, J. S. Thorson, *Current Opinion in Biotechnology* **2011**, *22*, 800-808; c) C. Breton, S. Fournel-Gigleux, M. M. Palcic, *Current Opinion in Structural Biology* **2012**, *22*, 540-549; d) J. I. Blanco Capurro, C. W. Hopkins, G. Pierdominici Sottile, M. C. González Lebrero, A. E. Roitberg, M. A. Marti, *The Journal of Physical Chemistry. B* **2017**, *121*, 471–478.

[31] a) D. Albesa-Jové, M. Á. Sainz-Polo, A. Marina, M. E. Guerin, *Angewandte Chemie* **2017**, *129*, 15049–15053; b) S. J. Charnock, G. J. Davies, *Biochemistry* **1999**, *38*, 6380–6385.

[32] a) R. E. Huber, M. T. Gaunt, K. L. Hurlburt, *Archives of Biochemistry and Biophysics* **1984**, *234*, 151-160; b) Y. Ooi, N. Mitsuo, T. Satoh, *Chemical & Pharmaceutical Bulletin* **1985**, *33*, 5547-5550; c) M. Okuyama, H. Mori, K. Watanabe, A. Kimura, S. Chiba, *Bioscience, Biotechnology, and Biochemistry* **2002**, *66*, 928-933; d) M. Jahn, J. Marles, R. A. J. Warren, S. G. Withers, *Angewandte Chemie International Edition* **2003**, *42*, 352-354.

[33] a) A. Aguirre, S. Peiru, F. Eberhardt, L. Vetcher, R. Cabrera, H. G. Menzella, *Applied Microbiology and Biotechnology* **2014**, *98*, 4033-4040; b) X. Peng, H. Su, S. Mi, Y. Han, *Biotechnology for Biofuels* **2016**, *9*, 98; c) U. R. Ezeilo, I. I. Zakaria, F. Huyop, R. A. Wahab, *Biotechnology & Biotechnological Equipment* **2017**, *31*, 647-662; d) H. Michlmayr, E. Varga, A. Malachova, N. T. Nguyen, C. Lorenz, D. Haltrich, F. Berthiller, G. Adam, *Applied and Environmental Microbiology* **2015**, *81*, 4885-4893.

[34] N. Mitsuo, H. Takeichi, T. Satoh, *Chemical & Pharmaceutical Bulletin* **1984**, *32*, 1183-1187.

[35] a) D. G. Drueckhammer, C. H. Wong, *The Journal of Organic Chemistry* **1985**, *50*, 5912-5913; b) A. M. Gold, M. P. Osber, *Biochemical and Biophysical Research Communications* **1971**, *42*, 469-474; c) S. J. Williams, S. G. Withers, *Carbohydrate Research* **2000**, *327*, 27-46; d) J. F. Tolborg, L. Petersen, K. J. Jensen, C. Mayer, D. L. Jakeman, R. A. J. Warren, S. G. Withers, *The Journal of Organic Chemistry* **2002**, *67*, 4143-4149.

[36] P. Fialová, A. T. Carmona, I. Robina, R. Ettrich, P. Sedmera, V. Přikrylová, L. Petrásková-Hušáková, V. Křen, *Tetrahedron Letters* **2005**, *46*, 8715-8718.

[37] E. Strahsburger, A. M. L. de Lacey, I. Marotti, D. DiGioia, B. Biavati, G. Dinelli, *Electronic Journal of Biotechnology* **2017**, *30*, 83-87.

[38] T. Yasukochi, K. Fukase, S. Kusumoto, *Tetrahedron Letters* **1999**, *40*, 6591-6593.

[39] V. Chiffoleau-Giraud, P. Spangenberg, C. Rabiller, *Tetrahedron: Asymmetry* **1997**, *8*, 2017-2023.

[40] G. Vic, D. H. G. Crout, *Carbohydrate Research* **1995**, *279*, 315-319.

[41] T. Hansson, M. Andersson, E. Wehtje, P. Adlercreutz, *Enzyme and Microbial Technology* **2001**, *29*, 527-534.

[42] H. A. Gavlighi, A. S. Meyer, J. D. Mikkelsen, *Biotechnology Letters* **2013**, *35*, 205-212.

[43] a) J. M. Cote, E. A. Taylor, *International Journal of Molecular Sciences* **2017**, *18*, 2256; b) S. Moréra, A. Imberty, U. Aschke-Sonnenborn, W. Rüger, P. S. Freemont, *Journal of Molecular Biology* **1999**, *292*, 717-730; c) T. Czabany, K. Schmölzer, C. Luley-Goedl, D. Ribitsch, B. Nidetzky, *Analytical Biochemistry* **2015**, *483*, 47-53; d) L. Ding, C. Zhao, J. Qu, Y. Li, G. Sugiarto, H. Yu, J. Wang, X. Chen, *Carbohydrate Research* **2015**, *408*, 127-133.

[44] K. Schmölzer, C. Luley-Goedl, T. Czabany, D. Ribitsch, H. Schwab, H. Weber, B. Nidetzky, *FEBS Letters* **2014**, *588*, 2978-2984.

[45] a) G. Sugiarto, K. Lau, J. Qu, Y. Li, S. Lim, S. Mu, J. B. Ames, A. J. Fisher, X. Chen, *ACS Chemical Biology* **2012**, *7*, 1232-1240; b) K. Schmölzer, M. Eibinger, B. Nidetzky, *ChemBioChem* **2017**, *18*, 1544-1550.

[46] C. Zhao, Y. Wu, H. Yu, I. M. Shah, Y. Li, J. Zeng, B. Liu, D. A. Mills, X. Chen, *Chemical Communications* **2016**, *52*, 3899-3902.

[47] a) R. Koszagova, T. Krajcovic, K. Palencarova-Talafova, V. Patoprsty, A. Vikartovska, K. Pospiskova, I. Safarik, J. Nahalka, *Microbial Cell Factories* **2018**, *17*, 139; b) L. Mestrom, S. R. Marsden, M. Dieters, P. Achterberg, L. Stolk, I. Bento, U. Hanefeld, P.-L. Hagedoorn, *Applied and Environmental Microbiology* **2019**, AEM.03084-03018.

[48] a) K. Schmölzer, A. Gutmann, M. Diricks, T. Desmet, B. Nidetzky, *Biotechnology Advances* **2016**, *34*, 88-111; b) L. Bungaruang, A. Gutmann, B. Nidetzky, *Advanced Synthesis & Catalysis* **2013**, *355*, 2757-2763; c) K. Schmölzer, M. Lemmerer, B. Nidetzky, *Biotechnology and Bioengineering* **2018**, *115*, 545-556.

[49] a) Q. M. Wang, R. B. Peery, R. B. Johnson, W. E. Alborn, W.-K. Yeh, P. L. Skatrud, *Journal of Bacteriology* **2001**, *183*, 4779-4785; b) L. Chen, P. Sun, Y. Li, M. Yan, L. Xu, K. Chen, P. Ouyang, *3 Biotech* **2017**, *7*, 356.

[50] a) D. H. Welner, D. Shin, G. P. Tomaleri, A. M. DeGiovanni, A. Y.-L. Tsai, H. M. Tran, S. F. Hansen, D. T. Green, H. V. Scheller, P. D. Adams, *PLOS ONE* **2017**, *12*, e0177591; b) M. E. Ortiz-Soto, J. Seibel, *PLOS ONE* **2016**, *11*, e0155410.

[51] O. J. Plante, E. R. Palmacci, P. H. Seeberger, *Science* **2001**, *291*, 1523-1527.

[52] S. G. Pistorio, S. S. Nigudkar, K. J. Stine, A. V. Demchenko, *The Journal of Organic Chemistry* **2016**, *81*, 8796-8805.

[53] T. Nokami, R. Hayashi, Y. Saigusa, A. Shimizu, C.-Y. Liu, K.-K. T. Mong, J.-i. Yoshida, *Organic Letters* **2013**, *15*, 4520-4523.

[54] a) T.-Y. Huang, M. M. L. Zulueta, S.-C. Hung, *Organic & Biomolecular Chemistry* **2014**, *12*, 376-382; b) C.-C. Wang, J.-C. Lee, S.-Y. Luo, S. S. Kulkarni, Y.-W. Huang, C.-C. Lee, K.-L. Chang, S.-C. Hung, *Nature* **2007**, *446*, 896; c) M. T. C. Walvoort, A. G. Volbeda, N. R. M. Reintjens, H. van den Elst, O. J. Plante, H. S. Overkleeft, G. A. van der Marel, J. D. C. Codée, *Organic Letters* **2012**, *14*, 3776-3779; d) A. Geert Volbeda, J. van Mechelen, N. Meeuwenoord, H. S. Overkleeft, G. A. van der Marel, J. D. C. Codée, *The Journal of Organic Chemistry* **2017**, *82*, 12992-13002; e) J. Zhou, S. Lv, D. Zhang, F. Xia, W. Hu, *The Journal of Organic Chemistry* **2017**, *82*, 2599-2621; f) S. Kaeothip, A. V. Demchenko, *The Journal of Organic Chemistry* **2011**, *76*, 7388-7398; g) O. Kanie, Y. Ito, T. Ogawa, *Journal of the American Chemical Society* **1994**, *116*, 12073-12074; h) S.-L. Tang, L. B. Linz, B. C. Bonning, N. L. B. Pohl, *The Journal of Organic Chemistry* **2015**, *80*, 10482-10489.

[55] a) D. Senf, C. Ruprecht, G. H. M. de Kruijff, S. O. Simonetti, F. Schuhmacher, P. H. Seeberger, F. Pfrengle, *Chemistry – A European Journal* **2017**, *23*, 3197-3205; b) K. Naresh, F. Schumacher, H. S. Hahm, P. H. Seeberger, *Chemical Communications* **2017**, *53*, 9085-9088; c) M. W. Weishaupt, H. S. Hahm, A. Geissner, P. H. Seeberger, *Chemical Communications* **2017**, *53*, 3591-3594; d) H. S. Hahm, M. Hurevich, P. H. Seeberger, *Nature Communications* **2016**, *7*, 12482; e) M. Hurevich, P. H. Seeberger, *Chemical Communications* **2014**, *50*, 1851-1853; f) J. Kandasamy, M. Hurevich, P. H. Seeberger, *Chemical Communications* **2013**, *49*, 4453-4455; g) S. Eller, M. Collot, J. Yin, H. S. Hahm, P. H. Seeberger, *Angewandte Chemie International Edition* **2013**, *52*, 5858-5861; h) L. Kröck, D. Esposito, B. Castagner, C.-C. Wang, P. Bindschädler, P. H. Seeberger, *Chemical Science* **2012**, *3*, 1617-1622.

[56] a) H. Tanaka, Y. Nishiura, T. Takahashi, *Journal of the American Chemical Society* **2006**, *128*, 7124-7125; b) D. Crich, W. Li, *The Journal of Organic Chemistry* **2007**, *72*, 2387-2391; c) D. Crich, B. Wu, *Organic Letters* **2008**, *10*, 4033-4035.

[57] Y. Hashimoto, S. Tanikawa, R. Saito, K. Sasaki, *Journal of the American Chemical Society* **2016**, *138*, 14840-14843.

[58] a) H.-T. Chiu, Y.-C. Lin, M.-N. Lee, Y.-L. Chen, M.-S. Wang, C.-C. Lai, *Molecular BioSystems* **2009**, *5*, 1192-1203; b) Q. Gao, C. Zhang, S. Blanchard, J. S. Thorson, *Chemistry & Biology* **2006**, *13*, 733-743; c) C. Sánchez, C. Méndez, J. A. Salas, *Natural Product Reports* **2006**, *23*, 1007-1045; d) C. Zhang, C. Albermann, X. Fu, N. R. Peters, J. D. Chisholm, G. Zhang, E. J. Gilbert, P. G. Wang, D. L. Van Vranken, J. S. Thorson, *ChemBioChem* **2006**, *7*, 795-804; e) A. P. Salas, L. Zhu, C. Sánchez, A. F. Braña, J. Rohr, C. Méndez, J. A. Salas, *Molecular Microbiology* **2005**, *58*, 17-27; f) C.-G. Hyun, T. Bililign, J. Liao, J. S. Thorson, *ChemBioChem* **2003**, *4*, 114-117; g) C. Sánchez, I. A. Butovich, A. F. Braña, J. Rohr, C. Méndez, J. A. Salas, *Chemistry & Biology* **2002**, *9*, 519-531; h) T. Ohuchi, A. Ikeda-Araki, A. Watanabe-Sakamoto, K. Kojiri, M. Nagashima, M. Okanishi, H. Suda, *The Journal of Antibiotics* **2000**, *53*, 393-403; i) Z. Guo, J. Li, H. Qin, M. Wang, X. Lv, X. Li, Y. Chen, *Angewandte Chemie International Edition* **2015**, *54*, 5175-5178; j) J. A. Gawthorne, N. Y. Tan, U.-M. Bailey, M. R. Davis, L. W. Wong, R. Naidu, K. L. Fox, M. P. Jennings, B. L. Schulz, *Biochemical and Biophysical Research Communications* **2014**, *445*, 633-638; k) A. Naegeli, C.-W. Lin, M. Aebi, G. Michaud, T. Darbre, J.-L. Reymond, M. Schubert, C. Lizak, *The Journal of Biological Chemistry* **2014**, *289*, 24521-24532; l) K.-J. Choi, S. Grass, S. Paek, J. W. St. Geme, III, H.-J. Yeo, *PLOS ONE* **2011**, *5*, e15888; m) S. Grass, C. F. Lichti, R. R. Townsend, J. Gross, J. W. St. Geme, III, *PLOS Pathogens* **2010**, *6*, e1000919; n) P. Zhao, L. Bai, J. Ma, Y. Zeng, L. Li, Y. Zhang, C. Lu, H. Dai, Z. Wu, Y. Li, X. Wu, G. Chen, X. Hao, Y. Shen, Z. Deng, H. G. Floss, *Chemistry & Biology* **2008**, *15*, 863-874; o) N. A. Magarvey, B. Haltli, M. He, M. Greenstein, J. A. Hucul, *Antimicrobial Agents and Chemotherapy* **2006**, *50*, 2167-2177.

[59] a) D. Chen, R. Chen, R. Wang, J. Li, K. Xie, C. Bian, L. Sun, X. Zhang, J. Liu, L. Yang, F. Ye, X. Yu, J. Dai, *Angewandte Chemie International Edition* **2015**, *54*, 12678-12682; b) D. Foshag, C. Campbell, P. D. Pawelek, *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **2014**, *1844*, 1619-1630; c) A. Gutmann, C. Krump, L. Bungaruang, B. Nidetzky, *Chemical Communications* **2014**, *50*, 5465-5468; d) A. Panek, O. Pietrow, J. Synowiecki, P. Filipkowski, *Food and Bioproducts Processing* **2013**, *91*, 632-637; e) T.-I. Tsai, H.-Y. Lee, S.-H. Chang, C.-H. Wang, Y.-C. Tu, Y.-C. Lin, D.-R. Hwang, C.-Y. Wu, C.-H. Wong, *Journal of the American Chemical Society* **2013**, *135*, 14831-14839; f) A. Gutmann, B. Nidetzky, *Angewandte Chemie International Edition* **2012**, *51*, 12879-12883; g) J. Härle, S. Günther, B. Lauinger, M. Weber, B. Kammerer, David L. Zechel, A. Luzhetskyy, A. Bechthold, *Chemistry & Biology* **2011**, *18*, 520-530; h) M. Mittler, A. Bechthold, G. E. Schulz, *Journal of Molecular Biology* **2007**, *372*, 67-76; i) I. Baig, M. Kharel, A. Kobylyanskyy, L. Zhu, Y. Rebets, B. Ostash, A. Luzhetskyy, A. Bechthold, V. A. Fedorenko, J. Rohr, *Angewandte Chemie International Edition* **2006**, *45*, 7842-7846; j) T. Liu, M. K. Kharel, C. Fischer, A. McCormick, J. Rohr, *ChemBioChem* **2006**, *7*, 1070-1077; k) M. A. Fischbach, H. Lin, D. R. Liu, C. T. Walsh, *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 571-576; l) T. Bililign, C.-G. Hyun, J. S. Williams, A. M. Czisny, J. S. Thorson, *Chemistry & Biology* **2004**, *11*, 959-969.

[60] a) J. Kopycki, E. Wieduwild, J. Kohlschmidt, W. Brandt, Anna N. Stepanova, Jose M. Alonso, M. Soledade C. Pedras, S. Abel, C. D. Grubb, *Biochemical Journal*

2013, *450*, 37-46; b) H. Wang, W. A. van der Donk, *Journal of the American Chemical Society* **2011**, *133*, 16394-16397; c) M. Almendros, D. Danalev, M. François-Heude, P. Loyer, L. Legentil, C. Nugier-Chauvin, R. Daniellou, V. Ferrières, *Organic & Biomolecular Chemistry* **2011**, *9*, 8371-8378; d) C. Douglas Grubb, B. J. Zipp, J. Ludwig-Müller, M. N. Masuno, T. F. Molinski, S. Abel, *The Plant Journal* **2004**, *40*, 893-908.

[61] a) A. D. Calderon, J. Zhou, W. Guan, Z. Wu, Y. Guo, J. Bai, Q. Li, P. G. Wang, J. Fang, L. Li, *Organic & Biomolecular Chemistry* **2017**, *15*, 7258-7262; b) I. A. Gagarinov, T. Li, J. S. Toraño, T. Caval, A. D. Srivastava, J. A. W. Kruijtzer, A. J. R. Heck, G.-J. Boons, *Journal of the American Chemical Society* **2017**, *139*, 1011-1018; c) Z. Xiao, Y. Guo, Y. Liu, L. Li, Q. Zhang, L. Wen, X. Wang, S. M. Kondengaden, Z. Wu, J. Zhou, X. Cao, X. Li, C. Ma, P. G. Wang, *The Journal of Organic Chemistry* **2016**, *81*, 5851-5865.

[62] C. Zhang, B. R. Griffith, Q. Fu, C. Albermann, X. Fu, I.-K. Lee, L. Li, J. S. Thorson, *Science* **2006**, *313*, 1291-1294.

[63] a) R. W. Gantt, P. Peltier-Pain, W. J. Cournoyer, J. S. Thorson, *Nature Chemical Biology* **2011**, *7*, 685; b) R. W. Gantt, J. S. Thorson, in *Methods in Enzymology, Vol. 516* (Ed.: D. A. Hopwood), Academic Press, **2012**, pp. 345-360; c) P. Peltier-Pain, K. Marchillo, M. Zhou, D. R. Andes, J. S. Thorson, *Organic Letters* **2012**, *14*, 5086-5089; d) M. Zhou, A. Hamza, C.-G. Zhan, J. S. Thorson, *Journal of Natural Products* **2013**, *76*, 279-286; e) R. W. Gantt, P. Peltier-Pain, S. Singh, M. Zhou, J. S. Thorson, *Proceedings of the National Academy of Sciences* **2013**, *110*, 7648-7653.

[64] P. S. Jourdan, R. L. Mansell, *Archives of Biochemistry and Biophysics* **1982**, *213*, 434-443.

[65] J. Shearer, T. E. Graham, *Exercise and Sport Sciences Reviews* **2004**, *32*, 120-126.

[66] a) M. L. Mavrovouniotis, *Journal of Biological Chemistry* **1991**, *266*, 14440-14445; b) M. D. Jankowski, C. S. Henry, L. J. Broadbelt, V. Hatzimanikatis, *Biophysical Journal* **2008**, *95*, 1487-1499.

[67] a) A. Bar-Even, A. Flamholz, D. Davidi, E. Noor, R. Milo, Y. Lubling, *Bioinformatics* **2012**, *28*, 2037-2044; b) E. Noor, H. S. Haraldsdóttir, R. Milo, R. M. T. Fleming, *PLOS Computational Biology* **2013**, *9*, e1003098.

[68] R. N. Goldberg, T. N. Bhat, Y. B. Tewari, *Bioinformatics* **2004**, *20*, 2874-2877.

[69] S. Minakami, H. Yoshikawa, *Biochemical and Biophysical Research Communications* **1965**, *18*, 345-349.

[70] C. Held, G. Sadowski, *Annual Review of Chemical and Biomolecular Engineering* **2016**, *7*, 395-414.

[71] a) S.-i. Ozaki, H. Imai, T. Iwakiri, T. Sato, K. Shimoda, T. Nakayama, H. Hamada, *Biotechnology Letters* **2012**, *34*, 475-481; b) M. Yahyaa, R. Davidovich-Rikanati, Y. Eyal, A. Sheachter, S. Marzouk, E. Lewinsohn, M. Ibdah, *Phytochemistry* **2016**, *130*, 47-55; c) T. Zhang, J. Liang, P. Wang, Y. Xu, Y. Wang, X. Wei, M. Fan, *Scientific Reports* **2016**, *6*, 35274; d) E.-K. Lim, C. J. Doucet, Y. Li, L. Elias, D. Worrall, S. P. Spencer, J. Ross, D. J. Bowles, *Journal of Biological Chemistry* **2002**, *277*, 586-592; e) K. Tadera, F. Yagi, A. Kobayashi, *J. Nutr. Sci. Vitaminol.* **1982**, *28*, 359-366; f) J.-S. Lin, X.-X. Huang, Q. Li, Y. Cao, Y. Bao, X.-F. Meng, Y.-J. Li, C. Fu, B.-K. Hou, *The Plant Journal* **2016**, *88*, 26-42; g) R. K. Ibrahim, H. Grisebach, *Archives of Biochemistry and Biophysics* **1976**, *176*, 700-708; h) J. Hyung Ko, B. Gyu Kim, A. Joong-Hoon, *FEMS Microbiology Letters* **2006**, *258*, 263-268; i) M. Ostrowski, A. Hetmann, A. Jakubowska, *Phytochemistry* **2015**, *117*, 25-33; j) K. Tahara, M. Nishiguchi, A. Frolov, J. Mittasch, C. Milkowski,

Phytochemistry **2018**, *152*, 154-161; k) Q. Wang, Y. Xu, J. Xu, X. Wang, C. Shen, Y. Zhang, X. Liu, B. Yu, J. Zhang, *Biotechnology Letters* **2017**, *39*, 1229-1235; l) H. Marcinek, W. Weyler, B. Deus-Neumann, M. H. Zenk, *Phytochemistry* **2000**, *53*, 201-207.

[72] a) S. S. Mandal, G. Liao, Z. Guo, *RSC Advances* **2015**, *5*, 23311-23319; b) D. S. Frear, *Phytochemistry* **1968**, *7*, 381-390.

[73] R. C. Martin, M. C. Mok, D. W. S. Mok, *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 284-289.

[74] a) C. Landmann, B. Fink, W. Schwab, *Planta* **2007**, *226*, 417-428; b) S. Lunkenbein, M. Bellido, A. Aharoni, E. M. J. Salentijn, R. Kaldenhoff, H. A. Coiner, J. Muñoz-Blanco, W. Schwab, *Plant Physiology* **2006**, *140*, 1047-1058.

[75] a) L. Trobo-Maseda, A. H. Orrego, S. Moreno-Pérez, G. Fernández-Lorente, J. M. Guisan, J. Rocha-Martin, *Applied Microbiology and Biotechnology* **2018**, *102*, 773-787; b) G. Dewitte, M. Walmagh, M. Diricks, A. Lepak, A. Gutmann, B. Nidetzky, T. Desmet, *Journal of Biotechnology* **2016**, *233*, 49-55.

[76] A. Gutmann, B. Nidetzky, *Advanced Synthesis & Catalysis* **2016**, *358*, 3600-3609.

[77] S.-I. Ryu, C.-S. Park, J. Cha, E.-J. Woo, S.-B. Lee, *Biochemical and Biophysical Research Communications* **2005**, *329*, 429-436.

[78] Y. B. Tewari, R. N. Goldberg, *Biophysical Chemistry* **1991**, *40*, 59-67.

[79] Q. Qu, S.-J. Lee, W. Boos, *Journal of Biological Chemistry* **2004**, *279*, 47890-47897.

[80] T. Kouril, M. Zaparty, J. Marrero, H. Brinkmann, B. Siebers, *Archives of Microbiology* **2008**, *190*, 355.

[81] C. Tian, J. Yang, Y. Zeng, T. Zhang, Y. Zhou, Y. Men, C. You, Y. Zhu, Y. Sun, *Applied and Environmental Microbiology* **2019**, *85*, e02306-02318.

[82] S. M. Resnick, A. J. B. Zehnder, *Applied and Environmental Microbiology* **2000**, *66*, 2045-2051.

[83] a) D. C. Crans, G. M. Whitesides, *The Journal of Organic Chemistry* **1983**, *48*, 3130-3132; b) G. Tasnádi, W. Jud, M. Hall, K. Baldenius, K. Ditrich, K. Faber, *Advanced Synthesis & Catalysis* **2018**, *360*, 2394-2401.

[84] S. T. Kulmer, A. Gutmann, M. Lemmerer, B. Nidetzky, *Advanced Synthesis & Catalysis* **2017**, *359*, 292-301.

[85] a) J. N. Andexer, M. Richter, *ChemBioChem* **2015**, *16*, 380-386; b) Z. Liu, J. Zhang, X. Chen, P. G. Wang, *ChemBioChem* **2002**, *3*, 348-355; c) K. Murata, T. Uchida, J. Kato, I. Chibata, *Agricultural and Biological Chemistry* **1988**, *52*, 1471-1477; d) E. Restiawaty, Y. Iwasa, S. Maya, K. Honda, T. Omasa, R. Hirota, A. Kuroda, H. Ohtake, *Process Biochemistry* **2011**, *46*, 1747-1752; e) M. Sato, Y. Masuda, K. Kirimura, K. Kino, *Journal of Bioscience and Bioengineering* **2007**, *103*, 179-184; f) S. Iwamoto, K. Motomura, Y. Shinoda, M. Urata, J. Kato, N. Takiguchi, H. Ohtake, R. Hirota, A. Kuroda, *Applied and Environmental Microbiology* **2007**, *73*, 5676-5678.

[86] a) T. Shiba, K. Tsutsumi, K. Ishige, T. Noguchi, *Biochemistry (Moscow)* **2000**, *65*, 315-323; b) T. Noguchi, T. Shiba, *Bioscience, Biotechnology, and Biochemistry* **1998**, *62*, 1594-1596.

[87] a) J. Nahálka, V. Pätoprstý, *Organic & Biomolecular Chemistry* **2009**, *7*, 1778-1780; b) S.-G. Lee, J.-O. Lee, J.-K. Yi, B.-G. Kim, *Biotechnology and Bioengineering* **2002**, *80*, 516-524.

[88] C. Rupprath, M. Kopp, D. Hirtz, R. Müller, L. Elling, *Advanced Synthesis & Catalysis* **2007**, *349*, 1489-1496.

[89] T. Fischöder, C. Wahl, C. Zerhusen, L. Elling, *Biotechnology Journal* **2019**, *14*.

[90] L. Wen, Y. Zheng, T. Li, P. G. Wang, *Bioorganic & Medicinal Chemistry Letters* **2016**, *26*, 2825-2828.
[91] J.-H. Lee, S.-W. Chung, H.-J. Lee, K.-S. Jang, S.-G. Lee, B.-G. Kim, *Bioprocess and Biosystems Engineering* **2009**, *33*, 71.
[92] C. Unverzagt, H. Kunz, J. C. Paulson, *Journal of the American Chemical Society* **1990**, *112*, 9308-9309.
[93] B. L. Hirschbein, F. P. Mazenod, G. M. Whitesides, *The Journal of Organic Chemistry* **1982**, *47*, 3765-3766.
[94] H. Yu, A. Santra, Y. Li, J. B. McArthur, T. Ghosh, X. Yang, P. G. Wang, X. Chen, *Organic & Biomolecular Chemistry* **2018**, *16*, 4076-4080.
[95] M. M. Muthana, J. Qu, M. Xue, T. Klyuchnik, A. Siu, Y. Li, L. Zhang, H. Yu, L. Li, P. G. Wang, X. Chen, *Chemical Communications* **2015**, *51*, 4595-4598.
[96] J. A. Read, R. A. Ahmed, M. E. Tanner, *Organic Letters* **2005**, *7*, 2457-2460.
[97] A. Alissandratos, K. Caron, T. D. Loan, J. E. Hennessy, C. J. Easton, *ACS Chemical Biology* **2016**, *11*, 3289-3293.
[98] C. H. Wong, S. L. Haynie, G. M. Whitesides, *The Journal of Organic Chemistry* **1982**, *47*, 5416-5418.
[99] F. Tomoike, N. Nakagawa, S. Kuramitsu, R. Masui, *Biochemistry* **2011**, *50*, 4597-4607.
[100] T. D. Loan, C. J. Easton, A. Alissandratos, *New Biotechnology* **2019**, *49*, 104-111.
[101] a) R. A. Scism, B. O. Bachmann, *ChemBioChem* **2010**, *11*, 67-70; b) J. Fernández-Lucas, *Applied Microbiology and Biotechnology* **2015**, *99*, 4615-4627.
[102] L. Glaser, D. H. Brown, *Journal of Biological Chemistry* **1957**, *228*, 729-742.
[103] B. Lougheed, H. D. Ly, W. W. Wakarchuk, S. G. Withers, *Journal of Biological Chemistry* **1999**, *274*, 37717-37722.
[104] A. Lepak, A. Gutmann, B. Nidetzky, *ACS Catalysis* **2018**, *8*, 9148-9153.
[105] L. L. Lairson, W. W. Wakarchuk, S. G. Withers, *Chemical Communications* **2007**, 365-367.
[106] a) C. Li, Z. Zhang, Q. Duan, X. Li, *Organic Letters* **2014**, *16*, 3008-3011; b) R. R. Hughes, K. A. Shaaban, J. Zhang, H. Cao, G. N. Phillips, J. S. Thorson, *ChemBioChem* **2017**, *18*, 363-367; c) A. A. Lee, Y. C. S. Chen, E. Ekalestari, S. Y. Ho, N. S. Hsu, T. F. Kuo, T. S. A. Wang, *Angewandte Chemie International Edition* **2016**, *55*, 12338-12342.
[107] F. Burkhart, Z. Zhang, S. Wacowich-Sgarbi, C. H. Wong, *Angewandte Chemie International Edition* **2001**, *40*, 1274-1277.
[108] C.-Y. Huang, D. A. Thayer, A. Y. Chang, M. D. Best, J. Hoffmann, S. Head, C.-H. Wong, *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103*, 15-20.
[109] D. M. Su, H. Eguchi, W. Yi, L. Li, P. G. Wang, C. Xia, *Organic Letters* **2008**, *10*, 1009-1012.
[110] H. Yu, Y. Li, J. Zeng, V. Thon, D. M. Nguyen, T. Ly, H. Y. Kuang, A. Ngo, X. Chen, *The Journal of Organic Chemistry* **2016**, *81*, 10809-10824.
[111] H. Yu, K. Lau, Y. Li, G. Sugiarto, X. Chen, *Current Protocols in Chemical Biology* **2012**, *4*, 233-247.
[112] S.-I. Nishimura, K. Yamada, *Journal of the American Chemical Society* **1997**, *119*, 10555-10556.
[113] X. Huang, Krista L. Witte, David E. Bergbreiter, C.-H. Wong, *Advanced Synthesis & Catalysis* **2001**, *343*, 675-681.
[114] a) F. A. Jaipuri, N. L. Pohl, *Organic & Biomolecular Chemistry* **2008**, *6*, 2686-2691; b) B. Yang, Y. Jing, X. Huang, *European Journal of Organic Chemistry* **2010**, *2010*, 1290-1298; c) L. Liu, N. L. B. Pohl, *Organic Letters* **2011**, *13*, 1824-1827; d) F. R.

Carrel, K. Geyer, J. D. C. Codée, P. H. Seeberger, *Organic Letters* **2007**, *9*, 2285-2288; e) F. R. Carrel, P. H. Seeberger, *The Journal of Organic Chemistry* **2008**, *73*, 2058-2065; f) F. Zhang, W. Zhang, Y. Zhang, D. P. Curran, G. Liu, *The Journal of Organic Chemistry* **2009**, *74*, 2594-2597; g) W. Huang, Q. Gao, G.-J. Boons, *Chemistry – A European Journal* **2015**, *21*, 12920-12926; h) G. Macchione, J. L. de Paz, P. M. Nieto, *Carbohydrate Research* **2014**, *394*, 17-25; i) Y.-H. Chai, Y.-L. Feng, J.-J. Wu, C.-Q. Deng, A.-Y. Liu, Q. Zhang, *Chinese Chemical Letters* **2017**, *28*, 1693-1700; j) J. Hwang, H. Yu, H. Malekan, G. Sugiarto, Y. Li, J. Qu, V. Nguyen, D. Wu, X. Chen, *Chemical Communications* **2014**, *50*, 3159-3162; k) H. Tanaka, Y. Tanimoto, T. Kawai, T. Takahashi, *Tetrahedron* **2011**, *67*, 10011-10016.

[115] H. Zhu, Z. Wu, M. R. Gadi, S. Wang, Y. Guo, G. Edmunds, W. Guan, J. Fang, *Bioorganic & Medicinal Chemistry Letters* **2017**, *27*, 4285-4287.

[116] A. Santra, Y. Li, H. Yu, T. J. Slack, P. G. Wang, X. Chen, *Chemical Communications* **2017**, *53*, 8280-8283.

[117] J. Zhang, C. Chen, M. R. Gadi, C. Gibbons, Y. Guo, X. Cao, G. Edmunds, S. Wang, D. Liu, J. Yu, L. Wen, P. G. Wang, *Angewandte Chemie International Edition* **2018**, *57*, 16638-16642.

[118] a) M. Schuster, P. Wang, J. C. Paulson, C.-H. Wong, *Journal of the American Chemical Society* **1994**, *116*, 1135-1136; b) R. L. Halcomb, H. Huang, C.-H. Wong, *Journal of the American Chemical Society* **1994**, *116*, 11315-11322.

[119] a) O. Blixt, T. Norberg, *Carbohydrate Research* **1999**, *319*, 80-91; b) B. T. Houseman, M. Mrksich, *Angewandte Chemie International Edition* **1999**, *38*, 782-785; c) K. Yamada, E. Fujita, S.-I. Nishimura, *Carbohydrate Research* **1997**, *305*, 443-461; d) F. Yan, W. W. Wakarchuk, M. Gilbert, J. C. Richards, D. M. Whitfield, *Carbohydrate Research* **2000**, *328*, 3-16.

[120] L. Wen, G. Edmunds, C. Gibbons, J. Zhang, M. R. Gadi, H. Zhu, J. Fang, X. Liu, Y. Kong, P. G. Wang, *Chemical Reviews* **2018**.

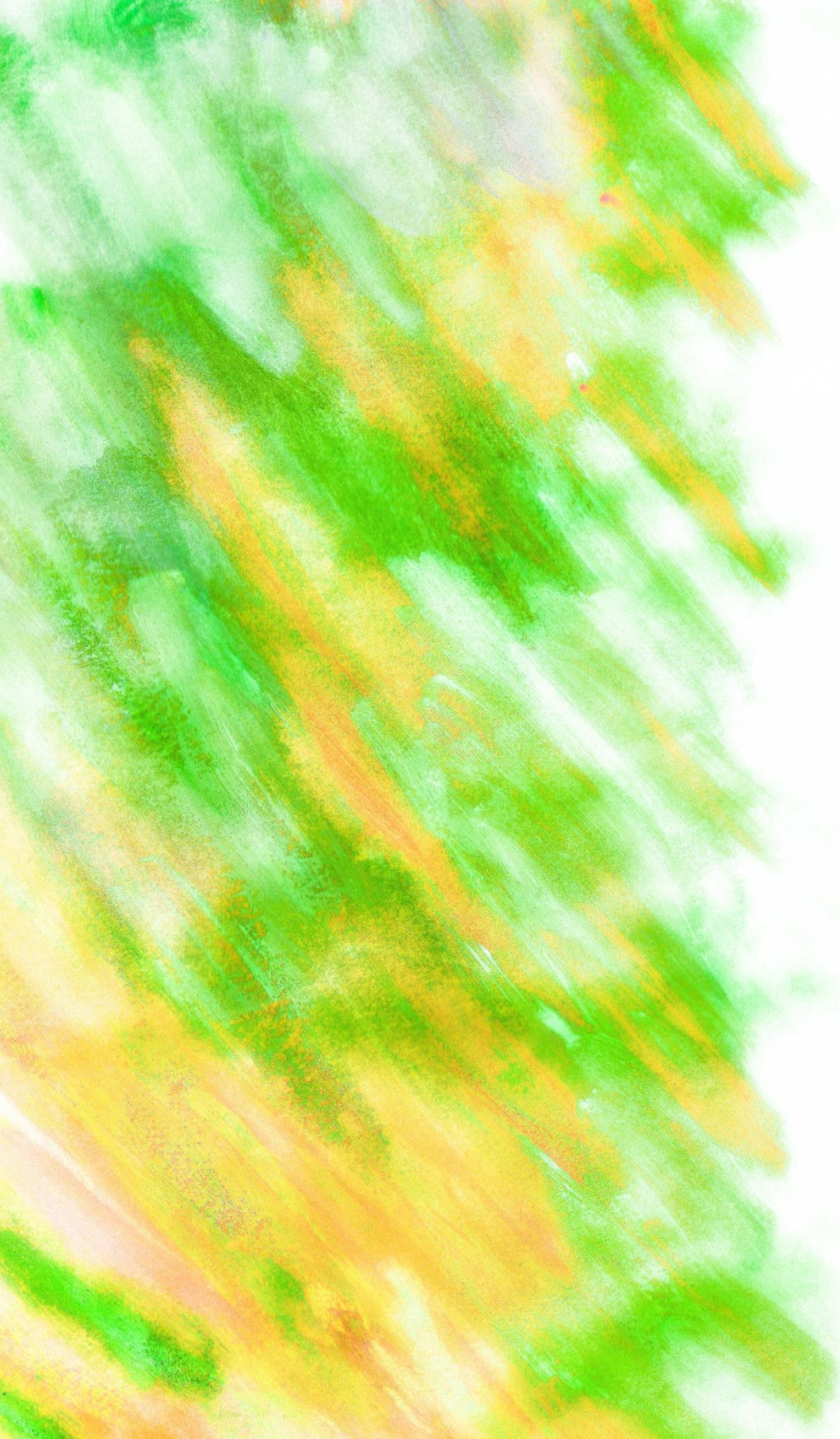
[121] a) P. Sears, C.-H. Wong, *Science* **2001**, *291*, 2344-2350; b) T. Ivannikova, F. Bintein, A. Malleron, S. Juliant, M. Cerutti, A. Harduin-Lepers, P. Delannoy, C. Augé, A. Lubineau, *Carbohydrate Research* **2003**, *338*, 1153-1161; c) C. Augé, R. Fernandez-Fernandez, C. Gautheron, *Carbohydrate Research* **1990**, *200*, 257-268; d) C. Augé, C. Gautheron, *Tetrahedron Letters* **1988**, *29*, 789-790; e) S. David, C. Auge, in *Pure and Applied Chemistry, Vol. 59*, **1987**, p. 1501; f) C. Augé, S. David, C. Mathieu, C. Gautheron, *Tetrahedron Letters* **1984**, *25*, 1467-1470; g) X. Chen, J. Fang, J. Zhang, Z. Liu, J. Shao, P. Kowal, P. Andreana, P. G. Wang, *Journal of the American Chemical Society* **2001**, *123*, 2081-2082; h) C.-C. Yu, Y.-Y. Kuo, C.-F. Liang, W.-T. Chien, H.-T. Wu, T.-C. Chang, F.-D. Jan, C.-C. Lin, *Bioconjugate Chemistry* **2012**, *23*, 714-724; i) J. Nahalka, Z. Liu, X. Chen, P. G. Wang, *Chemistry – A European Journal* **2003**, *9*, 372-377; j) P.-Y. Wang, S.-W. Tsai, T.-L. Chen, *Biotechnology and Bioengineering* **2008**, *101*, 460-469.

[122] a) Y. Zhang, J. Ge, Z. Liu, *ACS Catalysis* **2015**, *5*, 4503-4513; b) I. Chibata, T. Tosa, T. Sato, T. Mori, in *Methods in Enzymology, Vol. 44*, Academic Press, **1976**, pp. 746-759.

[123] a) Ž. Petronijević, S. Ristić, D. Pešić, A. Šmelcerović, *Enzyme and Microbial Technology* **2007**, *40*, 763-768; b) J. d. N. Schöffer, M. P. Klein, R. C. Rodrigues, P. F. Hertz, *Carbohydrate Polymers* **2013**, *98*, 1311-1316; c) J. d. N. Schöffer, C. R. Matte, D. S. Charqueiro, E. W. de Menezes, T. M. H. Costa, E. V. Benvenutti, R. C. Rodrigues, P. F. Hertz, *Process Biochemistry* **2017**, *58*, 120-127.

[124] J. Rakmai, B. Cheirsilp, P. Prasertsan, *Biocatalysis and Agricultural Biotechnology* **2015**, *4*, 717-726.

[125] D.-H. Jung, J.-H. Jung, D.-H. Seo, S.-J. Ha, D.-K. Kweon, C.-S. Park, *Bioresource Technology* **2013**, *130*, 801-804.
[126] A. H. Orrego, L. Trobo-Maseda, J. Rocha-Martin, J. M. Guisan, *Enzyme and Microbial Technology* **2017**, *105*, 51-58.
[127] K. De Winter, W. Soetaert, T. Desmet, *International Journal of Molecular Sciences* **2012**, *13*, 11333.
[128] a) J. Kaulpiboon, P. Pongsawasdi, W. Zimmermann, *The FEBS Journal* **2007**, *274*, 1001-1010; b) J. Sun, S. Wang, W. Li, R. Li, S. Chen, H. I. Ri, T. M. Kim, M. S. Kang, L. Sun, X. Sun, Q. Yuan, *Molecules* **2018**, *23*, 1087.
[129] a) R. A. Sheldon, S. van Pelt, *Chemical Society Reviews* **2013**, *42*, 6223-6235; b) A. Liese, L. Hilterhaus, *Chemical Society Reviews* **2013**, *42*, 6236-6249.
[130] U. Hanefeld, L. Gardossi, E. Magner, *Chemical Society Reviews* **2009**, *38*, 453-468.
[131] a) J. Rakmai, B. Cheirsilp, *Biochemical Engineering Journal* **2016**, *105*, 107-113; b) Y.-J. Cho, O.-J. Park, H.-J. Shin, *Enzyme and Microbial Technology* **2006**, *39*, 108-113.
[132] K. Szymańska, K. Odrozek, A. Zniszczoł, W. Pudło, A. B. Jarzębski, *Chemical Engineering Journal* **2017**, *315*, 18-24.
[133] G. Cattaneo, M. Rabuffetti, G. Speranza, T. Kupfer, B. Peters, G. Massolini, D. Ubiali, E. Calleri, *ChemCatChem* **2017**, *9*, 4614-4620.
[134] a) K. Szymańska, K. Odrozek, A. Zniszczoł, G. Torrelo, V. Resch, U. Hanefeld, A. B. Jarzębski, *Catalysis Science & Technology* **2016**, *6*, 4882-4888; b) K. Szymańska, W. Pudło, J. Mrowiec-Białoń, A. Czardybon, J. Kocurek, A. B. Jarzębski, *Microporous and Mesoporous Materials* **2013**, *170*, 75-82.
[135] J. Lawrence, B. O'Sullivan, G. J. Lye, R. Wohlgemuth, N. Szita, *Journal of Molecular Catalysis B: Enzymatic* **2013**, *95*, 111-117.
[136] D. Valikhani, J. M. Bolivar, M. Pfeiffer, B. Nidetzky, *ChemCatChem* **2017**, *9*, 161-166.
[137] D. J. Strub, K. Szymańska, Z. Hrydziuszko, J. Bryjak, A. B. Jarzębski, *Reaction Chemistry & Engineering* **2019**.
[138] R. A. Sheldon, J. M. Woodley, *Chemical Reviews* **2018**, *118*, 801-838.
[139] a) C.-J. Aurell, S. Karlsson, F. Pontén, S. M. Andersen, *Organic Process Research & Development* **2014**, *18*, 1116-1119; b) H. Mallin, J. Muschiol, E. Byström, U. T. Bornscheuer, *ChemCatChem* **2013**, *5*, 3529-3532; c) J. J. Carberry, *Industrial & Engineering Chemistry* **1964**, *56*, 39-46.
[140] a) J. M. Woodley, N. J. Titchener-Hooker, *Bioprocess Engineering* **1996**, *14*, 263-268; b) R. Wohlgemuth, I. Plazl, P. Žnidaršič-Plazl, K. V. Gernaey, J. M. Woodley, *Trends in Biotechnology* **2015**, *33*, 302-314; c) I. Rossetti, *Catalysis Today* **2018**, *308*, 20-31.



3

Artificial Fusion of mCherry Enhanced Solubility and Stability of Trehalose Transferase

This chapter is based on:

L. Mestrom, Stefan R. Marsden, M. Dieters, P. Achterberg, L. Stolk, I. Bento, U. Hanefeld, P.L. Hagedoorn. *Appl Environ Microbiol*., **2019**, 85 (8), e03084-18 Correction**:** L. Mestrom, Stefan R. Marsden, M. Dieters, P. Achterberg, L. Stolk, I. Bento, U. Hanefeld, P.L. Hagedoorn. *Appl Environ Microbiol*., **2019**, 85 (14), e00942-19

"Science, my lad, is made up of mistakes, but they are mistakes which it is useful to make, because they lead little by little to the truth."

Jules Verne*, A Journey to the Center of the Earth*, **1864**

Preface

After the publication of the article titled "*Artificial Fusion of mCherry Enhanced Solubility and Stability of Trehalose Transferase*" (vol 85, issue 8, e03084-18; manuscript AEM03084-18) an error was discovered in April 2019. The presented trehalose transferase from *Pyrobaculum yellowstonensis* was mislabeled (mCherry *Py*TreT) in our data file register. The correct name for the given protein sequence fused to mCherry was trehalose transferase from *Thermoproteus uzoniensis* (mCherry *Tu*TreT). Only the presented name was incorrect throughout the article. This is evident from the sequence we presented in the supplementary information for the mCherry-*Py*TreT protein in the original article, which is in fact the correct sequence of mCherry-*Tu*TreT. This mistake did not alter the main conclusion of the article. Upon discovery we have filed a correction in May 2019, which was published on July 2019 as "*Correction for Mestrom et al., "Artificial Fusion of mCherry Enhances Trehalose Transferase Solubility and Stability*" (vol 85, issue 14, e00942-19). For reasons of clarity, **chapter 3** is presented with the correct name and appropriate protein sequences have additionally been attached.

3.1 Introduction

Trehalose (α-D-glucopyranosyl-(1→1)-α-D-glucopyranoside) is a non-reducing disaccharide with an α,α-glycosidic linkage that has been identified in plants, insects, fungi, yeast, bacteria, and archaea[1]. The functional role of trehalose as an intracellular osmolyte is to manage the cell volume during exposure to intra- or extracellular osmotic-, thermal-, and oxidative stresses. Trehalose is a non-ionic kosmotrope which preserves the protein hydration-shell by reducing the water activity a_w[2]. Moreover, during anhydrobiosis trehalose protects the cell membranes by direct binding to phospholipids, preventing water leakage during rehydration[3]. Due to the absence of a free aldehyde moiety, trehalose is highly resistant to heat, changes in pH, and does not degrade *via* the Maillard reaction[4]. Unsurprisingly, trehalose is commonly found in extremophiles which have to withstand harsh growth conditions such as extreme temperatures, high ionic strengths and acidic or basic environments[1a].

Several metabolic pathways for the biosynthesis of trehalose have been found in nature (Fig. 1) and include: (i) trehalose synthase (TreS) interconverting maltose to trehalose[5]; (ii) maltooligosyltrehalose synthase (TreYZ) hydrolyzing maltodextrins to trehalose[6]; (iii) inverting trehalose phosphorylase ($TreP_{inv}$)[7] adding α-D-glucose-1-phosphate or (iv) retaining trehalose phosphorylase ($TreP_{ret}$)[8] adding β-D-glucose-1-phosphate to glucose producing trehalose and phosphate; (v) trehalose transferase (TreT) using D-glucose and a nucleotide diphosphate (NDP) sugar to produce D-trehalose[9]; (vi) trehalose phosphate synthase (OtsA) producing D-trehalose-6-phosphate from D-glucose-6-phosphate and a nucleotide sugar[10]. In comparison to trehalose phosphate synthase, the LeLoir glycosyltransferase TreT does not require the use of additional 6-phosphate (OtsA) avoiding sequential dephosphorylation of the non-reducing disaccharides, and therefore is of particular interest for industrial food applications[11].

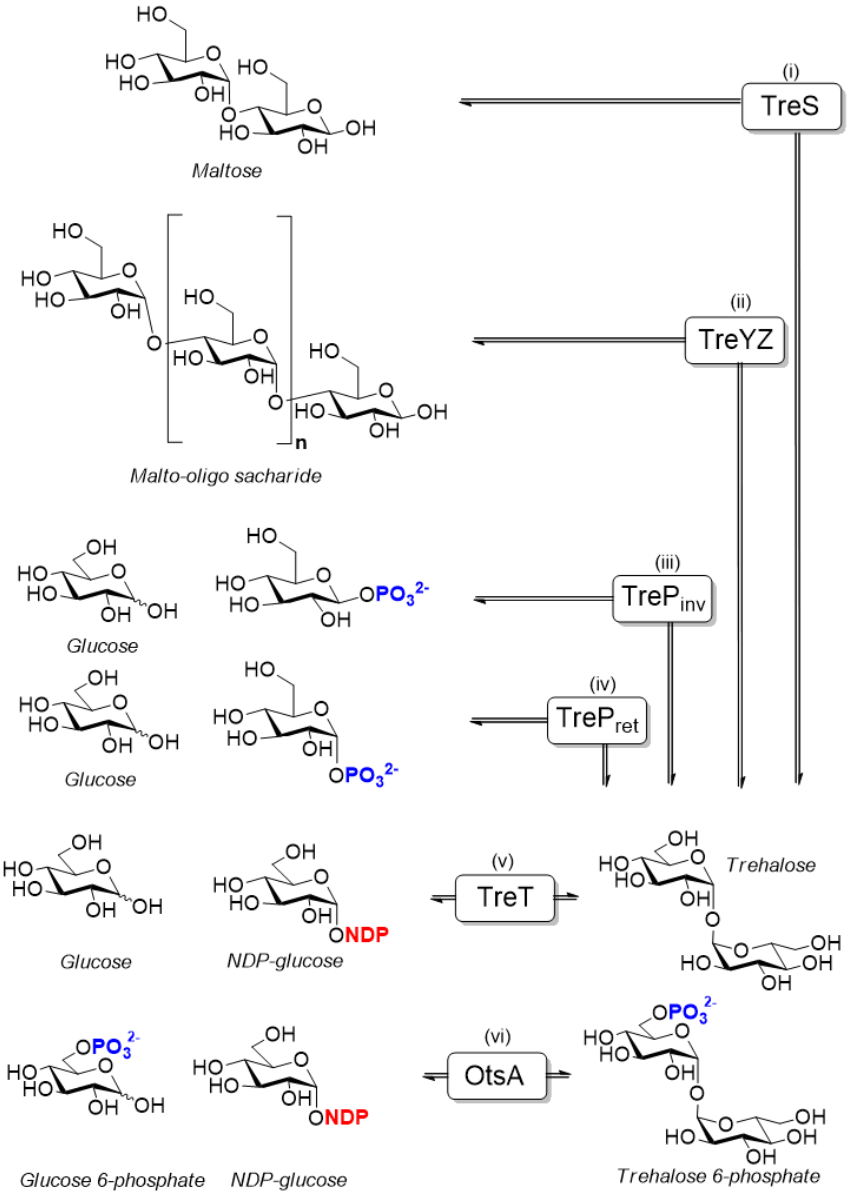


Figure 1. Biosynthesis of trehalose using trehalose isomerase (TreS), trehalose hydrolase/isomerase (TreYZ), inverting ($TreP_{inv}$) and retaining ($TreP_{ret}$) phosphorylases, trehalose transferase (TreT), and trehalose phosphate synthase (OtsA).

TreT from *Thermoproteus tenax* has previously been applied in enzymatic sugar coupling for the production of non-natural trehalose derivatives, but the variation of sugar acceptors was limited[12]. Currently, the main limitations for biotechnological applications of TreT are the low protein stability, solubility, and formation of inclusion bodies (IBs) during heterologous expression in *E. coli* [9b]. The formation of inclusion bodies is reported for 70% of all recombinant proteins[13], constituting as one of the major obstacles for heterologous expression systems emphasizing the requirement to use solubility tags. Despite these challenges trehalose transferase has been recognized for its high total turnover number during catalysis[9a, 14]. The aim of this study was to express a stable, robust trehalose transferase expression system for the chemical synthesis of trehalose derivatives. Since protein folding and aggregation are governed by hydrophobic and electrostatic interactions, the aggregation-prone behavior of several homologous trehalose transferases with different pI values from hyperthermophilic Crenarchaeota was investigated in order to address this issue. For this purpose, the TreT from *Pyrobaculum yellowstonensis* WP-30 (*Py*TreT), *Thermoproteus tenax* Kra1 (*Tt*TreT), and *Thermoproteus uzoniensis* 768-20 (*Tu*TreT) were selected.

Additionally, the fluorescent protein mCherry was fused to *Tu*TreT as a direct reporter for promoting protein solubility. Next to a short maturation time and an excellent photostability of the chromophore, the fluorescent protein possesses the required thermostable properties to match its fusion partner. While the complete mechanisms behind protein aggregation and the formation of inclusion bodies remain elusive, the fusion of fluorescent proteins to aggregation-prone enzymes to monitor protein solubility has previously been successfully applied for a variety of enzymes[15]. The use of a fusion complex has remained largely limited towards the visualization of proteins *in vivo*, rather than enzyme catalysis *in vitro*. In this study, we explored mCherry fusion to the aggregation-prone trehalose transferase as a solubility enhancement tag to address the challenging recombinant expression of archaeal glycosyltransferases.

3.2 Results and Discussion

3.2.1 Recombinant expression of TreT

The recombinant expression of *Tt*TreT from the pET302 plasmid in *E. coli* BL21(DE3) was previously reported to lead to the formation of insoluble inclusion bodies (IBs)[9b]. Therefore, *Tt*TreT was expressed using the pBAD/His A plasmid in *E. coli* Top 10 based on previous results, where the formation of insoluble inclusion bodies was not reported[16]. The protein solubility of TreT was evaluated in the expression host *E. coli* Top10 using the pBAD/His A vector in parallel expression experiments, harboring an empty plasmid or the genes encoding for *Tt*TreT, *Tu*TreT, *Py*TreT. The cells were harvested, lysed, and evaluated on protein content, purity, and activity. High overexpression of TreT was observed in all cases, but the enzymes were predominantly present in the insoluble cell debris as inclusion bodies (IBs) (Fig. 2).

Attempts to optimize the expression by varying the concentration of inducing agent, a change to auto-induction medium, or lower expression temperatures did not afford higher yields of soluble target protein according to SDS-PAGE analysis. Nevertheless, small fractions of the TreTs were soluble and enzyme activities were measured with HPLC by monitoring the production of trehalose from glucose and UDP-glucose. Cell-free extracts (CFEs) of *Tt*TreT showed higher activity than those of *Tu*TreT and *Py*TreT (table 1). Background glycosyltransferase activity from the expression host was ruled out with control experiments containing the empty plasmid. In order to purify from the soluble fraction, His_6-tags were introduced at the *C*- or *N*-terminus of *Py*TreT. The variant with a *N*-terminal His_6-tag did not bind to the nickel sepharose resin, the *C*-terminal tagged variant could be purified and showed TreT activity (table 1). However, rapid precipitation after affinity purification resulted in fibrillar protein aggregates which could not be prevented by buffer exchange to HEPES (50 mM, pH 7.0), with the addition of 300 mM NaCl, or sodium phosphate (50 mM, pH 7.0). Direct measurement of TreT activity upon purification at high temperatures showed a

specific activity of 20.8 U mg^{-1}, but the continuous precipitation of protein under given experimental conditions needs to considered.

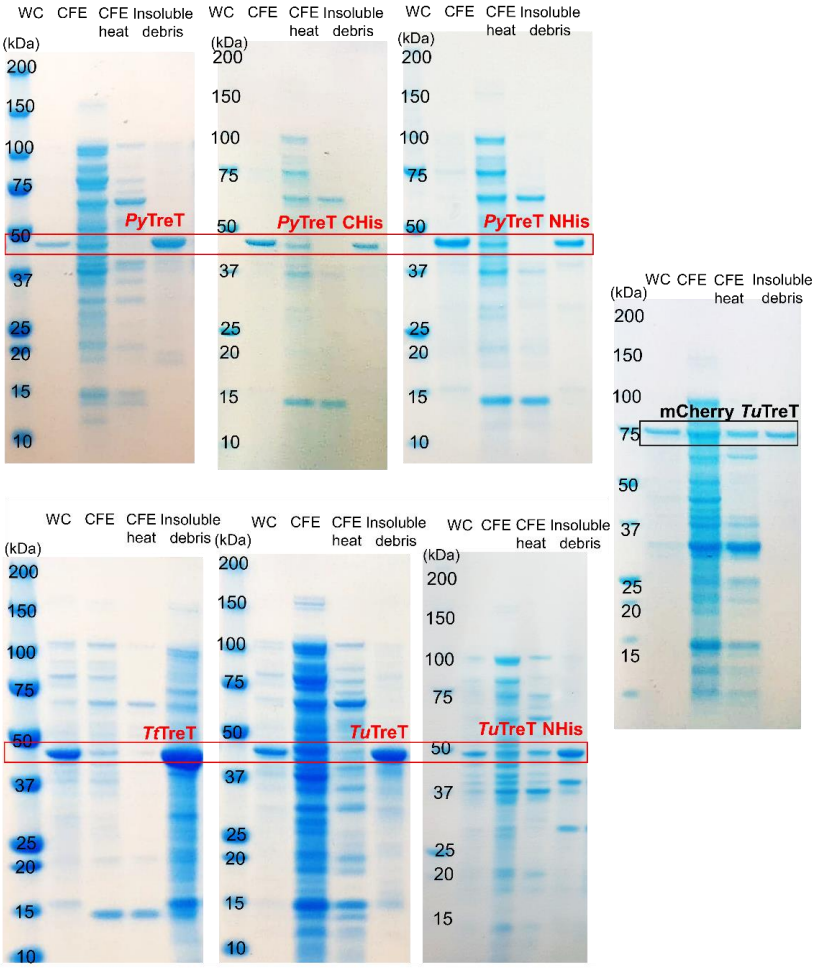


Figure 2: Comparison of whole cell expression and protein content in the soluble and insoluble fraction analysed with SDS-PAGE. No soluble *Tt*TreT, *Tu*TreT, *Py*TreT, or *Py*TreT with a *C*- or *N*-terminal HisTag was observed in the cell-free extract before and after heat-treatment (red box), while mCherry *Tu*TreT was highly soluble (black box). **Abbreviations**: WC: whole cells, CFE: Cell-free extract, CFE heat: after heat-treatment of cell-free extract, Insoluble debris: insoluble pellet after cell lysis.

Table 1: Specific activity of TreT of 400 mL of cell culture during purification. [a]not detected. [b]not applicable

Description	CFE (U mg^{-1})	Heat-treatment (U mg^{-1})	IB (A) (U mg^{-1})	IMAC (U mg^{-1})
*Tt*TreT	0.754	0.591	0.115	n.a.[b]
*Tu*TreT	0.463	0.681	0.172	n.a.[b]
*Py*TreT	n.d.[a]	n.d.[a]	0.029	n.a.[b]
*Py*TreT CHis	n.d.[a]	n.d.[a]	0.050	trace[†]
*Py*TreT NHis	n.d.[a]	n.d.[a]	0.088	n.a.[b]
*Tu*TreT NHis	0.705	1.268	0.119	n.a.[b]
mCherry *Tu*TreT	0.214	0.543	0.179	5.06
pBAD/His A	n.d.	n.d.	n.d.	n.d.

[†]*Py*TreT CHis was grown at a large scale (6 x 1L of TB-media), purified, and an activity of 20.8 U mg^{-1} was measured. The purified enzyme was completely precipitated after 1 hour after the purification. **Reaction conditions**: D-glucose (20 mM), UDP-D-glucose (40 mM), $MgCl_2$ (20 mM), HEPES (50 mM, pH 7.0), 80°C.

Hyperthermostable proteins are known to show folding energy landscapes which are different from their mesophilic counterparts, leading more rapidly towards the formation of oligomers and aggregates during heterologous expression in a mesophilic host[17]. Naturally, trehalose transferase is expressed in response to intra- or extracellular osmotic-, thermal, and oxidative stresses, which coincides with the expression of a large number of heat-shock proteins and chaperones that assist in the correct folding of the native TreT protein structure within *Thermoproteus* and *Pyrobaculum*[18]. *E. coli* however does not contain the same set of chaperones which would naturally occur in Crenarcheaota[19] leading to misfolded protein[18a-e, 18g]. Furthermore, a high intracellular concentration of 0.10 mg trehalose per mg of protein was reported for *Thermoproteus tenax*[1a], which potentially could stabilize TreT in their native hosts and could explain the low protein stability observed *in vitro*.

3.2.2 Purified mCherry *Tu*TreT shows high protein solubility and stability

Due to the poor solubility of *Py*TreT, a fusion construct mCherry *Tu*TreT containing a *C*-terminal His_6-tag was produced. The fusion with mCherry enables the direct quantitative spectrophotometric determination of *Tu*TreT in solution, allowing for rapid solubility and expression assays. To our satisfaction, expression of the fusion protein resulted in increased solubility for mCherry *Tu*TreT. Typically, ≤ 10 mg of purified mCherry *Tu*TreT was isolated per liter of TB medium, and the formation of catalytically active IBs could not be avoided during expression (Fig. 3).

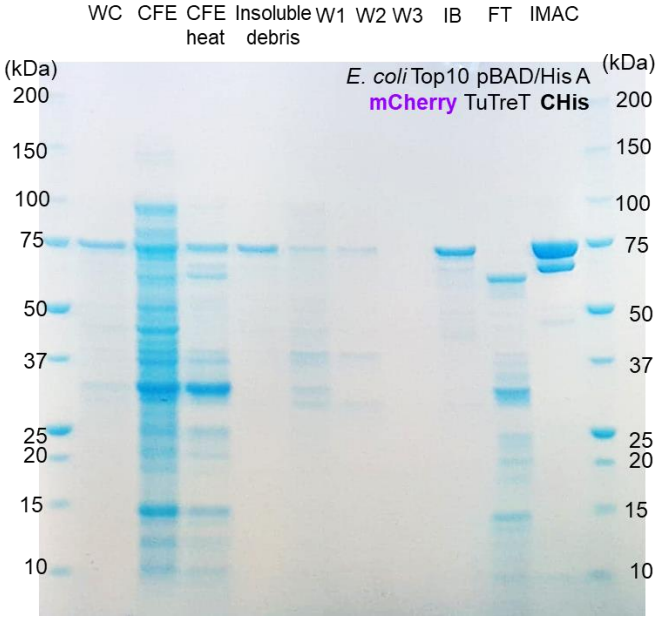


Figure 3: The increase in in protein solubility of mCherry *Tu*TreT, purification of inclusion bodies, and the purification of soluble protein using nickel sepharose were analyzed with SDS-PAGE. Abbreviations: WC: whole cells, CFE: Cell-free extract, CFE heat: after heat-treatment of cell-free extract, Insoluble debris: insoluble pellet after cell lysis, W1: first time washing insoluble debris with 1% DOC detergent, W2: second time washing insoluble debris with 1% DOC detergent, W3: third time washing insoluble debris with 1% DOC detergent, IB: purified inclusion bodies, FT: flow-through of nickel sepharose, IMAC: purified fraction of nickel sepharose.

The protein was purified *via* affinity chromatography without any concomitant precipitation occurring during purification, concentration, or repeated freezing and thawing steps at protein concentrations of up to 15 mg mL^{-1} demonstrating increased solubility and stability of the fusion construct (Fig. 4).

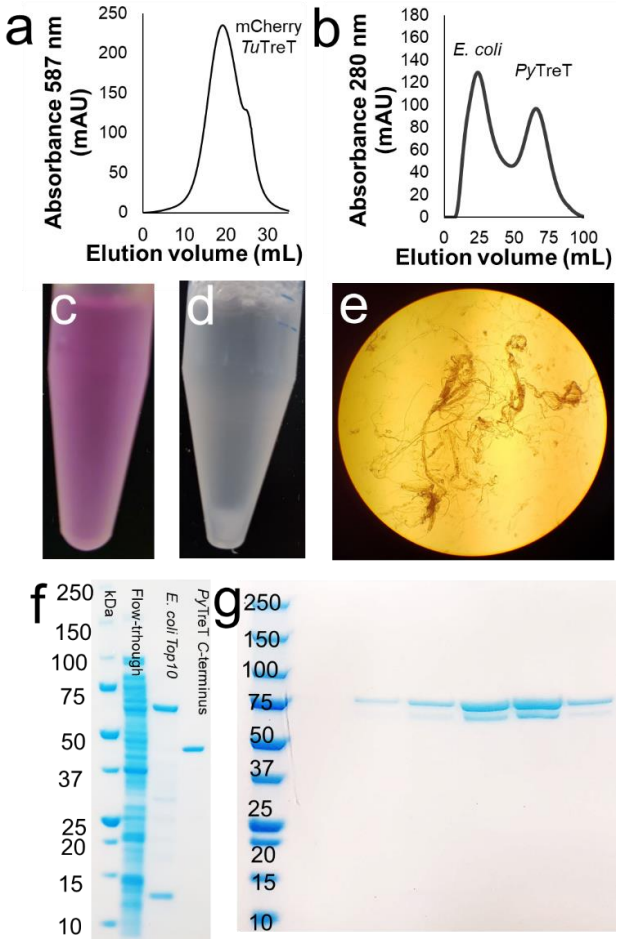


Figure 4: Nickel affinity chromatography purification of mCherry *Tu*TreT at 587 nm (Fig. a) and *Py*TreT at 280 nm (Fig. b). While mCherry *Tu*TreT remains soluble upon purification (c) the C-terminal His-tag *Py*TreT precipitates upon purification (Fig. d), adopting a fibrillar morphology according to optical microscopy (100X magnification) (Fig. e). Both *Py*TreT with a C-terminus Histag (Fig. f) and mCherry *Tu*TreT with C-terminus His-tag (Fig. g) were purified to homogeneity according to SDS-PAGE analysis.

The oligomerization state of mCherry *Tu*TreT was analysed by size exclusion chromatography and the theoretical molecular weight of 74 kDa for the fused protein was in agreement with a monomer of 73 kDa. However, dimerization and oligomerization were also observed for mCherry *Tu*TreT at elevated protein concentrations. Although there is not much known about the increase of protein stability due to the fusion of fluorescent proteins limiting the degree of aggregation in solution, there is another example where protein stability has been increased with fusion to yellow-fluorescent protein (YFP). Here a higher oligomerization state was hypothesized to increase protein stability[20]. In our case, *Tu*TreT is mostly a monomer and only a dimer at high protein concentrations.

Purified mCherry *Tu*TreT was further analysed by SDS-PAGE upon size exclusion chromatography and showed a single fluorescent purple band when the sample was not thermally denatured (Fig. 5). Upon thermal denaturation (SDS-PAGE sample buffer, 100°C), two bands were observed corresponding to residual native enzyme and denatured protein respectively. Variation of the incubation time for thermal denaturation in SDS-sample buffer did not lead to complete thermal unfolding of the protein, but showed increasing hydrolysis of the fusion protein (73 kDa) into its components mCherry (29 kDa) and *Tu*TreT (44 kDa), indicating that the amino acid linker GGSGGGGSGG was hydrolysed. In comparison, *Tu*TreT showed only a single band corresponding to the unfolded protein. The fusion protein therefore showed increased stability against denaturating agents, like sodium dodedecylsulfate (SDS), suggesting an increased protein stability for mCherry *Tu*TreT. Indeed, the purified soluble mCherry *Tu*TreT proved to be stable in 2% SDS in Tris buffer (50 mM, pH 8.0) when the absorption spectra were measured spectrophotometrically showing no protein denaturation.

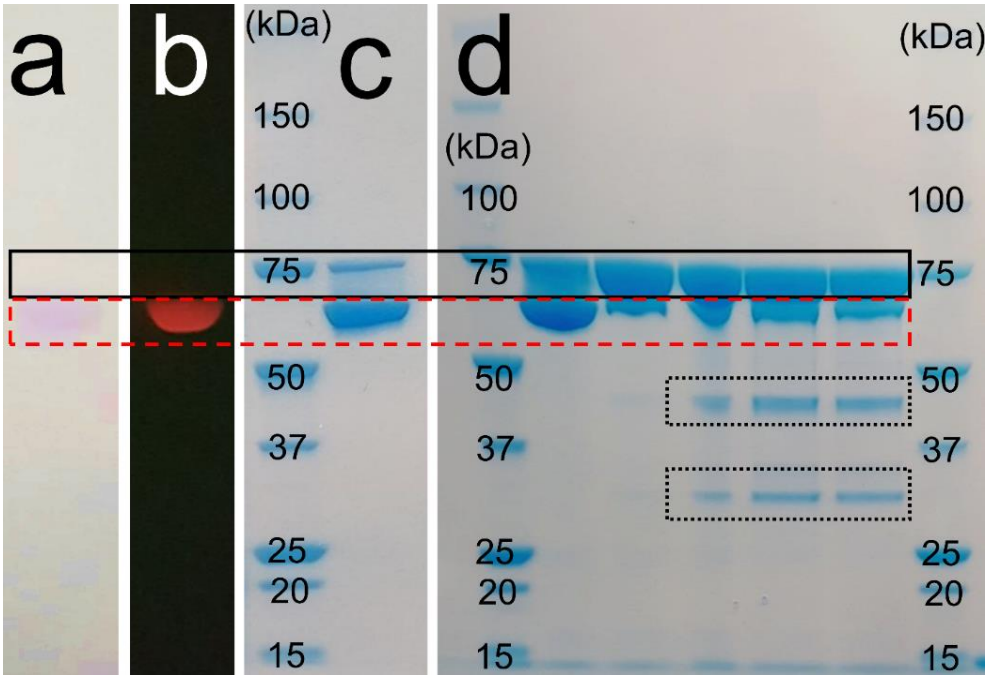


Figure 5. SDS-PAGE of SEC-purified mCherry *Tu*TreT without heat treatment showing native fusion protein in a) and fluorescence under illumination in b) (red dashed box). Subsequent staining demonstrates the appearance of the denatured protein at the expected molecular weight of mCherry *Tu*TreT in c) (black box). The native mCherry *Tu*TreT (red striped box) migrated further than the denatured protein (black box). Degradation of denatured mCherry *Tu*TreT over time at 90 °C resulted in protein sizes of 29 kDa and 44 kDa (dotted black box), indicating hydrolysis of the amino acid linker GGSGGGGSGG, as shown in d).

3.2.3 Direct spectrophotometric protein quantification of mCherry-*Tu*TreT fusion protein

The fusion of the thermostable protein mCherry to *Tu*TreT not only improved protein stability, but also provided a rapid spectroscopic method for protein quantitation. The molar extinction coefficient of the fusion protein mCherry *Tu*TreT was calculated from the protein concentration as determined by the BCA assay and the UV/Vis spectrum of the native protein, which showed an absorbance

maximum at 578 nm. Using the alkali denaturation method, the mCherry chromophore could be converted into the well-studied green fluorescent protein chromophore with a known molar extinction coefficient (ε) of 44000 M^{-1} cm^{-1} and a corresponding shift in the absorbance maximum from 587 nm to 455 nm. Based on these two values, a molar extinction coefficient of $\varepsilon_{mcherry\text{-}TuTreT}$ = 73476 M^{-1} cm^{-1} could be derived for the mCherry fusion protein, which matched the reported literature value of 72000 M^{-1} cm^{-1} for mCherry[21]. Using the calculated molar extinction coefficient of 73476 M^{-1} cm^{-1}, a protein concentration of 2.34 µM was determined spectrophotometrically in good agreement with a protein concentration of 2.50 µM in the BCA assay for mCherry *Tu*TreT.

3.2.4 Non-classical IBs of TreT show glycosyltransferase activity

The formation of insoluble aggregates during recombinant expression is driven by the association of correctly-, partially-, and misfolded proteins[22]. Classical inclusion bodies are described as aggregates of misfolded proteins with complete loss of function[23]. However, 'non-classical' IBs are described as aggregates that contain fully- or partially functional proteins which can be purified by the removal of contaminating membrane-bound proteins by mild solubilization agents like deoxycholic acid (DOC)[24].The removal of other proteins with 1% (v/v) DOC resulted in excellent purity of TreT in all cases (Fig. 6). The TreT content within inclusion bodies proved between 1 to 5% protein in wet inclusion bodies. In the case of the fluorescent mCherry *Tu*TreT, spectrophotometric analysis showed that 68% wt mCherry *Tu*TreT was correctly folded.

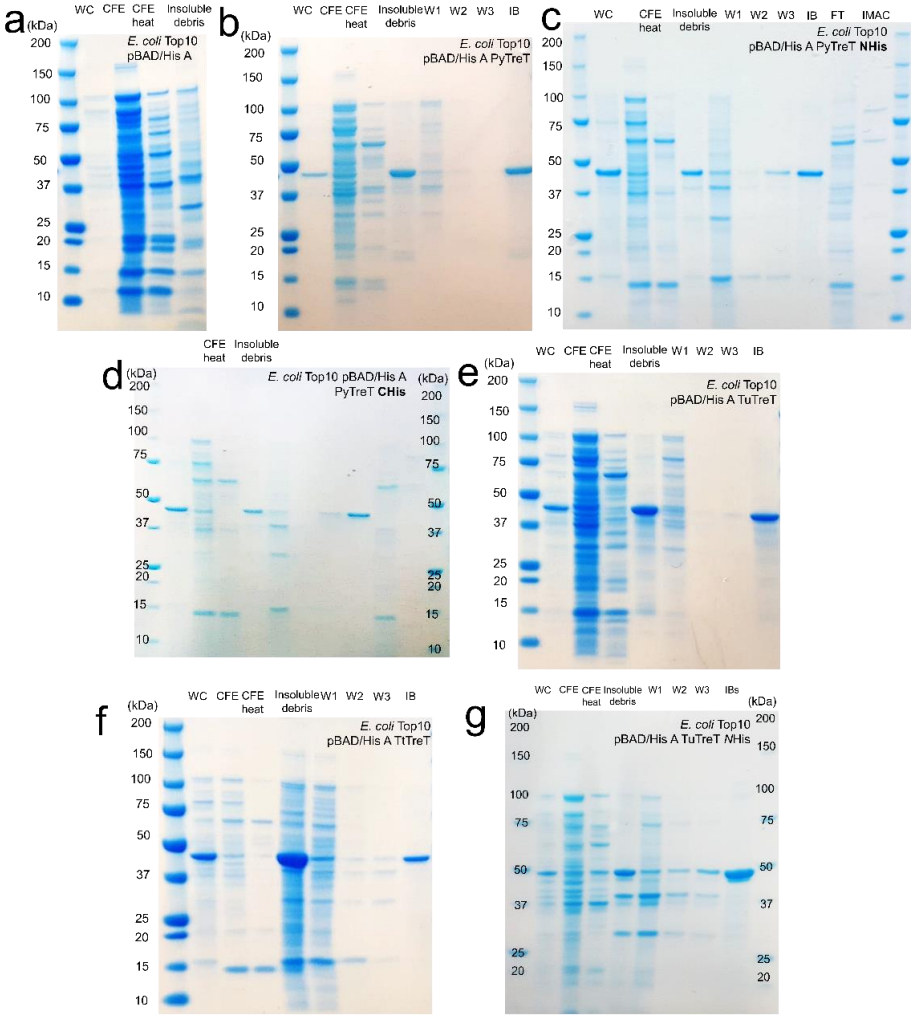


Figure 6: Protein purity of the expression, lysis, and purification of inclusion bodies and soluble (a) *Py*TreT, (b) empty pBAD/His A, (c) *Py*TreT with N-terminus, (d) *Py*TreT with *C*-terminus, (e) *Tu*TreT, (f) *Tt*TreT, and (g) *Tu*TreT *N*His analysed with SDS-PAGE. Abbreviations: WC: whole cells, CFE: Cell-free extract, CFE heat: after heat-treatment of cell-free extract, Insoluble debris: insoluble pellet after cell lysis, W1: first time washing insoluble debris with 1% DOC detergent, W2: second time washing insoluble debris with 1% DOC detergent, W3: third time washing insoluble debris with 1% DOC detergent, IB: purified inclusion bodies, FT: flow-through of nickel sepharose, IMAC: purified fraction of nickel sepharose.

IBs from all TreT variants showed high glycosyltransferase activity between 0.02 – 0.18 U mg^{-1}, as is shown in figure 7. Diffusion limitations within inclusion bodies could lower the observed reaction rate in comparison to the soluble protein, which could be optimized via the increase of temperature or formulation of inclusion bodies. Comparable amounts of IBs were utilized for the conversion of benzaldehyde to (*R*)-mandelonitrile with catalytically active inclusion bodies of hydroxynitrile lyases[25]. The feasibility of IBs as an immobilized biocatalyst has increased the interest for synthetic purposes[26]. However, this study aimed at the biochemical characterization of soluble TreT and the application, further optimization and formulation of TreT IBs was therefore not further pursued beyond the proof of concept.

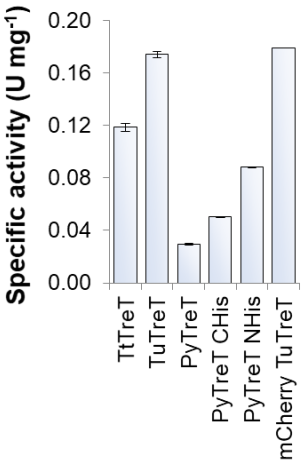


Figure 7. Specific activity of purified TreT IBs. **Reaction conditions**: HEPES (50 mM, pH 7.0), $MgCl_2$ (20 mM), UDP-glucose (40 mM), glucose (10 mM), 3.0 – 5.0 mg of TreT IB, 60°C.

3.2.5 Kinetic characterization and thermal stability of mCherry-*Tu*TreT

The glycosyltransferase activity of mCherry-*Tu*TreT was determined by HPLC analysis. A temperature optimum for the protein stability was determined to be 60 °C, and a V_{max} of 11.39 ± 0.29 U mg^{-1} and K_M of 0.61 ± 0.11 mM were obtained for UDP-glucose. The kinetic constants temperatures above 60 °C were

not investigated, due to the degradation of UDP-glucose under these conditions. In comparison, others have shown a V_{max} of 184 U mg^{-1} and K_M (UDP-glucose) of 0.23 mM for *Tt*TreT using a trienzymatic coupled assay at 80 °C[9b]. For D-glucose, non-competitive substrate inhibition was observed with a K_M (glucose) of 2.30 ± 0.58 mM, a K_i (glucose) of 10.63 ± 2.21 mM, and a V_{max} of 17.06 ± 2.22 U mg^{-1}. Substrate inhibition has not been reported for other TreTs[9b, 9f].

mCherry-*Tu*TreT was incubated for two hours at 60, 70, 80, and 90°C and residual enzyme activities are shown in Figure 8a. Residual protein activities were found to correlate with the residual absorbance from mCherry. The initial rate of the enzyme increased exponentially with temperature according to the Arrhenius equation, showing the highest activity at 80 °C (Fig. 8b) and a $\Delta G^{\#}$ = 92.7 kJ mol^{-1}, was determined from the Arrhenius plot (Fig. 8c).

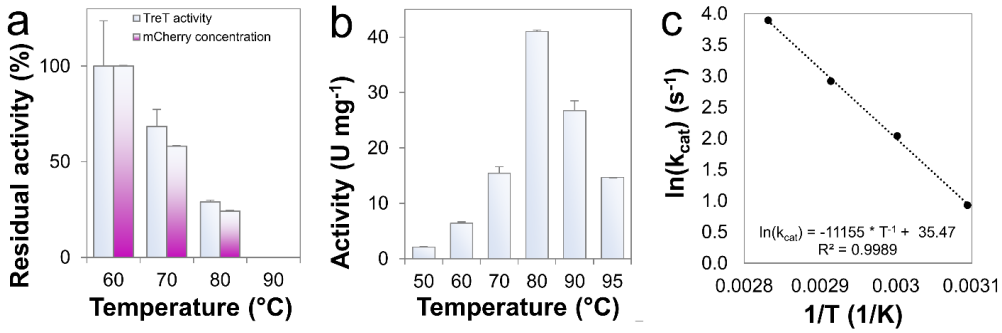


Figure 8. In a) the thermostability of mCherry *Tu*TreT after 2h incubation at 60, 70, 80, and 90 °C is shown in a) (1.0 mg mL^{-1} mCherry *Tu*TreT, 50 mM HEPES, pH = 7.0). In b) the initial activity of mCherry *Tu*TreT depending on temperature, showing deactivation after 80°C, where c) shows the Arrhenius plot of the initial enzyme activity from 50 to 80 °C. **Reaction condition**: HEPES (50 mM, pH 7.0), $MgCl_2$ (20 mM), UDP-glucose (40 mM), glucose (10 mM), and temperatures between 50 °C to 95 °C.

3.2.6 The effect of pH, cations, and anions on the activity of mCherry-*Tu*TreT

The fusion protein mCherry-*Tu*TreT showed a broad pH stability over a pH range from 5 – 9 (Fig. 9a). The optimal pH was found at pH 6.0, which is similar to the pH optimum of a trehalose transferase from *Thermococcus litoralis*[9f]. Divalent

cations could potentially form chelates with the phosphate group of the sugar donor substrate and thereby influence the substrate binding and enzyme activity. A wide range of metal salts were explored to account for potential chaotropic and kosmotropic effects of the counter anions. As can be seen in Figure 9b, the addition of sodium chloride slightly inhibited the enzyme, while sodium sulphate did not. The enzyme activity was found to increase with the addition of metal chloride salts in the following order: Mg^{2+} ~ Ni^{2+} < Co^{2+} < Ca^{2+} < Mn^{2+}, while addition of Zn^{2+} resulted in the complete inhibition of the enzyme. A similar behavior was reported for a trehalose transferase from *Thermococcus literoralis*, where zinc(II) chloride completely inhibited the enzyme[9f]. No difference in enzyme activity was observed when magnesium chloride was substituted with magnesium sulphate. Surprisingly, the strong chaotrope guanidine hydrochloride similarly increased the enzyme activity as magnesium(II).

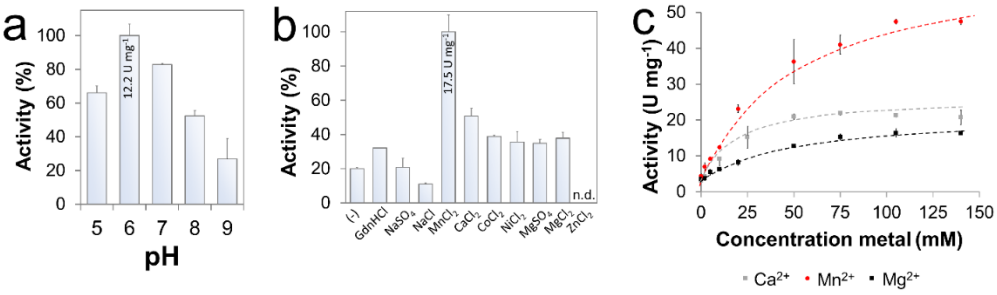


Figure 9. The enzyme activity of mCherry *Tu*TreT using a multi-component buffer shows a broad pH distribution in a). The effect of different cations and anions (20 mM) on enzyme activity is shown in b). Enzyme saturation kinetics of the three best metals were investigated in c). **Reaction conditions**: HEPES (50 mM, pH 7.0) or the multicomponent buffer, $MgCl_2$ (20 mM) for 20 mM of additive, UDP-glucose (40 mM), glucose (10 mM), 60°C.

To delineate the different effects of Mg(II), Ca(II), Mn(II) on the K_d and V_{sat} the protein stability of the mCherry *Tu*TreT metal complex was investigated. The observed protein melting temperatures T_m for Mn(II), Ca(II), and Mg(II) were 75 °C, 82 °C, and 82°C, respectively, indicating that chaotropic divalent cations reduce

the enzyme conformational stability. Moreover, the melting pattern showed faster denaturation of mCherry *Tu*TreT for Ca(II) in comparison with Mg(II) despite a similar T_m. This indicates, that calcium destabilizes the protein to a higher degree.

3.2.7 Reaction equilibrium of trehalose transferase catalyzed reactions is dependent on the nature of the nucleotides and nucleotide carbohydrates

Like any other catalyst, enzymes enhance the rate of reaction towards the thermodynamic equilibrium. However, it has been suggested that *Tt*TreT only catalyses the forward reaction in the synthesis direction when using UDP and UDP-glucose[9b], while the trehalose transferase from *Pyrococcus horikoshii* has been shown to catalyze the reaction reversibly using a wide range of nucleotide diphosphates (Fig. 10)[9e].

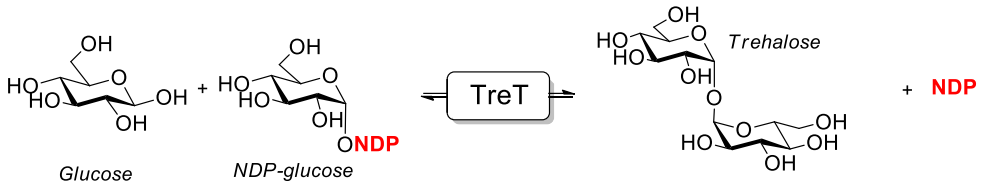


Figure 10. The overall reaction of α,α-(1→1) coupling of α-D-glucose and NDP-D-glucose to synthesise trehalose and nucleotide diphosphate (NDP), which is either adenosine- (ADP) or uridine diphosphate (UDP).

Overall, the thermodynamic equilibrium for the enzyme catalyzed synthesis of trehalose from glucose and a nucleotide sugar lies in favor of the product trehalose for both UDP and ADP as nucleotides, its exact position still depends on the respective nucleotide sugar utilized. The synthesis of trehalose using an excess of UDP- or ADP-glucose and glucose with mCherry *Tu*TreT, leads to quantitative conversion of glucose, while the reverse reaction of trehalose with UDP or ADP leads to the formation of glucose and UDP- or ADP-glucose. According to HPLC analysis, a low specific activity for ADP (98 mU mg^{-1}) *vs.* UDP (290 mU mg^{-1}) was observed and a K_{eq} of 157 for UDP and 30 for ADP was

determined experimentally. Analyzing previously reported data, reactions with TreT from *Pyrococcus horikoshii* were found to establish equilibrium concentrations with a K_{eq} of ~230 for UDP and ~29 for ADP[9e]. Hence, it was shown for different TreT enzymes that they catalyze both the forward and reverse reaction, and that the final conversions depend on the utilized nucleotide sugar, where product formation is more favored when UDP is used over ADP.

Overall, the thermodynamic equilibrium K_{eq} for the synthesis of trehalose and UDP from glucose and UDP-glucose also depends on the pH, metal ion composition, and ionic strength. Although chelation of divalent cations with the phosphate moiety of the nucleotide is not included in the estimation of Gibbs free energies of formation, the equilibrium constants K_{eq} for the formation of trehalose from glucose using uridine or adenosine diphosphate glucose at varying pH values could be estimated with the thermodynamic calculator eQuilibrator 2.0 (supplementary information)[27]. Equilibrium constants of 23 were calculated for ADP and 225 for UDP, which reasonably matched the values that were observed experimentally (K_{eq} = 30 and 157, respectively). Based on this, our calculations predict a shift of the equilibrium constant towards the starting materials under increasingly acidic conditions due to the protonation of uridine diphosphate (pKa 5.5 – 6.5 [28]) (Fig. 11). A similar pH dependence has been reported for UDP-dependent sucrose synthase[28b]. The application of eQuilibrator 2.0 to determine thermodynamic equilibria has successfully been implemented with sucrose synthase by others[28b].

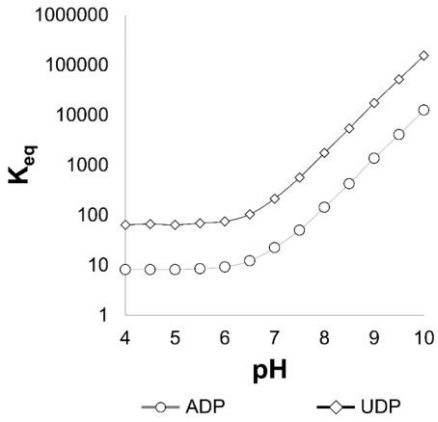


Figure 11. The influence of pH on the thermodynamic equilibrium of the coupling of glucose and UDP- or ADP-glucose to trehalose and UDP or ADP was calculated by eQuilibrator 2.0 (http://equilibrator.weizmann.ac.il/).

All the trehalose transferases investigated here, each having a different isoelectric point, did not show an increased solubility, resulting in the formation of inclusion bodies *in vivo* and aggregation of soluble protein *in vitro*. While our results clearly showed an increase in protein solubility and stability through the fusion of *Tu*TreT with mCherry, the exact mechanism behind these observations remains elusive. mCherry might function as a molecular chaperone by stabilizing aggregation-prone folding intermediates, which previously has also been suggested for the maltose binding protein[29]. Furthermore, self-oligomerization of correctly folded TreT might be reduced due to the increased size of the fusion protein. While other solubility tags potentially could have achieved similar results[30], the use of mCherry as a thermostable, fluorescent protein tag allowed rapid spectrophotometric protein quantification and exhibited excellent reaction compatibility with a thermostable enzyme.

The biochemical characterization of mCherry *Tu*TreT showed non-competitive substrate inhibition for glucose, which was not reported earlier for other trehalose transferases. A structural explanation for non-competitive inhibition would be the binding of sugar acceptor in the binding site for the nucleotide sugar donor. A comparison with untagged *Tu*TreT was not possible due to its aggregational

behavior and therefore we cannot exclude an effect of the mCherry fusion on the kinetic properties of TreT. Due to the length of the linker and the distance of mCherry to the active site of *Tu*TreT it is not likely that the kinetic parameters would be drastically different. The observed rate enhancement in the presence of chloride salts of guanidine, Mg(II), Ca(II), and Mn(II) could be explained by either (i) a decrease of enzyme rigidity by chaotropic agents; (ii) complexation of the metal with the diphosphate-group of UDP and/or UDP-glucose; or a combination of both. Considering the observation, that guanidinium hydrochloride similarly increased the enzyme activity, this study suggests that enzyme rigidity is a controlling factor in mCherry *Tu*TreT and that chaotropic agents can increase the enzyme flexibility and thereby the activity. This is not surprising, since meso- and thermophilic proteins have been shown earlier to demonstrate a catalytic enhancement by the addition of chaotropic agents leading to a decrease to the intrinsic high conformational rigidities of thermostable proteins[31]. Indeed, a conformational change of *ca* 4Å has been observed between the sugar donor and acceptor binding domains, for the protein crystal structure of TreT from *Pyrococcus horikoshii* (PDB: 2X6Q), which closes upon substrate binding (PDB: 2XMP) highlighting the importance of conformational flexibility[9d]. Moreover, conformational flexibility upon substrate binding has also been observed for trehalose phosphate synthase (OtsA) from *E. coli* K12[32], glycogen synthase[33] and α-fucosyltransferase V[34]. Assuming that enzyme mobility is rate-limiting, chaotropic reagents plausibly explain increased TreT activity. On the other side, Zn(II) completely inhibited the activity of mCherry *Tu*TreT and TreT from *Thermococcocus litoralis*[9f] emphasizing the effect that complexation of metals with the diphosphate moiety of the nucleotide sugar donor can have.

3.3 Conclusion

To conclude, the fusion of mCherry to *Tu*TreT showed that an intrinsically aggregation-prone protein could be stabilized in solution and offered a tool to monitor protein solubility by UV-VIS spectroscopy. This allowed for the biochemical

characterization of mCherry *Tu*TreT at elevated temperatures, which showed increased activity using manganese(II) and the chaotropic reagent guanidine hydrochloride. Furthermore, our results highlight that the equilibrium constant for the synthesis or hydrolysis of trehalose is determined by the composition of the reaction mixture, where the utilization of different nucleotides and pH values can substantially shift the equilibrium. Trehalose transferases are therefore not unidirectional, and the use of a specific nucleotide or nucleotide sugar determines the overall conversion. The production and biochemical characterization of the stable mCherry *Tu*TreT fusion protein addressed one of the major problems of archaeal glycosyltransferases and is therefore of particular relevance for the industrial production of novel disaccharides and nucleotide carbohydrates, gaining insight in the optimal process conditions and thermodynamic limitations when using trehalose transferases.

3.4 Materials and methods

3.4.1 Chemicals

Uridine 5'-diphosphate disodium salt (Carbosynth, 98%), D-glucose (Sigma-Aldrich, 99.5%), HEPES (Sigma-Aldrich, >99.5%,), $MgCl_2$ hexahydrate (VWR, >99.5%,), $CaCl_2$ dihydrate (Sigma-Aldrich, >99.0%), $NiCl_2$ hexahydrate (99.9%), guanidine hydrochloride (>99.5%, Sigma-Aldrich), sodium deoxycholate (>98%, Sigma-Aldrich), sodium sulphate (Sigma-Aldrich, >99.0%), zinc chloride (Sigma-Aldrich, >98%), magnesium sulphate heptahydrate (Sigma-Aldrich, 99%)), cobalt chloride hexahydrate (Sigma-Aldrich, >98%), glycerol (99.5%, Sigma-Aldrich), rubidium chloride (Sigma-Aldrich, > 99%), potassium acetate (Acros, >99%), sulphuric acid (Acros, 98%), agarose (Sigma-Aldrich, >99%), ampicillin (Sigma-Aldrich), Tris(hydroxymethyl)aminomethane (Tris, 99%, Sigma-Aldrich), glycine (Sigma-Aldrich, >99%), pyridine (Sigma-Aldrich, >99%), sodium chloride (J.T. Baker, 99.5%), Bis-Tris (Sigma-Aldrich, >99%), acetonitrile (Sigma-Aldrich, >99.5%), uridine 5'-diphosphate disodium salt (Sigma-Aldrich, >96%), adenosine 5'-

diphosphate disodium salt (Sigma-Aldrich, bacterial, >95%), adenosine 5'-diphosphoglucose disodium salt (Sigma-Aldrich, >93%).

3.4.2 Materials

The QIAprep Miniprep kit was purchased from Qiagen. The high-fidelity restriction endonucleases *KpnI* HF, SacI HF, *BamHI* HF, *NcoI* HF were used with a tandard protocol using 10x CutSmart buffer (New England Biolabs). Purification of plasmids from agarose gel was performed with the Monarch DNA gel extraction kit (New England Biolabs) using the standard protocol. Ligation was performed using standard protocol with T7 DNA ligase and T7 DNA ligase reaction buffer (New England Biolabs).

3.4.3 Analytical instruments

Chromatographic analysis of reactions was performed using a Shimadzu high-performance liquid chromatography (HPLC) system equipped with a Imtakt Unison-UK amino column (0.4 x 25 cm, 60 °C), an ELSD detector (Schimadzu ELSD-LTII), UV-detector (SPD-20A), and acetonitrile:water:formic acid 80:20:0.1 as mobile phase (1 mL min^{-1}). The samples were calibrated using external calibration curve.

3.4.4 BCA assay

Protein content was determined with the BCA protein quantitation kit (Thermo Scientific, Carlsbad, USA). Standard curves were prepared with BSA in the range of 0.01 – 2 mg mL^{-1} in (poly)styrene 96 wells plate. Samples were measured in triplicates and monitored at 562 nm utilizing a microtiter plate spectrophotometer (Synergy 2, BioTek). A protocol for the solubilisation of inclusion bodies in 2% SDS in Tris HCl buffer (50 mM, pH 8.0) was adopted from literature (54). The negative control containing 2% SDS in Tris HCl buffer (50 mM, pH 8.0) does not show a background absorption with the BCA reagent buffer.

3.4.5 SDS-PAGE

Protein samples were denatured using XT sample buffer (BioRad) supplied with XT reducing agent (BioRad) at 95 °C for 15 minutes. Gel electrophoresis was performed with Criterion XT 4-12% Bis-Tris precast gels (Bio-Rad) using MOPS buffer (Bio-Rad). The gels were run at 150 V for 40 to 60 minutes and stained with SimplyBlue SafeStain (Novex). The Precision Plus Protein Unstained Standard (Bio-Rad) was used to determine the relative molecular mass of the protein.

3.4.6 Spectrophotometric measurements

The absorbance of mCherry-TuTreT at $\lambda_{190nm-800nm}$ was measured utilizing a 1 cm quartz cuvette. All measurements for the determination of the molar extinction coefficient were performed in triplicates.

3.4.7 Growth Media

Terrific broth medium consists of 1.20% (w/w) tryptone, 2.40% (w/w) yeast extract, 53 mM K_2HPO_4, 16 mM KH_2PO_4, 4% (w/w) glycerol, and was autoclaved at 121°C for 20 minutes. Auto-induction medium ZYM-5052 was prepared according to literature protocols (67). LB-medium consists out of 1.00% (w/w) tryptone, 0.5% (w/w) yeast extract, 1% NaCl, and was autoclaved at 121°C for 20 minutes. All media was supplemented with 100 µg mL^{-1} ampicillin.

3.4.8 Bacterial plasmids and strains

The plasmid pBAD/His A (Invitrogen) was provided by their commercial suppliers. The strains *E. coli* DH5α with the genotype F - ϕ80*lac*Z ∆M15∆ (lacZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(r_k^- , m_k^+) *gal*$^-$ *pho*A *sup*E44 λ$^-$ thi$^-$1 *gyr*A96 *rel*A1, and *E. coli* Top10 with the genotype F^- *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara leu*) 7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G were ordered from New England Biolabs and Invitrogen respectively.

3.4.9 Preparation of competent cells with rubidium chloride

Competent cells of *E. coli* DH5α and *E. coli* Top10 were prepared with rubidium chloride. Cells from an overnight culture were grown to OD600 of 0.5 in LB medium and centrifuged (425 rcf, 15 minutes, 4°C). The LB-medium was decanted and the cells were washed in 30 mL of freshly prepared ice-cold solution rubidium chloride (100 mM), manganese(II) chloride (10 mM), potassium acetate (3 mM), calcium chloride (1 mM), glycerol (165 mM), followed by centrifugation at 2000 rpm (5 minutes, 4°C). The cells were resuspended in 4 mL MOPS buffer (100 mM, pH 7.0) containing RbCl (10 mM), $CaCl_2$ (5 mM), and glycerol (165 mM) and 0.1 mL was aliquoted in ice-cold polypropylene Eppendorf tubes. The competent cells were stored at -80°C.

3.4.10 Transformation

The synthesized, lyophilized DNA (Baseclear, Leiden) was briefly centrifuged (425 rcf, 30 seconds), and resuspended in 40 µL Tris buffer (10 mM, pH 8.5) and diluted 1:10. The measured DNA concentrations measured via the ratio of λ = 280/260 nm showed a similar DNA concentration of 200 ng μL^{-1} and 20 ng μL^{-1} respectively. The competent cells were thawed and the pUC-SP plasmids containing the synthesized gene was added to reach a final concentration of ~8 ng μL^{-1} and ~1.6 ng μL^{-1} respectively. After 30 minutes of incubation on ice, the competent cells were heat-treated at 42 °C for 30 seconds. To the solution 500 µL of sterile LB-medium was added and incubated for 1h at 37°C, followed by plating on agar plates containing the ampicillin (100 µg mL^{-1}).

3.4.11 Cloning, and expression of TreT from in *E. coli* Top10

The pUC-SP vector containing the codon-optimized and synthetic TreT genes were transformed in chemically competent *E. coli* DH5α strains and stored as glycerol stocks at -80 °C. The *treT* gene *E. coli* DH5α was grown in LB-medium containing 100 µg mL^{-1} ampicillin overnight at 37 °C and the plasmid was isolated (QIAprep

Miniprep, Qiagen). The *treT* containing pUC-SP and pBAD/His A plasmids were digested with *Kpn*I and *Nco*I. The digested fragments were purified on a 1% agarose gel after gel electrophoresis (BioRad, 120V) using the standard protocol of the Monarch DNA gel extraction kit (New England BioLabs). After ligation with T7 ligase at 16 °C overnight using the provided protocol (New England BioLabs), the plasmid was transformed into competent *E. coli* Top10 cells and sequenced (BaseClear, Leiden).

3.4.12 Production and purification of recombinant TreT from in *E. coli* Top10 pBAD/His A

(i) Preparation of cell-free extract The 5 mL inocula of *E. coli* Top10 pBAD/His A containing *Tt*TreT, *Tu*TreT, *Py*TreT, *N*- or C-terminus His-tagged *Py*TreT, *N*-terminus His-tagged *Tu*TreT, and mCherry *Tu*TreT genes were grown in LB-medium containing 100 µg mL^{-1} ampicillin at 37 °C overnight. To seven 2 L baffled Erlenmeyer flask containing 400 mL TB-medium was 5 mL inoculum added (1.3% v/v) and induced with L-arabinose to a final concentration of 0.02% (w/w) after reaching an OD600 of 0.6 – 0.8. The cells were harvested at an OD600 after 14 hours by centrifugation (17 000 g, 15 min, 4°C) followed by resuspension of wet cell pellet in 25 mL lysis buffer containing Tris HCl buffer (50 mM, pH 7.4), imidazole (20 mM), lysozyme (0.5 mg mL^{-1}), DNaseI (0.1 mg mL^{-1}) per gram of wet cells. After 30 minutes of incubation on ice the cells were passed through the cell disruptor (1.35 kbar) for three consecutive rounds. The cell debris was collected via centrifugation 12 000 rpm (Sorvall, Fiberlite F12-6x500 LEX, 10 min, 20 °C), and the CFE was obtained via decantation.

(ii) Immobilized nickel affinity chromatography The CFE was heat-treated at 60 °C for 20 minutes in a water bath. The precipitates in the CFE were removed via centrifugation 12 000 rpm (Sorvall, Fiberlite F12-6x500 LEX, 10 min, 20 °C), and the heat-treated CFE was obtained via decanting. The heat-treated CFE was purified using affinity chromatography on a 1 mL Nickel Sepharose column by charging CFE on the column for at least three consecutive rounds using a

peristaltic pump (Bio-Rad). The column was washed with binding buffer (20 mM Tris HCl, 500 mM NaCl, 20 mM imidazole, pH 7.4) until no protein eluted any longer. The bound enzyme was eluted using elution buffer (20 mM Tris HCl, 500 mM NaCl, 500 mM imidazole, pH 7.4). Protein samples were concentrated in a 12 mL Amicon Ultra Centrifugal filter (Merck, 30 kDa). Elution buffer was exchanged for HEPES (50 mM, pH 7.0) containing $MgCl_2$ (20 mM) by washing three consecutive rounds with 12 mL Amicon Ultra Centrifugal filters (Merck, 30 kDa) analysed with SDS-PAGE and HPLC.

(iii) Purification of inclusion bodies A literature protocol was adopted (68). The insoluble debris was homogenized in 20 mL Tris HCl buffer (50 mM, pH 8.5) containing 1% (w/w) deoxycholic acid. The solubilized trehalose transferases were separated from the inclusion bodies via centrifugation (20 000 g, 15 min, 20 °C) and is referred to as a washing solution. The solubilisation and centrifugation were repeated two times, resulting in solubilized washing solution 1, 2 and 3. Next, Tris HCl buffer (50 mM, pH 8.5) was utilized to remove remaining DOC. The inclusion bodies were harvested via centrifugation (20 000 g, 15 min, 20 °C) and the solutions were analysed with SDS-PAGE for protein purity.

3.4.13 Production and purification of soluble recombinant *C*-His-tagged TreT from *Pyrobaculum yellowstonensis WP30* in *E. coli* Top10 pBAD/His A The protocol described above was repeated six-fold for 5 liter Erlenmeyer flasks containing 1 L of TB media. The isolated cell-free extract was purified using a prepacked 5 mL HisTrap FF column (GE Healthcare) and analysed with SDS-PAGE and HPLC. Reaction conditions: D-glucose (10 mM), UDP-glucose (0 – 50 mM), HEPES (50 mM, pH 7.0), *Py*TreT (0.02 mg mL^{-1}), and 20 mM magnesium(II) chloride, 80 °C.

3.4.14 Production and purification of soluble recombinant mCherry-*Tu*TreT in *E. coli* Top10 pBAD/His A The cell-free extract was prepared as was described above, containing 1 L of TB media in an Erlenmeyer flask. All of the other steps

were sized accordingly, using a 12 mL column prepacked HisTap FF Column for purification. As an additional purification, the protein was purified using a Superdex 200 Increase 10/300 GL column with HEPES (50 mM, pH 7.4) buffer containing 300 mM NaCl as a mobile phase. The column was calibrated with a Gel Filtration Standard (Bio-Rad, cat. nr. 151-1901) containing a lyophilized mix of thyroglobulin, bovine γ-globulin, chicken ovalbumin, equine myoglobin, and vitamin B12 (M_w 670 – 1.35 kDa) before use. The eluate was concentrated using 12 mL Amicon Ultra Centrifugal filter (Merck, 30 kDa) yielding 1 mg per liter of LB media. The sample was analysed with SDS-PAGE and HPLC.

3.4.15 Quantification of D-glucose and D-trehalose with HPLC Samples during activity assays were quenched by the addition of 50 µL of reaction solution to an equal volume of ice-cold HPLC-grade acetonitrile and incubated at -80 °C for one hour. The samples were centrifuged at 14000 rpm for 10 min at 4 °C. The supernatant was collected and analysed by HPLC (column: Imtakt UK-Amino 250 x 4.6 mm, 50 °C, ELSD, 80:20 ACN:H_2O, 1.0 mL min^{-1}). Enzyme activity was calculated with external standards for trehalose using the slope of at least three different substrate concentrations. The enzyme activity was determined in duplicates.

For D-glucose, the reaction was varied for D-glucose (0 – 35 mM), UDP-glucose (40 mM), HEPES (50 mM, pH 7.0), mCherry *Tu*TreT (0.02 mg mL^{-1}), and $MgCl_2$ (20 mM). Reaction was incubated at 60 °C with gentle shaking. The data was fitted (Gnuplot 5.2).

For the kinetic analysis of UDP-glucose, the reaction conditions were varied with D-glucose (10 mM), UDP-glucose (0 – 50 mM), HEPES (50 mM, pH 7.0), mCherry *Tu*TreT (0.02 mg mL^{-1}), and 20 mM magnesium(II) chloride. Reaction was incubated at 60 °C with gentle shaking. The data was fitted (Gnuplot 5.2).

For the evaluation of enzyme activity and kinetic analysis different cations or anions, the reaction conditions were varied for D-glucose (10 mM), UDP-glucose (40 mM), HEPES (50 mM, pH 7.0), mCherry *Tu*TreT (0.02 mg mL^{-1}), and either

guanidine hydrochloride (20 mM), sodium sulphate (20 mM), sodium chloride (20 mM), manganese(II) chloride (0 – 140 mM), calcium(II) chloride (0 – 140 mM), cobalt(II) chloride (20 mM), nickel(II) chloride (20 mM), magnesium(II) sulphate (20 mM), magnesium(II) chloride (0 – 140 mM), and zinc(II) chloride (20 mM). Reaction was incubated at 60 °C with gentle shaking. The data was fitted (Gnuplot 5.2). The effect of pH was evaluated using a multicomponent buffer containing D-glucose (10 mM), UDP-glucose (40 mM), mCherry *Tu*TreT (0.02 mg mL^{-1}), pyridine (15 mM), Bis-Tris (15 mM), HEPES (15 mM), glycine (15 mM), $MgCl_2$ (20 mM), NaCl (150 mM). Reaction was incubated at 60 °C with gentle shaking. The enzyme activity was determined in duplicates.

The reaction temperature was varied between 50 - 95 °C with gentle shaking for D-glucose (10 mM), UDP-glucose (40 mM), HEPES (50 mM, pH 7.0), mCherry *Tu*TreT (0.02 mg mL^{-1}), and $MgCl_2$ (20 mM). The enzyme activity was determined in duplicates.

To assess the stability of mCherry *Tu*TreT a 1 mL stock solution containing 1.0 mg mL^{-1} was incubated between 50 – 80 °C and the absorbance (587 nm) was measured in a 1 cm polyacrylate cuvette. The enzyme activity was measured with D-glucose (10 mM), UDP-glucose (40 mM), HEPES (50 mM, pH 7.0), mCherry *Tu*TreT (0.02 mg mL^{-1}), and $MgCl_2$ (20 mM), 60 °C. The enzyme activity was determined in duplicates using HPLC analysis.

3.4.16 Thermal shift assays The melting temperature in different solution conditions containing Mn(II), Ca(II), or Mg(II) were determined by using the thermal shift assay (or differential scanning fluorimetry (DSF)). Briefly, the mCherry *Tu*TreT was diluted in HEPES buffer (50 mM, pH 7.0) containing 20 mM of divalent cation, 300 mM of NaCl, and SYPRO Orange solution (Thermofischer Scientific, S-6651). The microplate was sealed with an adhesive optical clear seal (MicroAmp Optical Adhesive Film), centrifuged at 4 °C for 30 s, and heated from 5 to 95°C, with increments of 1 °C/min, using an RT-PCR instrument (StepOnePlus Real-Time PCR System, Applied Biosystems). Fluorescence in each well was followed by

applying excitation and emission wavelengths of 485 nm and 530 nm, respectively. The melting temperature (T_m) corresponds to the temperature where the protein is 50% unfolded.

3.4.17 Determination of thermodynamic equilibrium of the trehalose transferase reaction using mCherry *Tu*TreT Reaction equilibria was determined via the addition of enzyme. For the forward reaction D-glucose (10 mM), UDP- or ADP-glucose (40 mM), HEPES (50 mM, pH 7.0), mCherry-*Tu*TreT (1.0 mg mL^{-1}), and $MgCl_2$ (20 mM) were evaluated monitoring the production of trehalose. For the reverse reaction, D-trehalose (10 mM), UDP or ADP (40 mM), HEPES (50 mM, pH 7.0), mCherry-*Tu*TreT (1.0 mg mL^{-1}), and $MgCl_2$ (20 mM) were utilised following the production of D-glucose. The enzyme activity was determined in duplicates using HPLC analysis. The thermodynamic equilibrium was determined from the last three datapoints.

3.5 Sequences

Trehalose transferase from *Pyrobaculum yellowstonesis* WP30

Non-codon-optimized gene (GenBank: CP012158.1)

```
ATGATCGAGCGGTATATTCAGTTCGTGGGGGAGGACGAGGTAGACGCCATCGTCAAGTTGGCGGAGCGGCTACAAGACCTCTCCATA
CTTCACGTCAACTCCACCGCGGCGGGGGGCGGCGTGGCTGAGATCCTCAACCGCCTAGTTCCGCTCATGCGGGAGCTGGGCCTCAGG
GTGGACTGGAGGGTGATCAGAGGCGACCAGGAGTTCTTCACAGTGACGAAGACCTTCCACAACGCCCTGCAGGTCGGCGCCGTGGAG
GTGCCTAGACGCTTCTACGAAATATACGAGAGGTGGCAGGAGATAAACGCCAACGAGCTAGACCTCGACTACGACGTGGTCTTCATC
CACGACCCCCAGCCTGCCGGCCTTGTGAAGTACAGAAAAAGAGGGCTTTGGATCTGGCGTTGCCACATCGACGTCTCCACGCCGAAT
CCAGAGGTGTGGGGCTTCCTCCGGAGGTACGTGTCGCAGTACGACGCCGCCATCTTCCACATACCGGAGTTCGCCCGAGACGACCTC
GAAATACCGCAGTTGCTCATCCCGCCGTCTATAGACCCCCTGAGCCCTAAAAACGTGGAGCTACCCCCCACGGCGGTGGAGCGGGTG
GTGCAGAAATTCGACGTAGACCCAGAGCGGCCCATTCTTCTGCAAGTGTCGCGGTTCGACAGAGCTAAGGACCCCCTGGGCGTCGTG
GAGGCGTACCGACTCGCCAGACGACGCGTGCCTGGCCTCCAGCTGGTGTACCTCGGTAGCCCCGCCCACGATGACCCGGAGGGCGAG
GCCGTGTACAGAGAGACCGTAGAGGCGGCGGGCGGCGACCCCGACATACACCTGCTCATGCTCCCGCCCGACAGCCACTACGAGGTA
AACGCCTTCCAACGCGCCGCCACAGTCGTCATGCAGAAGTCCATAAGAGAAGGCTTCGGCCTCACCGTCAGCGAAGCCCTTTGGAAG
AAGAGGCCAGTGATAGGCGGCAAGACGGGCGGCATAAAGATCCAGGTGATACACGGCGTCACTGGCTTCCTCGCCACCTCGCCGAGA
ACCGCGGCCCACTACGCCGTATATCTGCTCAGAGAGAAAAGATTAAGGGAAGAAATGGGCGCCGCAGGCAGAGAACACGTCAGACGC
AACTTCCTCATCACACATCAACTACGCAGATACTTAATGGCCATAGCGTACGTGGCTAAGAGAGCGATGTGGAGTGACTGA
```

Codon-optimized synthetic DNA sequence

```
ccatggGCATGATTGAACGTTACATCCAGTTCGTTGGTGAAGATGAAGTTGATGCTATTGTGAAACTGGCGGAACGTCTGCAGGATC
TGTCTATCCTGCACGTTAACAGCACCGCTGCGGGCGGTGGTGTTGCTGAAATTCTGAACCGCCTGGTTCCGCTGATGCGTGAACTGG
GTCTGCGTGTTGATTGGCGTGTTATTCGCGGTGATCAGGAATTCTTTACCGTGACCAAAACCTTTCATAATGCCCTGCAAGTGGGCG
CCGTGGAAGTTCCGCGCCGCTTCTATGAAATTTACGAACGCTGGCAAGAAATCAACGCGAACGAACTGGACCTGGACTATGATGTTG
TTTTCATCCACGATCCGCAGCCGGCGGGCCTGGTCAAATACCGTAAACGTGGTCTGTGGATTTGGCGCTGCCATATCGATGTGTCTA
CCCCGAACCCGGAAGTGTGGGGCTTCCTGCGTCGTTACGTTAGCCAGTACGATGCTGCGATTTTCCACATTCCAGAATTCGCGCGTG
ACGACCTGGAAATTCCGCAGCTGCTGATTCCGCCGTCCATCGATCCGCTGTCCCCGAAGAACGTTGAACTGCCGCCGACCGCGGTTG
AACGCGTGGTGCAAAAATTCGATGTCGACCCGGAAAGACCGATCCTGCTGCAGGTCTCGCGCTTCGATCGTGCGAAAGATCCGCTGG
GCGTTGTCGAGGCGTACCGTCTGGCGCGTCGTCGCGTTCCGGGTCTGCAGCTGGTTTATCTGGGCTCTCCGGCGCACGACGATCCGG
AAGGCGAAGCGGTTTATCGTGAGACGGTAGAAGCAGCGGGCGGTGATCCGGATATTCACCTGCTGATGCTGCCGCCGGACTCTCACT
ACGAAGTGAACGCGTTCCAGCGTGCTGCGACCGTTGTTATGCAGAAAAGCATCCGTGAAGGTTTTGGTCTGACCGTTAGCGAAGCGC
TGTGGAAAAAACGTCCGGTTATTGGCGGCAAAACCGGTGGTATCAAAATCCAGGTGATCCACGGCGTGACCGGCTTCCTGGCTACCA
GCCCGCGCACCGCGGCGCACTACGCGGTTTACCTGCTGCGTGAAAAACGTCTGCGTGAAGAAATGGGCGCGGCGGGCCGTGAACACG
TTCGTCGCAACTTCCTGATCACCCACCAGCTGCGTCGTTACCTGATGGCCATCGCGTACGTTGCGAAACGTGCGATGTGGAGCGATT
AATAAggtacc
```

Protein sequence

```
MGMIERYIQFVGEDEVDAIVKLAERLQDLSILHVNSTAAGGGVAEILNRLVPLMRELGLRVDWRVIRGDQEFFTVTKTFHNALQVGA
VEVPRRFYEIYERWQEINANELDLDYDVVFIHDPQPAGLVKYRKRGLWIWRCHIDVSTPNPEVWGFLRRYVSQYDAAIFHIPEFARD
DLEIPQLLIPPSIDPLSPKNVELPPTAVERVVQKFDVDPERPILLQVSRFDRAKDPLGVVEAYRLARRRVPGLQLVYLGSPAHDDPE
GEAVYRETVEAAGGDPDIHLLMLPPDSHYEVNAFQRAATVVMQKSIREGFGLTVSEALWKKRPVIGGKTGGIKIQVIHGVTGFLATS
PRTAAHYAVYLLREKRLREEMGAAGREHVRRNFLITHQLRRYLMAIAYVAKRAMWSD**
```

Protparam(3)

Number of amino acids:	405
Molecular weight:	46229.21
Theoretical pI:	6.66
Ext. coefficient	60850 (M^{-1} cm^{-1}, 280 nm, measured in water).
Abs 0.1% (=1 g/l)	1.316, assuming all pairs of Cys residues form cysteines
The instability index (II)	52.27 (unstable).
Aliphatic index:	98.67
GRAVY:	-0.208

Trehalose transferase DNA sequence from *Thermoproteus tenax* Kra 1

Non-codon optimized DNA sequence (GenBank: FN869859.1)

```
CTAGGGAGCCGAATGCCGCCCCGTTAAATAGAGGATCGTCATGAGGTAGCGCCTGAGCTGTTGAGTTATCAAGAAGTTGCGCCTCAC
GTGCTCTCTGCCTGCGGCGCCCATCTCCCTCCTCAGTCTCTTGTTCTTCAGCAGGTAGACAATATAGTGGGCCGCCGCCTTGGGGGA
GTCCACGAGGAAGCCAGTGACTCCGTGTATCACTTGGATTCTGATGCCTCCAGTGTTGCCTCCTATGACCGGCCTCCTCTTCCACAG
AGCCTCGCTGACAGTGAGCCCGAACCCCTCTCTTATCGACTTCTGTAAAACTACGGCGGCGGCCCTCTGGAAGGCGTTGACCTCTAT
GTGGCTGTTGGGCGGCAACATCAACAAGTGTATGTCTTTATCGTCTCCGGCCGCCCTAAGGGCCTCTCTATAGACTTCCTCCCCCTC
TGGGTCGTCCGAGGCGGGGCTGCCTAGGTAGACCAGCTGTACGTCTACATGCCGGCGGGCCAACTTGTAGGCCTCAATGACTCCAAC
AGGGTCCTTCGCTCTATCGAACCGAGATACTTGGAGCACTATAGGGCGCTCTGGATCAACGCCGTATTTCCTCACTATTCTGTCCAC
AGTGGCTCTGGGCAGAGGCACGTTCTTGGGGCTCAGAGGGTCTATCGAGGGCGGTATAGATATCTGGGGCACATCTAGATCGTCTCT
GGCAAACTCTGGTATCGAGACTATTACTCCATCGTAGGCCGATATATACCTCTTGAGGAAGGCCCAGACCTCGGGGTGGGGGTTGCT
GATGTCGATATGGCATCGCCAGATCCAGACCCCCCTCCTTTTATATCTAATGAGGCCAGCTGGCTGCGGATCGTGTATGAACACTAC
GTCGTAGTCCAGAGGGATCTCGCCGGCGTTGATCTCCTGCCATCTGTCGTATATCTCGAAGTACTCCCTGGGTATCGAGCCGGCGCC
TGTCTGAAGCGCGTTGTGAAACGACTTGGTGACTCTGAAGAACTCCTCGTTCCCCCTTATCACTTTCCACTCTACGTTGAGCCCCAG
CTCCCTCATCAGAGGGATCAATCTGTGGAGTATCTCGGCGACTCCGCCGCCGGCCGCCGTAGAGTTGATGTGTAGGATCGACAGATC
CCTCAGCCTTTCGGCGTATTTAAATATGGCGTTCAGCTCATGCTCGCCTATAAATTCGACGTAACGCTCTATCAT
```

Codon-optimized synthetic DNA sequence

```
ccatggGCATGATTGAACGTTACGTTGAATTCATCGGTGAACATGAACTGAACGCCATTTTTAAGTATGCTGAGCGTCTGCGCGATC
TGTCCATCCTGCACATTAATTCCACCGCCGCAGGTGGAGGCGTTGCGGAAATCCTGCATCGTCTGATTCCGCTGATGCGCGAACTGG
GCCTGAACGTAGAATGGAAAGTGATTCGTGGTAACGAGGAGTTTTTCCGCGTGACTAAAAGCTTCCACAATGCACTGCAGACCGGTG
CCGGCTCCATCCCACGTGAATACTTCGAAATCTATGACCGTTGGCAGGAAATCAACGCGGGTGAAATCCCACTGGATTACGATGTTG
TTTTCATTCATGACCCGCAACCGGCGGGCCTGATCCGTTATAAACGTCGTGGTGTGTGGATCTGGCGTTGCCACATCGATATCAGCA
ATCCGCACCCGGAAGTGTGGGCTTTCCTGAAACGTTACATCAGCGCGTACGATGGTGTGATCGTTTCTATTCCGGAATTCGCTCGTG
ATGATCTGGATGTTCCGCAGATTTCTATCCCGCCGAGCATCGATCCGCTGTCCCCGAAAAACGTGCCGCTGCCGCGTGCGACCGTTG
ATCGTATCGTTCGCAAATACGGCGTTGATCCGGAACGTCCGATCGTTCTGCAGGTTTCTCGCTTCGATCGTGCGAAAGACCCGGTTG
GCGTTATCGAAGCGTACAAACTGGCTCGTCGCCACGTGGATGTTCAGCTGGTGTACCTGGGCAGCCCGGCGTCTGATGACCCGGAAG
GCGAAGAAGTTTATCGTGAAGCGCTGCGCGCGGCGGGCGATGATAAAGATATCCATCTGCTGATGCTGCCGCCGAACTCTCACATCG
AAGTGAACGCGTTCCAGCGTGCGGCAGCGGTGGTGCTGCAGAAAAGCATTCGTGAAGGTTTCGGCCTGACCGTTTCTGAAGCGCTGT
GGAAACGCCGTCCGGTGATCGGCGGCAACACCGGCGGTATCCGTATCCAGGTTATCCACGGCGTTACCGGTTTCCTGGTTGATAGCC
CGAAAGCTGCTGCGCACTACATCGTTTACCTGCTGAAAAACAAACGTCTGCGTCGTGAAATGGGTGCGGCGGGCCGTGAACACGTTC
GTCGTAACTTCCTGATCACCCAGCAGCTGCGCCGTTACCTGATGACCATCCTGTACCTGACCGGTCGTCACTCTGCGCCGTAATAAgg
gtacc
```

Protein sequence

```
MGMIERYVEFIGEHELNAIFKYAERLRDLSILHINSTAAGGGVAEILHRLIPLMRELGLNVEWKVIRGNEEFFRVTKSFHNALQTGA
GSIPREYFEIYDRWQEINAGEIPLDYDVVFIHDPQPAGLIRYKRRGVWIWRCHIDISNPHPEVWAFLKRYISAYDGVIVSIPEFARD
DLDVPQISIPPSIDPLSPKNVPLPRATVDRIVRKYGVDPERPIVLQVSRFDRAKDPVGVIEAYKLARRHVDVQLVYLGSPASDDPEG
EEVYREALRAAGDDKDIHLLMLPPNSHIEVNAFQRAAAVVLQKSIREGFGLTVSEALWKRRPVIGGNTGGIRIQVIHGVTGFLVDSP
KAAAHYIVYLLKNKRLRREMGAAGREHVRRNFLITQQLRRYLMTILYLTGRHSAP**
```

Protparam (3)	
Number of amino acids:	403
Molecular weight:	45923.94
Theoretical pI:	9.09
Ext. coefficient	56840 (M^{-1} cm^{-1}, 280 nm, measured in water).
Abs 0.1% (=1 g/l)	1.238, assuming all pairs of Cys residues form cysteines
The instability index (II)	50.60 (unstable).
Aliphatic index:	101.84
GRAVY:	-0.206

Trehalose transferase DNA sequence from *Thermoproteus uzoniensis* 768-20

Non-codon optimized DNA sequence (GenBank: CP002590.1)

```
ATGATCGAGCGCTACGTCGAGTTCGTAGGGGAGCACGAGATAGACGCCATATTCAAATACGCCGAGAAGTTGAAGGACCTCTCCATA
CTGCACGTGAACTCGACGGCTGCCGGCGGCGGCGTCGCCGAGATACTGCACAGGCTGGTGCCTCTGATGAGGGAGCTGGGGCTCAAC
GCCGAGTGGAAAGTGATACGGGGGAGCCAAGACTTCTTCACCGTCACTAAGTCCTTCCACAACGCCCTCCAGACGGGGAAGGGCGAG
ATACCCGACGAGTACTTCAAGATATACGACGAGTGGCAGGAAATAAACGCCGGCGAGATACCGCTGGACTACGACGTCGTCTTCATA
CACGACCCCCAGCCCGCCGGCTTGGTCAAGTACAGGAAGAAGGGGACGTGGATATGGCGTTGCCACATAGACATAAGCAATCCACAC
CCCAAGGTCTGGGGCTTCCTGCGGGGCTACATATCCAAGTACGACGGCATGATAGTGTCCATACCCGAGTTCGCCAGAGACGACCTG
GACATCCCCCAGATAGCGATCCCGCCGTCTATAGACCCACTGAGCCCTAAGAACATGCCTCTGCCCCAGACGACCGTGGACAGAATA
GTCGACAAGTTCGGCGTGGATAGGGAGAGGCCCATAATTCTGCAGGTCTCCAGATACGACAGGGCCAAAGACCCCGTAGGCGTCATA
GAGTCCTTCAGGCTGGCGAAGCGCCACGTGCCCGACGCCCAGCTGGTCTACCTCGGGAGCCCCGCCACCGACGACCCGGAGGGCGAG
GTGGTCTACCGCGAGACTGTCGAGGCGGCTCGCGGCGAGAAGGACGTGCACCTCCTCATGTTGCCCCCCAACAGCCACGTGGAGGTT
AACGCGTTCCAGAGGGCGGCCACCGTGGTGATGCAGAAGTCTATAAAGGAGGGCTTCGGCCTCACGGTCAGCGAGGCTCTCTGGAAG
CGCAAGCCCGTCATAGGCGGCAACACGGGCGGCATAAGGATACAGGTGATAAACGGCGTCACGGGCTTCTTGGTGGACAGCCCCAAG
GCCGCCGCCTACTACCTCGTCTACCTCCTCAAGAACAAAAAGGTGAGGGAGGAGATGGGGGAGGCCGGCCGCGACCACGTCAGGAGG
AACTTCTTGATAACGCAACAGCTGAGGCGCTATTTAATGGCCATACTCTACCTCACCAAACGCCACGCCTCCTGA
```

Codon-optimized synthetic DNA sequence

```
ccatggAGCATCCTGCACGTTAACTCCACCGCGGCGGGTGGTGGTGTTGCGGAAATCCTGCACCGCCTGGTTCCGCTGATGCGTGAA
CTGGGTCTGAACGCTGAATGGAAAGTGATCCGTGGTAGCCAGGATTTCTTCACCGTTACCAAAAGCTTCCACAACGCGCTGCAGACC
GGTAAAGGTGAAATCCCGGATGAATATTTTAAAATCTACGACGAATGGCAGGAAATCAATGCAGGCGAAATACCGCTGGACTATGAT
GTTGTTTTCATTCACGATCCGCAGCCGGCAGGTCTGGTTAAATACCGCAAAAAAGGCACCTGGATTTGGCGTTGCCACATTGATATT
AGCAACCCGCACCCGAAAGTTTGGGGCTTCCTGCGCGGCTACATCAGCAAATATGACGGCATGATCGTTAGTATTCCGGAATTCGCA
CGTGATGATCTGGATATTCCGCAGATCGCGATCCCGCCATCTATTGACCCGCTGTCTCCGAAAAACATGCCGCTGCCGCAGACTACC
GTTGATCGTATCGTTGATAAATTCGGTGTTGATCGTGAACGTCCGATTATCCTGCAGGTTTCTCGTTACGACCGTGCGAAAGATCCC
GTTGGTGTGATTGAATCTTTCCGTCTGGCAAAACGCCATGTGCCGGATGCACAGCTGGTTTACCTGGGTTCGCCGGCGACTGATGAC
CCGGAAGGTGAAGTTGTGTATCGTGAAACCGTTGAAGCGGCTCGTGGTGAAAAAGACGTGCATCTGCTGATGCTGCCGCCGAACTCT
CACGTGGAAGTTAACGCTTTCCAGCGTGCGGCTACCGTAGTTATGCAGAAATCCATCAAAGAAGGCTTCGGTCTGACCGTGAGCGAA
GCTCTGTGGAAACGTAAACCTGTGATCGGCGGCAACACCGGTGGCATCCGTATTCAGGTAATCAACGGTGTTACCGGTTTCCTGGTT
GATTCTCCGAAAGCTGCAGCGTACTATCTGGTTTATCTGCTGAAAAACAAAAAAGTGCGTGAAGAAATGGGCGAAGCGGGCCGCGAT
CACGTTCGTCGTAACTTCCTGATCACCCAGCAGCTGCGTCGTTACCTGATGGCAATCCTGTACCTGACTAAACGTCATGCTAGCTAA
TAAggtacc
```

Protein sequence

```
MGMIERYVEFVGEHEIDAIFKYAEKLKDLSILHVNSTAAGGVAEILHRLVPLMRELGLNAEWKVIRGSQDFFTVTKSFHNALQTGK
GEIPDEYFKIYDEWQEINAGEIPLDYDVVFIHDPQPAGLVKYRKKGTWIWRCHIDISNPHPKVWGFLRGYISKYDGMIVSIPEFARD
DLDIPQIAIPPSIDPLSPKNMPLPQTTVDRIVDKFGVDRERPIILQVSRYDRAKDPVGVIESFRLAKRHVPDAQLVYLGSPATDDPE
GEVVYRETVEAARGEKDVHLLMLPPNSHVEVNAFQRAATVVMQKSIKEGFGLTVSEALWKRKPVIGGNTGGIRIQVINGVTGFLVDS
PKAAAYYLVYLLKNKKVREEMGEAGRDHVRRNFLITQQLRRYLMAILYLTKRHAS**
```

Protparam(3)

Number of amino acids:	403
Molecular weight:	45713.72
Theoretical pI:	7.81
Ext. coefficient	56840 (M^{-1} cm^{-1}, 280 nm, measured in water).
Abs 0.1% (=1 g/l)	1.243, assuming all pairs of Cys residues form cysteines
The instability index (II)	35.25 (stable).
Aliphatic index:	95.51
GRAVY:	-0.250

C-terminal *Py*TreT

Codon-optimized synthetic DNA sequence

```
CCATGGGCATGATTGAACGTTACATCCAGTTCGTTGGTGAAGATGAAGTTGATGCTATTGTGAAACTGGCGGAACGTCTGCAGGATC
TGTCTATCCTGCACGTTAACAGCACCGCTGCGGGCGGTGGTGTTGCTGAAATTCTGAACCGCCTGGTTCCGCTGATGCGTGAACTGG
GTCTGCGTGTTGATTGGCGTGTTATTCGCGGTGATCAGGAATTCTTTACCGTGACCAAAACCTTTCATAATGCCCTGCAAGTGGGCG
CCGTGGAAGTTCCGCGCCGCTTCTATGAAATTTACGAACGCTGGCAAGAAATCAACGCGAACGAACTGGACCTGGACTATGATGTTG
TTTTCATCCACGATCCGCAGCCGGCGGGCCTGGTCAAATACCGTAAACGTGGTCTGTGGATTTGGCGCTGCCATATCGATGTGTCTA
CCCCGAACCCGGAAGTGTGGGGCTTCCTGCGTCGTTACGTTAGCCAGTACGATGCTGCGATTTTCCACATTCCAGAATTCGCGCGTG
ACGACCTGGAAATTCCGCAGCTGCTGATTCCGCCGTCCATCGATCCGCTGTCCCCGAAGAACGTTGAACTGCCGCCGACCGCGGTTG
AACGCGTGGTGCAAAAATTCGATGTCGACCCGGAAAGACCGATCCTGCTGCAGGTCTCGCGCTTCGATCGTGCGAAAGATCCGCTGG
GCGTTGTCGAGGCGTACCGTCTGGCGCGTCGTCGCGTTCCGGGTCTGCAGCTGGTTTATCTGGGCTCTCCGGCGCACGACGATCCGG
AAGGCGAAGCGGTTTATCGTGAGACGGTAGAAGCAGCGGGCGGTGATCCGGATATTCACCTGCTGATGCTGCCGCCGGACTCTCACT
ACGAAGTGAACGCGTTCCAGCGTGCTGCGACCGTTGTTATGCAGAAAAGCATCCGTGAAGGTTTTGGTCTGACCGTTAGCGAAGCGC
TGTGGAAAAAACGTCCGGTTATTGGCGGCAAAACCGGTGGTATCAAAATCCAGGTGATCCACGGCGTGACCGGCTTCCTGGCTACCA
GCCCGCGCACCGCGGCGCACTACGCGGTTTACCTGCTGCGTGAAAAACGTCTGCGTGAAGAAATGGGCGCGGCGGGCCGTGAACACG
TTCGTCGCAACTTCCTGATCACCCACCAGCTGCGTCGTTACCTGATGGCCATCGCGTACGTTGCGAAACGTGCGATGTGGAGCGATC
TGGGCCACCATCACCATCACCATTAATAAGGTACC
```

Protein sequence

```
MGMIERYIQFVGEDEVDAIVKLAERLQDLSILHVNSTAAGGGVAEILNRLVPLMRELGLRVDWRVIRGDQEFFTVTKTFHNALQVGA
VEVPRRFYEIYERWQEINANELDLDYDVVFIHDPQPAGLVKYRKRGLWIWRCHIDVSTPNPEVWGFLRRYVSQYDAAIFHIPEFARD
DLEIPQLLIPPSIDPLSPKNVELPPTAVERVVQKFDVDPERPILLQVSRFDRAKDPLGVVEAYRLARRRVPGLQLVYLGSPAHDDPE
GEAVYRETVEAAGGDPDIHLLMLPPDSHYEVNAFQRAATVVMQKSIREGFGLTVSEALWKKRPVIGGKTGGIKIQVIHGVTGFLATS
PRTAAHYAVYLLREKRLREEMGAAGREHVRRNFLITHQLRRYLMAIAYVAKRAMWSDLGHHHHHH**
```

Protparam(3)

Number of amino acids:	413
Molecular weight:	47222.27
Theoretical pI:	6.82
Ext. coefficient	60850 (M^{-1} cm^{-1}, 280 nm, measured in water).
Abs 0.1% (=1 g/l)	1.289, assuming all pairs of Cys residues form cysteines
The instability index (II)	51.45 (unstable).
Aliphatic index:	97.7
GRAVY:	-0.242

N-terminal *Py*TreT

Codon-optimized synthetic DNA sequence (digested with SacI-HF and KpnI-HF)

*gagctc*ATGATCGAACGTTACATCCAGTTCGTTGGTGAAGATGAAGTTGATGCTATTGTTAAACTGGCGGAACGTCTGCAGGACCTG
TCTATCCTGCACGTGAACAGCACTGCGGCTGGCGGTGGCGTGGCGGAAATCCTGAACCGCCTGGTTCCGCTGATGCGTGAACTGGGT
CTGCGCGTTGATTGGCGTGTTATCCGTGGTGATCAGGAATTCTTCACTGTTACCAAAACCTTCCACAACGCTCTGCAGGTTGGTGCG
GTGGAAGTGCCGCGTCGCTTCTACGAAATTTACGAACGTTGGCAGGAAATTAACGCTAACGAACTGGACCTGGATTACGATGTTGTT
TTCATCCATGATCCGCAGCCGGCGGGTCTGGTGAAATACCGCAAACGTGGCCTGTGGATCTGGCGTTGTCACATCGATGTTAGCACC
CCGAACCCGGAAGTGTGGGGTTTCCTGCGCCGCTACGTGTCTCAGTATGACGCTGCGATCTTCCACATCCCGGAATTCGCGCGTGAT
GATCTGGAAATCCCGCAGCTGCTGATCCCGCCGTCTATCGATCCGCTGAGCCCGAAAAACGTTGAACTGCCGCCGACCGCGGTTGAA
CGTGTTGTTCAGAAATTCGATGTTGATCCGGAACGTCCGATTCTGCTGCAGGTTTCTCGTTTCGATCGCGCTAAAGATCCGCTGGGC
GTTGTTGAAGCATATCGTTTAGCGCGTCGTCGTGTGCCAGGCCTGCAGTTAGTGTATTTGGGTTCCCCGGCACATGATGATCCGGAA
GGTGAAGCAGTGTATCGCGAAACCGTTGAAGCAGCTGGCGGTGATCCTGACATCCATCTGCTGATGCTGCCGCCGGATTCTCACTAT
GAAGTTAACGCTTTCCAGCGTGCAGCAACTGTTGTGATGCAGAAATCTATTCGTGAAGGCTTTGGCCTGACCGTTAGTGAAGCTCTG
TGGAAAAAACGTCCGGTTATTGGTGGCAAAACTGGTGGTATCAAAATCCAGGTTATCCACGGTGTTACCGGTTTCCTGGCGACCTCT
CCGCGTACCGCTGCGCACTACGCTGTTTATCTGCTGCGTGAAAAACGTCTGCGCGAAGAAATGGGGGCTGCGGGCCGTGAACACGTT
CGTCGTAACTTCCTGATCACCCACCAGCTGCGTCGTTACCTGATGGCAATCGCGTACGTTGCGAAACGTGCGATGTGGTCTGATTAA
TAA*ggtac*

Protein sequence

*MGGSHHHHHHGMASMTGGQQMGRDLYDDDDKDRWGSEL*MIERYIQFVGEDEVDAIVKLAERLQDLSILHVNSTAAGGGVAEILNRLV
PLMRELGLRVDWRVIRGDQEFFTVTKTFHNALQVGAVEVPRRFYEIYERWQEINANELDLDYDVVFIHDPQPAGLVKYRKRGLWIWR
CHIDVSTPNPEVWGFLRRYVSQYDAAIFHIPEFARDDLEIPQLLIPPSIDPLSPKNVELPPTAVERVVQKFDVDPERPILLQVSRFD
RAKDPLGVVEAYRLARRRVPGLQLVYLGSPAHDDPEGEAVYRETVEAAGGDPDIHLLMLPPDSHYEVNAFQRAATVVMQKSIREGFG
LTVSEALWKKRPVIGGKTGGIKIQVIHGVTGFLATSPRTAAHYAVYLLREKRLREEMGAAGREHVRRNFLITHQLRRYLMAIAYVAK
RAMWSD**

Protparam(3)	
Number of amino acids:	441
Molecular weight:	50313.53
Theoretical pI:	6.32
Ext. coefficient	67840 (M^{-1} cm^{-1}, 280 nm, measured in water).
Abs 0.1% (=1 g/l)	1.348, assuming all pairs of Cys residues form cysteines
The instability index (II)	51.28 (unstable).
Aliphatic index:	92.6
GRAVY:	-0.318

Fusion protein of mCherry *Tu*TreT

Protein sequence mCherry (PDB:2H5Q)

MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDY
LKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKD
GGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYK

Codon optimized synthetic mCherry *Tu*TreT DNA sequence

ccATGGTGTCTAAAGGTGAAGAAGATAACATGGCGATCATCAAAGAGTTCATGCGTTTCAAAGTTCACATGGAAGGTAGCGTTAACG
GTCACGAATTTGAAATCGAAGGTGAAGGTGAAGGTCGTCCGTACGAAGGTACTCAGACCGCTAAACTGAAAGTAACTAAAGGCGGCC
CGCTGCCGTTCGCGTGGGATATCCTGAGCCCGCAGTTTATGTACGGTTCTAAAGCATATGTTAAACACCCGGCGGACATCCCGGATT
ACCTGAAACTGTCCTTCCCAGAAGGCTTCAAATGGGAACGTGTGATGAACTTCGAAGATGGTGGTGTTGTTACCGTTACCCAGGATA
GCAGCCTGCAGGACGGTGAGTTCATCTATAAAGTTAAACTGCGTGGTACTAACTTCCCGTCTGATGGTCCGGTTATGCAGAAGAAAA
CTATGGGCTGGGAAGCGTCTTCTGAACGTATGTACCCGGAAGACGGCGCACTGAAAGGTGAAATCAAACAGCGTCTGAAACTGAAAG
ACGGCGGTCACTATGACGCTGAAGTTAAAACCACCTATAAAGCTAAAAAACCGGTTCAGCTGCCAGGTGCATACAACGTGAACATTA
AACTGGACATCACCTCCCACAACGAAGACTACACCATCGTTGAACAGTACGAACGTGCTGAAGGTCGCCATTCCACTGGTGGTATGG
ATGAACTGTATAAAGGTGGCTCTGGCGGCGGTGGCTCTGGTGGTATGATTGAACGTTACGTGGAATTTGTTGGTGAACACGAAATCG
ACGCAATCTTCAAATACGCGGAAAAACTGAAAGACCTGTCTATCCTGCACGTTAACTCCACCGCTGCGGGTGGTGGTGTAGCTGAAA
TCCTGCACCGTCTGGTGCCGCTGATGCGTGAACTGGGCCTGAACGCTGAATGGAAAGTAATCCGTGGTTCTCAAGATTTCTTCACTG
TAACTAAATCTTTCCATAACGCGCTGCAGACTGGTAAAGGTGAAATCCCGGACGAATACTTTAAAATCTATGATGAATGGCAGGAAA
TTAACGCGGGTGAAATCCCGCTTGATTACGACGTTGTGTTCATCCACGATCCGCAGCCGGCTGGTCTGGTTAAATACCGTAAAAAAG
GCACTTGGATCTGGCGTTGCCACATCGACATTAGCAACCCGCACCCGAAAGTCTGGGGTTTCCTGCGTGGTTACATTAGCAAATACG
ACGGCATGATCGTGTCCATCCCGGAGTTCGCTCGTGACGACCTGGACATCCCGCAGATCGCAATCCCGCCGTCCATTGATCCGCTGA
GCCCGAAAAACATGCCACTGCCGCAGACCACCGTTGATCGTATCGTTGACAAATTCGGTGTTGATCGTGAACGTCCGATCATCCTGC
AGGTTAGCCGTTATGATCGTGCTAAAGACCCGGTTGGTGTTATTGAATCCTTTCGCCTGGCTAAACGCCACGTTCCGGACGCGCAGC
TGGTTTACCTGGGCTCTCCGGCTACTGATGACCCGGAAGGTGAAGTTGTTTACCGTGAAACCGTTGAAGCGGCACGTGGTGAAAAAG
ACGTACATCTGCTGATGCTGCCGCCGAACAGCCACGTTGAAGTTAACGCTTTCCAGCGTGCGGCTACCGTTGTTATGCAGAAATCCA
TTAAAGAAGGCTTCGGCCTGACCGTTTCTGAAGCGCTGTGGAAACGTAAACCGGTGATCGGTGGTAACACCGGTGGTATCCGTATCC
AGGTTATTAACGGTGTGACCGGTTTCCTGGTTGATAGCCCGAAAGCGGCGGCTTACTACCTGGTATACCTGCTGAAAAACAAAAAAG
TTCGTGAAGAAATGGGTGAAGCAGGTCGTGACCACGTTCGTCGTAACTTCCTGATCACCCAGCAGCTGCGTCGTTACCTGATGGCAA
TCCTGTACCTGACCAAACGTCACGCGTCTGGTCACCACCACCACCATCACTAAtaaggtacc

Protein sequence

MVSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDY
LKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLKLKD
GGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELYKGGSGGGGSGGMIERYVEFVGEHEID
AIFKYAEKLKDLSILHVNSTAAGGGVAEILHRLVPLMRELGLNAEWKVIRGSQDFFTVTKSFHNALQTGKGEIPDEYFKIYDEWQEI
NAGEIPLDYDVVFIHDPQPAGLVKYRKKGTWIWRCHIDISNPHPKVWGFLRGYISKYDGMIVSIPEFARDDLDIPQIAIPPSIDPLS
PKNMPLPQTTVDRIVDKFGVDRERPIILQVSRYDRAKDPVGVIESFRLAKRHVPDAQLVYLGSPATDDPEGEVVYRETVEAARGEKD
VHLLMLPPNSHVEVNAFQRAATVVMQKSIKEGFGLTVSEALWKRKPVIGGNTGGIRIQVINGVTGFLVDSPKAAAYYLVYLLKNKKV
REEMGEAGRDHVRRNFLITQQLRRYLMAILYLTKRHASGHHHHHH**

Protparam(3)

Number of amino acids:	654
Molecular weight:	73740.12
Theoretical pI:	6.58
Ext. coefficient	91220
Abs 0.1% (=1 g/l)	1.237, assuming all pairs of Cys residues form cysteines
The instability index (II)	34.19 (stable).
Aliphatic index:	80.90
GRAVY:	-0.462

N-terminal *Tu*TreT

Non-codon optimized DNA sequence (GenBank: CP002590.1):

ATGATCGAGCGCTACGTCGAGTTCGTAGGGGAGCACGAGATAGACGCCATATTCAAATACGCCGAGAAGTTGAAGGACCTCTCCATA
CTGCACGTGAACTCGACGGCTGCCGGCGGCGGCGTCGCCGAGATACTGCACAGGCTGGTGCCTCTGATGAGGGAGCTGGGGCTCAAC
GCCGAGTGGAAAGTGATACGGGGGAGCCAAGACTTCTTCACCGTCACTAAGTCCTTCCACAACGCCCTCCAGACGGGGAAGGGCGAG
ATACCCGACGAGTACTTCAAGATATACGACGAGTGGCAGGAAATAAACGCCGGCGAGATACCGCTGGACTACGACGTCGTCTTCATA
CACGACCCCCAGCCCGCCGGCTTGGTCAAGTACAGGAAGAAGGGGACGTGGATATGGCGTTGCCACATAGACATAAGCAATCCACAC
CCCAAGGTCTGGGGCTTCCTGCGGGGCTACATATCCAAGTACGACGGCATGATAGTGTCCATACCCGAGTTCGCCAGAGACGACCTG
GACATCCCCCAGATAGCGATCCCGCCGTCTATAGACCCACTGAGCCCTAAGAACATGCCTCTGCCCCAGACGACCGTGGACAGAATA
GTCGACAAGTTCGGCGTGGATAGGGAGAGGCCCATAATTCTGCAGGTCTCCAGATACGACAGGGCCAAAGACCCCGTAGGCGTCATA
GAGTCCTTCAGGCTGGCGAAGCGCCACGTGCCCGACGCCCAGCTGGTCTACCTCGGGAGCCCCGCCACCGACGACCCGGAGGGCGAG
GTGGTCTACCGCGAGACTGTCGAGGCGGCTCGCGGCGAGAAGGACGTGCACCTCCTCATGTTGCCCCCCAACAGCCACGTGGAGGTT
AACGCGTTCCAGAGGGCGGCCACCGTGGTGATGCAGAAGTCTATAAAGGAGGGCTTCGGCCTCACGGTCAGCGAGGCTCTCTGGAAG
CGCAAGCCCGTCATAGGCGGCAACACGGGCGGCATAAGGATACAGGTGATAAACGGCGTCACGGGCTTCTTGGTGGACAGCCCCAAG
GCCGCCGCCTACTACCTCGTCTACCTCCTCAAGAACAAAAAGGTGAGGGAGGAGATGGGGGAGGCCGGCCGCGACCACGTCAGGAGG
AACTTCTTGATAACGCAACAGCTGAGGCGCTATTTAATGGCCATACTCTACCTCACCAAACGCCACGCCTCCTGA

Codon optimized synthetic DNA sequence (*Sac*I (*gagctc*); *Kpn*I (*ggtac*)

*gagctc*ATGATTGAACGTTACGTTGAATTTATCGGCGAACATGAACTGAACGCGATTTTTAAATACGCTGAACGTCTGCGTGATCTG
AGCATCCTGCACATTAATTCCACTGCAGCGGGTGGCGGTGTTGCTGAAATCCTGCACCGCTTAATCCCGCTGATGCGTGAACTGGGT
CTGAACGTTGAATGGAAAGTGATCCGTGGTAACGAAGAATTCTTTCGTGTTACGAAATCTTTTCATAACGCACTGCAGACCGGTGCA
GGTTCTATCCCGCGTGAATACTTCGAAATCTATGATCGTTGGCAGGAAATTAACGCGGGCGAAATCCCGCTGGACTATGATGTTGTA
TTCATTCATGATCCGCAGCCTGCGGGTCTGATTCGTTATAAACGTCGTGGCGTTTGGATTTGGCGTTGCCACATTGACATCTCCAAT
CCGCACCCGGAAGTGTGGGCATTTCTGAAACGTTACATTTCTGCATATGACGGTGTTATCGTTTCCATTCCGGAATTCGCGCGTGAT
GACCTGGATGTTCCGCAGATCAGCATCCCGCCGAGCATCGACCCGCTGAGCCCGAAAAACGTGCCGCTCCCGCGCGCTACCGTTGAC
CGTATCGTGCGTAAATACGGCGTGGATCCGGAACGTCCGATCGTTCTGCAGGTTAGCCGTTTCGATCGTGCCAAAGATCCGGTGGGT
GTTATCGAAGCGTACAAACTGGCGCGCCGTCACGTGGACGTTCAGCTGGTGTACCTGGGTTCCCCGGCGAGCGATGATCCGGAAGGT
GAAGAAGTATACCGCGAAGCTCTGCGTGCGGCGGGTGATGATAAAGATATCCACCTCCTGATGCTGCCGCCGAACTCCCACATCGAA
GTTAACGCGTTCCAGCGCGCGGCGGCGGTTGTTCTGCAGAAATCTATCCGTGAAGGCTTCGGCCTGACCGTGTCTGAAGCGCTGTGG
AAACGTCGCCCGGTTATCGGTGGTAACACCGGCGGCATCCGTATTCAGGTTATCCACGGCGTTACCGGCTTCCTGGTAGATAGCCCG
AAAGCAGCGGCGCACTACATCGTGTATCTGCTGAAAAACAAACGTCTGCGCCGTGAAATGGGCGCTGCTGGCCGTGAACACGTTCGT
CGTAACTTCCTGATCACCCAGCAGCTGCGTCGTTACCTGATGACCATCCTGTACCTGACCGGTCGTCACAGCGCGCCCTAATAA*ggt*
ac

Protein sequence

MGGSHHHHHHGMASMTGGQQMGRDLYDDDDKDRWGSELMIERYVEFIGEHELNAIFKYAERLRDLSILHINSTAAGGGVAEILHRLI
PLMRELGLNVEWKVIRGNEEFFRVTKSFHNALQTGAGSIPREYFEIYDRWQEINAGEIPLDYDVVFIHDPQPAGLIRYKRRGVWIWR
CHIDISNPHPEVWAFLKRYISAYDGVIVSIPEFARDDLDVPQISIPPSIDPLSPKNVPLPRATVDRIVRKYGVDPERPIVLQVSRFD
RAKDPVGVIEAYKLARRHVDVQLVYLGSPASDDPEGEEVYREALRAAGDDKDIHLLMLPPNSHIEVNAFQRAAAVVLQKSIREGFGL
TVSEALWKRRPVIGGNTGGIRIQVIHGVTGFLVDSPKAAAHYIVYLLKNKRLRREMGAAGREHVRRNFLITQQLRRYLMTILYLTGR
HSAP**

Protparam(3)

Number of amino acids:	403
Moleuclar weight:	50008.26
Theoretical pI:	7.92
Ext. coefficient:	63830 (M^{-1} cm^{-1}, measured in water).
Abs 0.1% (=1 g/l)	1.276 (assuming all pairs of Cys residues from cysteines)
The instability index(II)	49.74 (unstable)
Aliphatic index:	95.49
GRAVY:	-0.25

References

[1] a) L. O. Martins, R. Huber, H. Huber, K. O. Stetter, M. S. Da Costa, H. Santos, *Applied and Environmental Microbiology* **1997**, *63*, 896-902; b) L. K. Conlin, H. C. M. Nelson, *Molecular and Cellular Biology* **2007**, *27*, 1505-1515; c) O. Fernandez, L. Béthencourt, A. Quero, R. S. Sangwan, C. Clément, *Trends in Plant Science* **2010**, *15*, 409-417; d) K. Liu, Y. Dong, Y. Huang, J. L. Rasgon, P. Agre, *Proceedings of the National Academy of Sciences* **2013**, *110*, 17504-17509; e) D. D. Martin, R. A. Ciulla, M. F. Roberts, *Applied and Environmental Microbiology* **1999**, *65*, 1815-1825.

[2] S. Ajito, M. Hirai, H. Iwase, N. Shimizu, N. Igarashi, N. Ohta, *Physica B: Condensed Matter* **2018**, *551*, 249-255.

[3] L. Cordone, G. Cottone, S. Giuffrida, *Journal of Physics: Condensed Matter* **2007**, *19*, 205110.

[4] C. Schebor, L. Burin, M. a. d. P. Buera, J. Chirife, *LWT - Food Science and Technology* **1999**, *32*, 481-485.

[5] R. Zhang, Y. T. Pan, S. He, M. Lam, G. D. Brayer, A. D. Elbein, S. G. Withers, *Journal of Biological Chemistry* **2011**, *286*, 35601-35609.

[6] a) R. K. Asthana, S. Srivastava, A. P. Singh, A. M. Kayastha, S. P. Singh, *Journal of Plant Physiology* **2005**, *162*, 1030-1037; b) A. Wolf, R. Krämer, S. Morbach, *Molecular Microbiology* **2003**, *49*, 1119-1134.

[7] a) K. Saito, T. Kase, E. Takahashi, S. Horinouchi, *Applied and Environmental Microbiology* **1998**, *64*, 4340-4345; b) C. Goedl, A. Schwarz, M. Mueller, L. Brecker, B. Nidetzky, *Carbohydrate Research* **2008**, *343*, 2032-2040; c) W. J. B. Wannet, H. J. M. Op den Camp, H. W. Wisselink, C. van der Drift, L. J. L. D. Van Griensven, G. D. Vogels, *Biochimica et Biophysica Acta (BBA) - General Subjects* **1998**, *1425*, 177-188; d) C. Goedl, R. Griessler, A. Schwarz, B. Nidetzky, *The Biochemical Journal* **2006**, *397*, 491-500.

[8] a) H. Kizawa, K.-i. Miyagawa, Y. Sugiyama, *Bioscience, Biotechnology, and Biochemistry* **1995**, *59*, 1908-1912; b) K. Maruta, H. Watanabe, T. Nishimoto, M. Kubota, H. Chaen, S. Fukuda, M. Kurimoto, Y. Tsujisaka, *Journal of Bioscience and Bioengineering* **2006**, *101*, 385-390; c) Y. Inoue, K. Ishii, T. Tomita, T. Yatake, F. Fukui, *Bioscience, Biotechnology, and Biochemistry* **2002**, *66*, 1835-1843; d) J. Van der Borght, C. Chen, L. Hoflack, L. Van Renterghem, T. Desmet, W. Soetaert, *Applied and Environmental Microbiology* **2011**, *77*, 6939-6944.

[9] a) S.-I. Ryu, J.-E. Kim, N. T. Huong, E.-J. Woo, S.-K. Moon, S.-B. Lee, *Enzyme and Microbial Technology* **2010**, *47*, 249-256; b) T. Kouril, M. Zaparty, J. Marrero, H. Brinkmann, B. Siebers, *Archives of Microbiology* **2008**, *190*, 355-369; c) A. Nobre, S. Alarico, C. Fernandes, N. Empadinhas, M. S. Da Costa, *Journal of Bacteriology* **2008**, *190*, 7939-7946; d) E.-J. Woo, S.-I. Ryu, H.-N. Song, T.-Y. Jung, S.-M. Yeon, H.-A. Lee, B. C. Park, K.-H. Park, S.-B. Lee, *Journal of Molecular Biology* **2010**, *404*, 247-259; e) S.-I. Ryu, J.-E. Kim, E.-J. Kim, S.-K. Chung, S.-B. Lee, *Process Biochemistry* **2011**, *46*, 128-134; f) Q. Qu, S. J. Lee, W. Boos, *Journal of Biological Chemistry* **2004**, *279*, 47890-47897.

[10] a) D. Tischler, S. Niescher, S. R. Kaschabek, M. Schlömann, *FEMS Microbiology Letters* **2013**, *342*, 113-122; b) S. S. Lee, S. Y. Hong, J. C. Errey, A. Izumi, G. J. Davies, B. G. Davis, *Nature Chemical Biology* **2011**, *7*, 631; c) M. Zaparty, A. Hagemann, C. Bräsen, R. Hensel, A. N. Lupas, H. Brinkmann, B. Siebers, *PLOS ONE* **2013**, *8*, e61354; d) Y. Miao, J. L. Tenor, D. L. Toffaletti, S. A. Maskarinec, J. Liu, R. E. Lee, J. R. Perfect, R. G. Brennan, *mBio* **2017**, *8*; e) F. S. Cardoso, R. F. Castro, N. Borges, H. Santos, *Microbiology* **2007**, *153*, 270-280.

[11] a) X. Cai, I. Seitl, W. Mu, T. Zhang, T. Stressler, L. Fischer, B. Jiang, *Applied Microbiology and Biotechnology* **2018**, *102*, 2965-2976; b) M. K. O'Neill, B. F. Piligian, C. D. Olson, P. J. Woodruff, B. M. Swarts, *Pure and Applied Chemistry. Chimie pure et appliquee* **2017**, *89*, 1223-1249.

[12] B. L. Urbanek, D. C. Wing, K. S. Haislop, C. J. Hamel, R. Kalscheuer, P. J. Woodruff, B. M. Swarts, *ChemBioChem* **2014**, *15*, 2066-2070.

[13] Z. Yang, L. Zhang, Y. Zhang, T. Zhang, Y. Feng, X. Lu, W. Lan, J. Wang, H. Wu, C. Cao, X. Wang, *PLOS ONE* **2011**, *6*, e22981.

[14] B. Nidetzky, A. Gutmann, C. Zhong, *ACS Catalysis* **2018**, *8*, 6283-6300.

[15] a) A. Sabogal, D. C. Rio, *Protein Science* **2010**, *19*, 2210-2218; b) S. J. Davis, R. D. Vierstra, *Plant Molecular Biology* **1998**, *36*, 521-528; c) M. A. Coleman, V. H. Lao, B. W. Segelke, P. T. Beernink, *Journal of Proteome Research* **2004**, *3*, 1024-1032; d) W. W. Su, *Microbial Cell Factories* **2005**, *4*, 12-12.
[16] L. M. Meints, A. W. Poston, B. F. Piligian, C. D. Olson, K. S. Badger, P. J. Woodruff, B. M. Swarts, *Journal of Visualized Experiments* **2017**, 54485.
[17] T. Lazaridis, I. Lee, M. Karplus, *Protein Science* **1997**, *6*, 2589-2605.
[18] a) J. D. Trent, *FEMS Microbiology Reviews* **1996**, *18*, 249-258; b) L. Ditzel, J. Löwe, D. Stock, K.-O. Stetter, H. Huber, R. Huber, S. Steinbacher, *Cell* **1998**, *93*, 125-138; c) K. Kagawa Hiromi, T. Yaoi, L. Brocchieri, R. A. McMillan, T. Alton, D. Trent Jonathan, *Molecular Microbiology* **2003**, *48*, 143-156; d) J. D. Trent, E. Nimmesgern, J. S. Wall, F. U. Hartl, A. L. Horwich, *Nature* **1991**, *354*, 490; e) *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **1993**, *339*, 313-326; f) S. Marco, D. Ureña, J. L. Carrascosa, T. Waldmann, J. Peters, R. Hegerl, G. Pfeifer, H. Sack-Kongehl, W. Baumeister, *FEBS Letters* **1994**, *341*, 152-155; g) J. M. Archibald, J. M. Logsdon, W. F. Doolittle, *Current Biology* **1999**, *9*, 1053-S1056.
[19] M. Klumpp, W. Baumeister, *FEBS Letters* **1998**, *430*, 73-77.
[20] K. E. Scholz, B. Kopka, A. Wirtz, M. Pohl, K.-E. Jaeger, U. Krauss, *Applied and Environmental Microbiology* **2013**, *79*, 4727-4733.
[21] N. C. Shaner, R. E. Campbell, P. A. Steinbach, B. N. G. Giepmans, A. E. Palmer, R. Y. Tsien, *Nature Biotechnology* **2004**, *22*, 1567.
[22] S. Jevševar, V. Gaberc-Porekar, I. Fonda, B. Podobnik, J. Grdadolnik, V. Menart, *Biotechnology Progress* **2008**, *21*, 632-639.
[23] E. García-Fruitós, N. González-Montalbán, M. Morell, A. Vera, R. M. Ferraz, A. Arís, S. Ventura, A. Villaverde, *Microbial Cell Factories* **2005**, *4*, 27.
[24] J. Seras-Franzoso, S. Peternel, O. Cano-Garrido, A. Villaverde, E. García-Fruitós, *Methods in Molecular Biology* **2015**, *1258*, 293-305.
[25] M. Diener, B. Kopka, M. Pohl, K. E. Jaeger, U. Krauss, *ChemCatChem* **2015**, *8*, 142-152.
[26] a) R. Kloss, M. H. Limberg, U. Mackfeld, D. Hahn, A. Grünberger, V. D. Jäger, U. Krauss, M. Oldiges, M. Pohl, *Scientific Reports* **2018**, *8*, 5856; b) U. Krauss, V. D. Jäger, M. Diener, M. Pohl, K.-E. Jaeger, *Journal of Biotechnology* **2017**, *258*, 136-147.
[27] E. Noor, H. S. Haraldsdóttir, R. Milo, R. M. T. Fleming, *PLOS Computational Biology* **2013**, *9*, e1003098.
[28] a) A. Gutmann, C. Krump, L. Bungaruang, B. Nidetzky, *Chemical Communications* **2014**, *50*, 5465-5468; b) A. Gutmann, B. Nidetzky, *Advanced Synthesis & Catalysis* **2016**, *358*, 3600-3609.
[29] J. D. Fox, D. S. Waugh, in *E. coli Gene Expression Protocols* (Ed.: P. E. Vaillancourt), Humana Press, Totowa, NJ, **2003**, pp. 99-117.
[30] M. E. Kimple, A. L. Brill, R. L. Pasker, *Current Protocols in Protein science***2013**, *73*, Unit-9.9.
[31] G. Žoldák, R. Šut'ák, M. Antalík, M. Sprinzl, E. Sedlák, *European Journal of Biochemistry* **2003**, *270*, 4887-4897.
[32] R. P. Gibson, C. A. Tarling, S. Roberts, S. G. Withers, G. J. Davies, *The Journal of Biological Chemistry* **2004**, *279*, 1950-1955.
[33] K.-P. Huang, J. C. Robinson, *The Journal of Biological Chemistry* **1977**, *252*, 3240-3244.
[34] B. W. Murray, S. Takayama, J. Schultz, C.-H. Wong, *Biochemistry* **1996**, *35*, 11183-11195.

4

Comparison of Enzymes Immobilized on Immobeads and Inclusion Bodies: A Case Study of a Trehalose Transferase

This chapter is based on:

L. Mestrom, S. R. Marsden, D. McMillan, Rob Schoevaart, Peter-Leon Hagedoorn, Ulf Hanefeld, **2020**. *ChemCatChem*, **2020**, *12*, 3249-3256.

"Purple haze, all in my brain
Lately things they don't seem the same
Actin' funny, but I don't know why
Excuse me while I kiss the sky"
*Jimi Hendrix, "*Heaven Research*",* Purple Haze, **1967**

4.1 Introduction

Successful application of enzymes for the production of complex food products and chemicals depends on the recyclability of the biocatalyst and its ease of separation.[1] Enzyme immobilization is a popular strategy for enzyme recycling and improvement of downstream processing. Since different enzyme immobilization procedures influence the stability, activity, and selectivity of biocatalysts, a wide number of different methodologies have been developed.[2] The use of either carrier-free aggregates or carrier-attached enzymes are two of the most common techniques of enzyme immobilization (Fig. 1).[3] Catalytically active inclusion bodies (CatIBs) have been described as new form of carrier-free immobilization[4] which has been successful for different enzyme classes[5], such as hydrolases[6], oxidoreductase[7], lyases[8], and transferases[9]. The simplicity of chromatography-free production and purification of CatIBs has been attributed the be the key to their success.[8]

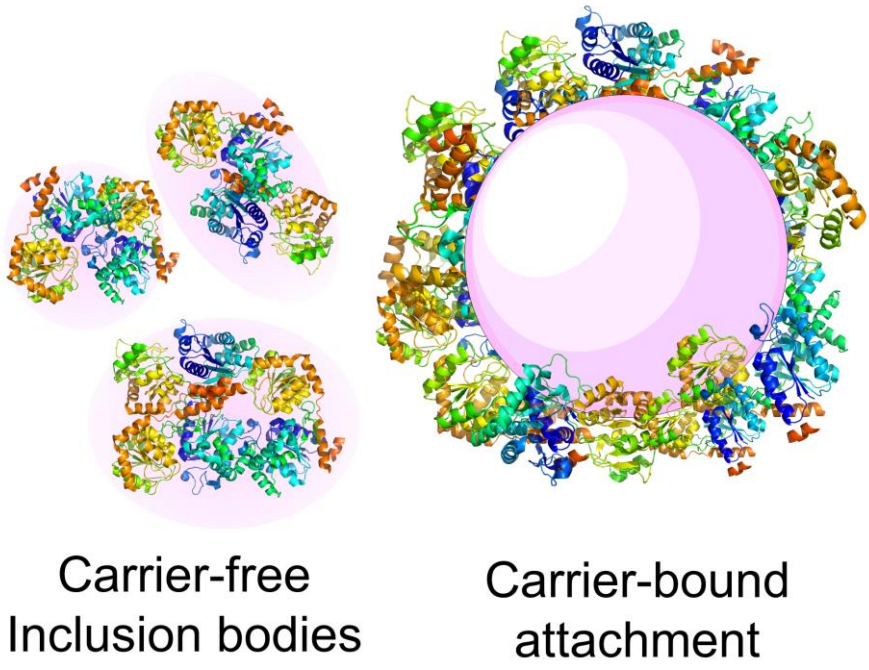


Figure 1. Schematic depiction of carrier-free CatIBs and carrier-attached attachment of enzyme.

To our knowledge, a direct comparison of enzymatic catalytic performance using the same enzyme as carrier-attached or carrier-free biocatalytic formulation has not been performed yet. One of the challenges in such a comparison is the characterization of carrier-free CatIBs and their material properties. A high polydispersity in size and morphology of CatIBs complicates the analysis of diffusion limitation and the effect on catalytic activity within these particles. The CatIBs particles contain (partially) misfolded protein[10] possibly resulting in lower catalytic activity, independent of mass transfer limitations. The use of soft CatIBs can be disadvantageous, as continuous processes in a packed-bed plug flow reactor setup leads to pressure drops with compressible materials. It is not surprising that typically (fed-)batch processes have been reported with CatIBs.[5] Additional formulation steps are required to engineer the mechanical properties of CatIBs to broaden the choice of reactors.[11] In contrast, for carrier-attached enzymes the choice of reactor and the carriers dictate the material properties of the immobilization matrix. Depending on the properties of the carrier, they can be used for different reactor types. A case by case optimization of enzyme immobilization with different attachment modes of carriers are required to guarantee optimal enzyme stability and activity.[12]

Despite the plethora of enzyme immobilization methodologies, including their optimization strategies to increase their performance[13], it remains challenging to assess the reduction in catalytic activity of immobilized enzymes. The immobilization procedure for attaching enzymes to carriers might affect the stability, activity, selectivity, and can influence the apparent inhibition.[14] Therefore well-characterized, commercial carriers were utilized in this study. The screening conditions were kept similar according to a standard immobilization protocol. Different binding interactions can lead to (partial) protein denaturation, and enzymes can be distributed in homogenously within an immobilization matrix. The use of fluorescent proteins provides insight in these aspects during enzyme immobilization. The enzyme used in this study, trehalose transferase from *Thermoproteus uzoniensis* (*Tu*TreT), was fused to fluorescent protein mCherry.[15] *Tu*TreT couples a nucleotide sugar donor and sugar acceptor in a (1→1)-α,α-glycosidic bond resulting in the formation of trehalose.[15] Although TreT has been applied for the synthesis of trehalose and its analogues[16], TreT has been proven difficult to express as soluble enzyme.[15, 17] The fusion of mCherry to *Tu*TreT resulted in increased solubility although still a large part of the expressed protein

was in the form CatIBs.[15] Upon denaturation of mCherry *Tu*TreT, the chromophore of mCherry changes color from purple to green.[18] Using the two different excitation and emission spectra, the quantitative and qualitative assessment between native and denatured mCherry *Tu*TreT is possible.[15] The potential of visualizing protein aggregation and the distribution of native and denatured protein with and without carriers allow the evaluation of different protein immobilization procedures.

The aim of this study was to the comparison of the performance of a carrier-attached and carrier-free biocatalytic formulation of mCherry *Tu*TreT. The characterization of the two biocatalytic formulations of mCherry *Tu*TreT was combined with essential parameters to measure the performance of each formulation: catalytic activity, operational stability, and ease of biocatalyst production. For the carrier-attached enzyme, twelve preexisting carriers with mCherry *Tu*TreT were explored with covalent, hydrophobic, or electrostatic interactions as attachment methodology. The aim of this screening of carrier materials was to select the immobilized enzyme with the highest catalytic activity for further comparison. The CatIBs were extensively characterized, assessing the quality and quantity of mCherry *Tu*TreT and the effects of the binding interactions to various carriers. For both immobilization techniques the fluorescent protein was used as a probe to assess the distribution and quality of the immobilized enzyme. The fusion of mCherry to *Tu*TreT allowed direct spectrophotometric quantification and visualization of the enzyme in the native and denatured state. The CatIBs outperformed immobilized enzyme in their simplicity of biocatalyst production resulting in high enzyme productivity, while enzyme immobilized on carrier materials showed a higher catalytic activity and a more robust performance under batch process conditions.

4.2 Results and discussion

To test the fluorescent protein mCherry to *Tu*TreT as a probe the immobilization on twelve different commercial carriers was performed. All carriers were organic polymers with similar morphology, size, and porosity. These carrier materials utilize different types of attachment interactions: covalent linkages using epoxide-functionalized polymers, absorption on hydrophobic materials, and electrostatic interactions with ionic carriers. mCherry *Tu*TreT was produced and purified as was described previously.[15] The progress of immobilization was determined by visual inspection, since the intensity of purple color of the

immobilized enzyme and supernatant is proportional to the protein content.[15] Classification into ‘high’ (Fig. 2a) and ‘low’ (Fig. 2b) immobilization efficiency was straightforward, and denaturation was readily identified by observing a change in color from purple (folded mCherry) to green (denatured mCherry) (Fig. 2c). With fluorescence microscopy three main states of immobilized *Tu*TreT fused to mCherry were observed: (i) uniform distribution on the surface without denaturation (Fig. 2d); (ii) inhomogeneous distribution of native and denatured protein (Fig. 2e); (iii) or the occurrence of fibrillar denatured protein aggregates on the surface of the carrier (Fig. 2f).

The conventional characterization of the immobilization of enzymes on carriers relies on accurate protein quantification on the carrier material and the specific activity of the immobilized enzyme. Loss of specific enzyme activity upon immobilization is often attributed to protein denaturation or diffusion limitation of the substrate. We performed this conventional characterization together with fluorescence microscopy using mCherry as a reporter for the enzyme. 100 mg of enzyme carrier was added to 5.0 mg mCherry *Tu*TreT in 1.0 mL of HEPES buffer (50 mM, pH 7.0). Both the activity and amount of any remaining soluble enzyme was measured before and after immobilization. mCherry *Tu*TreT is a monomer in solution,[15] and harbors 50 lysine residues per monomer corresponding to 0.664 mmol of free amino groups per g of protein added to the carrier.

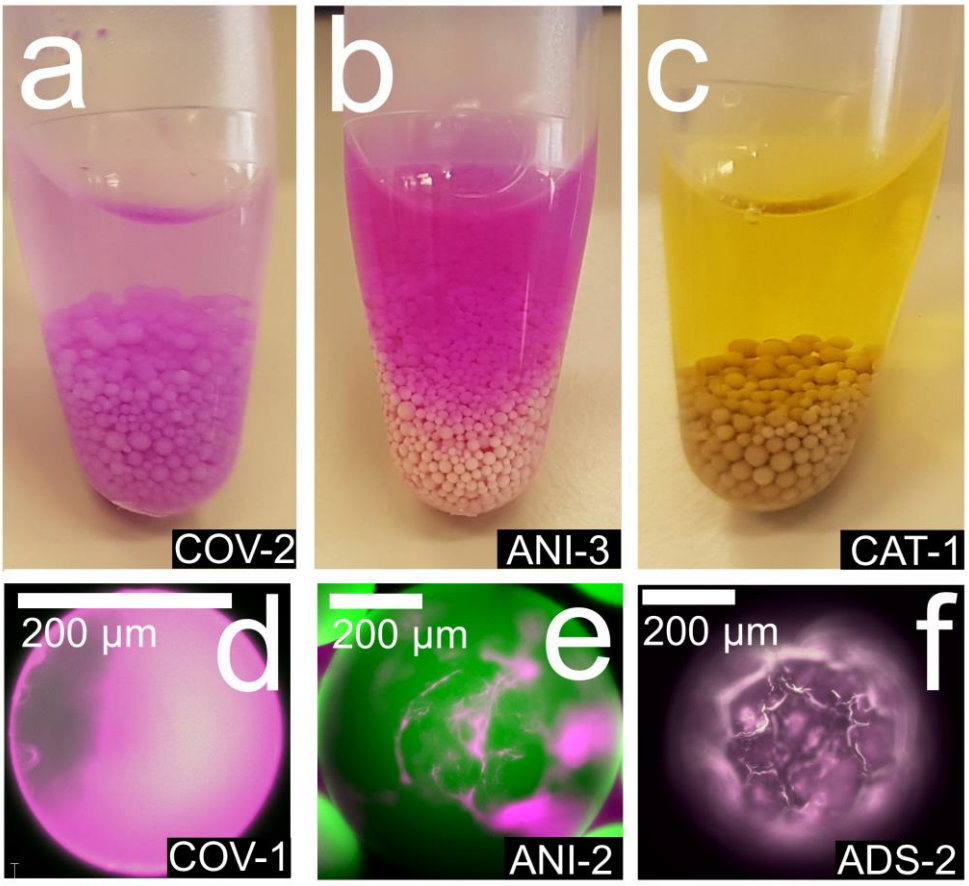


Figure 2. Visual inspection of the immobilisation procedure showing a high immobilisation yield of mCherry *Tu*TreT COV-2 (95% yield, 48 µg mg^{-1} carrier) in (**a**), moderate loading of ANI-3 in (**b**) (49%, 19 µg mg^{-1} carrier), and denatured mCherry *Tu*TreT on CAT-1 (**c**). Fluorescence microscopy of carrier-attached mCherry *Tu*TreT on COV-1 was homogenously distributed over the surface and inside the particles (**d**), where ANI-2 shows inhomogeneous distribution of native versus denatured enzyme (**e**). Aggregation of GFP-like fibrillar mCherry *Tu*TreT was observed on the surface of ADS-2 in (**f**).

The highest immobilization yields and specific activities of carrier-attached mCherry *Tu*TreT were found for immobilization using covalent interactions (Fig 3a). Fluorescence microscopy showed a homogenous distribution of protein over the surface without protein denaturation. Lower catalytic activity and lower protein

yields were observed after immobilization using hydrophobic interactions (Fig. 3b). Green fibrillar protein aggregates were observed on the hydrophobic surface of these carriers, suggesting that the lower catalytic activity observed was due to enzyme denaturation of the immobilized enzyme. This is in agreement with previous reports of protein aggregation or adverse folding effects using hydrophobic carriers.[2a, 19] For carriers with electrostatic attachment modes, the cationic carrier (CAT) showed complete denaturation of the protein without any recovery of the enzyme activity after immobilization (Fig. 3). The anionic carrier materials displayed an inhomogeneous distribution of native and denatured enzyme on the carrier materials. ATR-FTIR, a widely used to measure the presence of proteins on the carriers, showed the characteristic amide I and II vibrations on the carriers, showing protein presence for all the carriers tested.[20] Clearly the use of fluorescence microscopy showed more than just the presence of immobilized protein, as it yielded information on the native or denatured state of the protein and its distribution on the carrier material.

In light of the above described results, COV-1 was selected as the model system for the carrier-attached mCherry *Tu*TreT. The rate of immobilization on different amounts of carrier material with a fixed amount of soluble mCherry *Tu*TreT (1.0 mg mL^{-1}) was measured using spectroscopic UV-analysis. This characterization allows the determination of the surface coverage of the spherical particles and the immobilization process over time. Based on the ratio of amino groups (mCherry *Tu*TreT) to epoxide groups (COV-1), we determined the amount of accessible epoxide groups to be consistent with a surface coverage of approximately 40%.

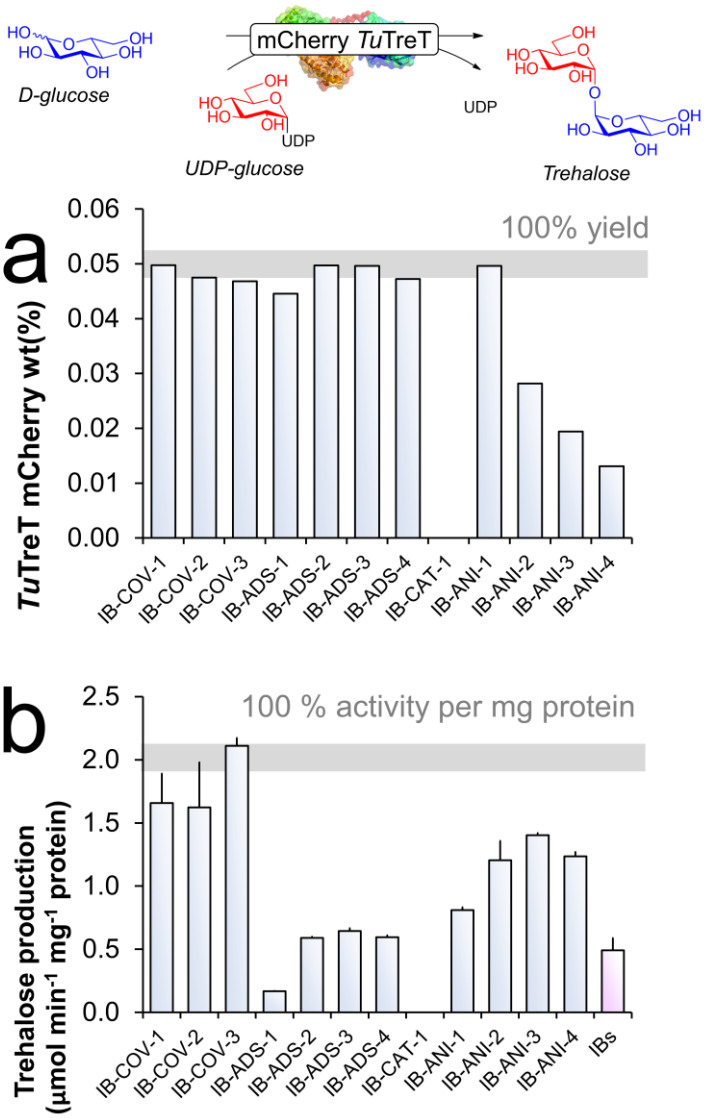


Figure 3. Immobilisation of mCherry *Tu*TreT on a wide range of carriers using either covalent (COV), hydrophobic (ADS), cationic (CAT), or anionic (ANI) binding modes. The immobilisation efficiency is high except for anionic binding modes (**a**), as was determined by relative decrease in protein content in solution (ε_{587nm} = 0.9979 mg^{-1} mL cm^{-1}). The the specific activity is highest for covalent binding modes (**b**), as was determined by measuring the activity per amount of protein on the carrier material. **Reaction conditions**: D-glucose (10 mM), UDP-D-glucose (40 mm), $MgCl_2$ (20 mM), HEPES (50 mM, pH 7.0).

The characterisation of carrier-free CatIBs as biocatalyst showed that mCherry could be used effectively as a reporter for the rapid analysis of both protein content and the state of denaturation of *Tu*TreT. This is particular useful in complex mixtures like inclusion bodies, since they are often contaminated with variable quantities of *E. coli* cell debris (i.e. other proteins)[8, 21]. When protein expression in *E. coli* is high, the mCherry *Tu*TreT inclusion bodies showed low amounts of contaminating proteins.[15] After separation from other cellular material, we took advantage of the sodium dodecyl sulphate (SDS) stability of *Tu*TreT and inclusion bodies were solubilised using 2% wt. SDS.[15] The solubilized protein yield was then determined spectrophotometrically ($\varepsilon_{587\ nm}$ = 0.9979 mg mL^{-1} cm^{-1}) to correspond to 3 % wt. mCherry *Tu*TreT in the CatIBs. Lyophilisation of the CatIBs allowed concentration of mCherry *Tu*TreT to 10 % wt. without loss of enzyme activity due to denaturation. Further analysis of the lyophilized CatIBs with ATR-FTIR showed characteristic amide I, II, III, and A vibrations which are typically observed for proteins (Fig. 4a). Structural analysis with powder X-ray diffraction (XRD) of the CatIBs revealed the presence of (poorly) crystalline cross-β sheet interactions, with an interstrand distance of 4.7 Å and intersheet distance of 10 Å (Fig. 4b-c). Similar cross β-sheet interaction distances have been reported for other IBs.[10] The hydrogen bonding of cross β-sheet interactions might be the major interaction governing the protein aggregation resulting in the carrier-free CatIBs. Fluorescence microscopy (Fig. 4d-f) revealed that mCherry *Tu*TreT contains mostly the native state within the CatIBs. Unfortunately, the physical size of CatIBs could not be determined due to their polydispersity.

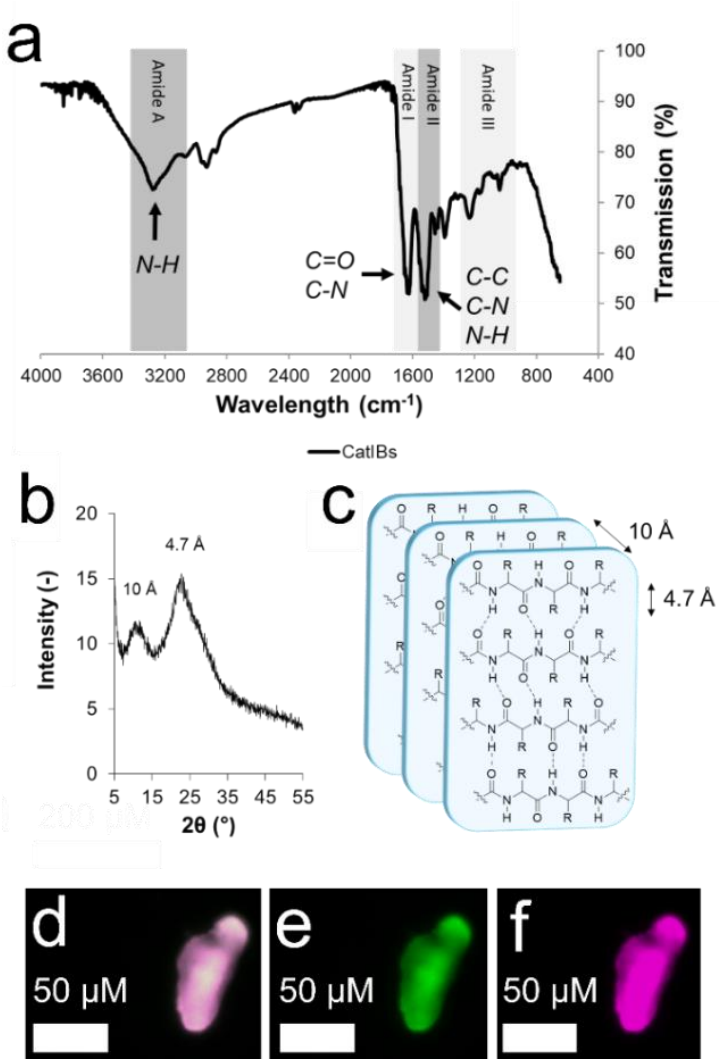


Figure 4. FTIR spectra of CatIBs of mCherry *Tu*TreT showing the distinctive amide A, I, II, and III vibrations (**a**). XRD analysis of cross-β interactions of mCherry *Tu*TreT CatIBs (**b**). The stacked β-sheets with an interstrand distance of 4.7 Å and intersheet distance of ~10 Å (R = amino acid residue) (**c**). Fluorescence microscopy of carrier-free CatIBs of mCherry *Tu*TreT showing the presence of denatured, GFP-like (green) and native (purple) mCherry protein (**d**). The GFP-like mCherry is measured with an excitation filter 488 nm - emission 522/35 nm (**e**) and for native mCherry an excitation filter of 568 - 585 nm (**f**) was used.

The thermostability of mCherry *Tu*TreT in the soluble, carrier-attached (COV-1), and carrier-free CatIBs was investigated. After 2 hours of incubation in HEPES buffer (50 mM, pH 7.0) at 60°C, 35% of enzyme activity was lost for the carrier attached enzyme. Minor loss of enzyme activity was observed for the soluble form (0%) and CatIBs (5%) at 60°C. The step-wise loss of enzyme activity from 70°C and to 80°C was similar for soluble, carrier-free, and carrier-attached mCherry *Tu*TreT. At 90°C, mCherry *Tu*TreT completely denatured. This was also evident from the loss of the purple color. We hypothesize that the large loss of activity at 60°C for the mCherry *Tu*TreT on COV-1 is due to the presence of unreacted epoxide groups. Upon heating at higher temperatures additional covalent linkages to mCherry *Tu*TreT can be formed, limiting structural mobility or causing the enzyme to denature. Enzyme deactivation when attached to epoxy functionalized carriers has been observed for multiple biocatalysts, where optimization of their stability might be achieved by blocking agents or favoring multipoint covalent attachment.[22]

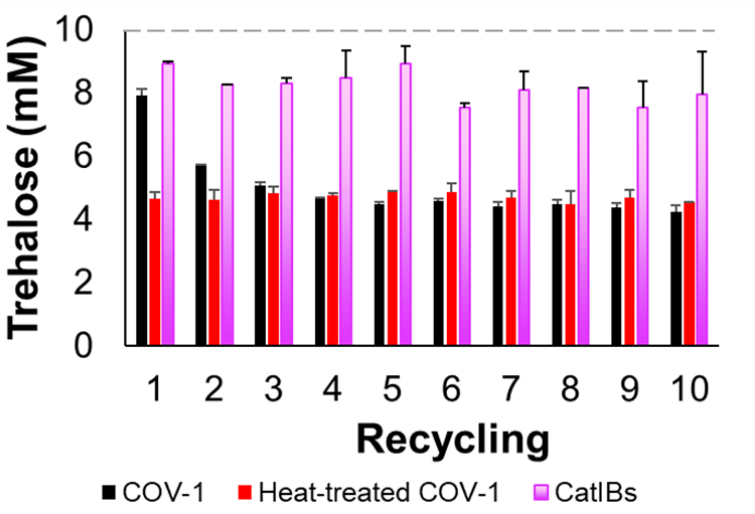


Figure 5. Recyclability of the immobilised mCherry *Tu*TreT COV-1 with or without heat-treatment (60°C), or its inclusion bodies in batch operation. The dotted grey line indicates the maximum achievable conversion. **Reaction conditions**: Glucose (10 mM), UDP-glucose (40 mM), HEPES (50 mM, pH 7.5), $MgCl_2$ (20 mM), temperature 60°C.

The recyclability of mCherry *Tu*TreT immobilized on a carrier or CatIBs is essential to maintain a high catalyst productivity in a batch process. The recyclability of enzyme on COV-1 before and after a heat-treatment was compared to CatIBs in ten consecutive cycles. Between each cycle, the biocatalyst was washed with buffer to remove the substrate and product. In line with the thermostability results, carrier-attached enzyme deactivated partially within the first 3 cycles at 60°C (Fig. 5). For the heat-treated carrier-attached mCherry *Tu*TreT, a stable performance during ten consecutive cycles was observed without leaching of any biocatalyst. The heat-treatment showed that the thermal deactivation happened only in the initial phase of the recycling due to the unreacted epoxide groups on the surface of the carrier, converging after a few cycles into similar conversion as samples that were not heat treated. The CatIBs showed no catalyst deactivation, which is consistent with the inactivation being linked to the carrier. The lower batch reproducibility of CatIBs arises from the difficulty to reproducibly resuspend the CatIBs after centrifugation. The sedimentation of the CatIBs particles typically lasted at least 5 minutes. To ensure optimal separation and no leaching of CatIBs, the solution was centrifuged at the end of each recycling step. This leads to differences in particle size distribution and therefore also a larger variation in catalytic activities, making it challenging to obtain a reproducible batch process with this procedure.

The next aspect of the comparison was the assessment of the catalytic performance of the immobilized enzymes. Both the enzyme activity during process conditions and the apparent kinetic parameters of the biocatalyst were investigated. During the biocatalytic operating process, the specific space-time yields (STY) per gram of catalyst or protein is higher for immobilized enzyme on COV-1 than for CatIBs, indicating that the carrier-immobilized mCherry *Tu*TreT demonstrates superior catalytic performance. The kinetic analysis of carrier-attached and carrier-free enzyme (table 1) gives the apparent catalytic turnover number ($k_{cat,\ app}$). This was 11-fold higher for carrier-attached enzyme than for the carrier-free CatIBs. Besides the catalytic activity, diffusion limitations due to the inaccessibility to the active site leads to a higher dissociation constant of the immobilized catalyst. Indeed, the apparent $K_{M,\ app}$ and $K_{i,\ app}$ increased with the same order of magnitude for CatIBs and COV-1 in comparison to the soluble enzyme (Table 1). The catalytic performance the carrier-attached on COV-1 is significantly better for than carrier-free CatIBs.

Table 1. Comparing the performance of the biocatalytic formulation of carrier-attached mCherry TuTreT to COV-1 or carrier-free CatIBs.

Name	Soluble	CatIBs	COV-1
Batch (nr)	1	10	10
Reaction volume (mL)	1	10 x 1 = 10	10 x 1 = 10
Reaction time (min)	30	10 x 15 = 150	10 x 15 = 150
Temperature (°C)	60	60	60
Conversion (% over nr batches)	100	82	50
Trehalose (mmol of all batches)	0.010	0.082	0.050
Trehalose (mg over number of batches)	3.42	28.1	17.11
Catalyst (mg catalyst mL^{-1})	0.50	40	5.30
$k_{cat,\ app}$ (s^{-1})	14 ± 0.36[a]	0.49 ± 0.10[b]	5.7 ± 2.0[b]
$K_{M,\ app}$ (mM)	2.3 ± 0.58[a]	10.3 ± 3.2[b]	20.1 ± 10.0[b]
$K_{I,\ app}$ (mM)	17 ± 2.2[a]	17.2 ± 5.2[b]	19.8 ± 9.7[b]
STY (g L^{-1} d^{-1})	164	135	82
STY (g L^{-1} d^{-1}) per g catalyst	n.a.[c]	3	15
STY (g L^{-1} d^{-1}) per g protein	327	130	329

[a] Ref. [15] [b] Reported $k_{cat,\ app}$, $K_{M,\ app}$, and $K_{I,\ app}$ are apparent kinetic parameters. [c] n.a.; not applicable.

The simplicity of production of the immobilized enzyme is one of the key contributors to a successful implementation in applied biocatalysis. In order to compare the amount of product produced of carrier-free or carrier-attached biocatalyst, the enzymatic productivity $g_{product}$ per amount of catalyst or bacterial culture was determined. Due to the ease of protein production as insoluble CatIBs, the enzymatic productivity expressed as $g_{product}$ per liter of culture, is higher for CatIBs. Besides the increased protein production in the form of insoluble CatIBs, simple chromatography-free down-stream processing and the abolishment of additional contaminating material (i.e. unreacted monomers of polymeric carrier materials) is beneficial for potential pharmaceutical and food-grade applications. It is important to note that we have presented a methodology to evaluate the performance of one single CatIBs formulation, while the stability and enzymatic activity of other CatIBs can vary. For instance, different stabilities and activities have been reported when biocatalysts contained a fusion domain or peptide for tailor-made CatIBs.[5, 23]

4.3 Experimental section

4.3.1 Materials

Ampicillin (Sigma-Aldrich), bovine serum albumin (ThermoFischer), DNAse I (bovine pancreas, Sigma-Aldrich), uridine 5'-diphosphate D-glucose disodium salt (Carbosynth, 98%), D-glucose (Sigma-Aldrich, 99.5%), HEPES (Sigma-Aldrich, >99.5%,), formic acid (VWR), $Mg(II)Cl_2$ hexahydrate (VWR, >99.5%), tris(hydroxymethyl)aminomethane (Tris, 99%; Sigma-Aldrich), L-arabinose (Sigma-Aldrich, ≥99%), acetonitrile (ACN) (>99.5%; Sigma-Aldrich), Immobead kit (ChiralVision), acetone (Sigma-Aldrich).

The pH was adjusted with 0.014 ΔpKa/°C for HEPES buffer. Terrific broth medium consists of 1.20% (w/w) tryptone, 2.40% (w/w) yeast extract, 53 mM K_2HPO_4, 16 mM KH_2PO_4, 4% (w/w) glycerol, and autoclaved at 121°C for 20 minutes.

4.3.2 Analytical equipment

Samples containing mCherry *Tu*TreT were analysed using an Axioplan 2 microscope (Zeiss, Mannheim, Germany), equipped with filter set XF108-2. Images were obtained using a Krypton/Argon laser using excitation 488 nm - emission 522/35 nm for denatured mCherry and excitation 568 - 585 nm long pass emission for mCherry. The projections of the individual channels were merged using the scientific image-analysis program ImageJ[24]. X-ray diffractometry (XRD) measurements were performed on a Bruker D8 Advance X-ray diffractometer using Co Ka radiation (1.78886 Å) at 35 kV and 40 mA equipped with a LynxEye detector. The data was collected from 5° to 80° 2 θ with a step size of 0.05° 2 θ and a counting time of 0.5 s per step. ATR-FTIR spectroscopy was performed with a Nicolet™ 6700 FT-IR spectrometer from Thermo Electron Corporation equipped with OMNIC Software, which were recorded at a wavenumber range from 4000 – 400 cm^{-1} (4 cm^{-1} resolution). UV-VIS spectroscopy was carried out with a Cary 60 UV-vis spectrophotometer (Agilent Technologies) connected to a Cary single cell Peltier

accessory (Agilent Technologies). A laboratory alpha 2-4 Freeze Dryer (Christ) was used for lyophilisation of CatIBs of mCherry *Tu*TreT. All reactions were performed in an Eppendorf Thermomixer. Chromatographic analysis of reaction products was performed using a Shimadzu high-performance liquid chromatography (HPLC) system equipped with an Imtakt Unison-UK amino column (0.4 by 25 cm, 60°C), an evaporative light-scattering detector (ELSD) (Shimadzu ELSD-LTII), a UV detector (SPD-20A), and acetonitrile-water-formic acid at 80:20:0.1 as the mobile phase (1 ml min^{-1}). The product formation was quantified using an external calibration curves.

4.3.3 Protein homology model of mCherry *Tu*TreT

The protein crystal homology model was constructed using 4Q7U[25] and 2XA9[26] for mCherry *Tu*TreT from the Protein Databank.[27] The surface potential was determined using the Adaptive Poisson-Boltzmann Solver plugin[28] in PyMOL Molecular Graphics system.[29]

4.3.4 Expression and purification of soluble and CatIBs of mCherry *Tu*TreT

The soluble protein and inclusion bodies of mCherry *Tu*TreT in *E. coli* Top10 pBAD/His A was expressed and purified as was described previously with minor changes[15].

(i) Preparation of cell-free extract 5 mL precultures of *E. coli* Top10 pBAD/His A containing the mCherry *Tu*TreT genes were grown in LB-medium containing 100 µg mL^{-1} ampicillin at 37 °C overnight. To a 3 L baffled Fernbach flask containing 1 L TB-medium 20 mL preculture was added and induced by addition of L-arabinose to a final concentration of 0.02% (w/w) after reaching an OD600 of 0.6 – 0.8. The cells were harvested at after 14 hours by centrifugation (24515 g, 15 min, 4°C) followed by resuspension of wet cell pellet in 4 mL lysis buffer containing Tris HCl buffer (50 mM, pH 8.0), imidazole (20 mM), lysozyme (0.5 mg mL^{-1}), DNaseI (0.1 mg mL^{-1}) per gram of wet cells. After 30 minutes of incubation on ice the cells were passed through the cell disruptor (1.35 kbar, Constant systems) for three consecutive rounds. The cell debris was collected via centrifugation 24515 g (Sorvall, Fiberlite F12-6x500 LEX, 10 min, 20°C), and the CFE was obtained *via*

decantation.

(ii) Immobilised nickel affinity chromatography The CFE was heat-treated at 60°C for 20 minutes in a water bath. The precipitates were removed by centrifugation at 24515 g (Sorvall, Fiberlite F12-6x500 LEX, 10 min, 20°C), and the heat-treated CFE was obtained via decanting. The heat-treated CFE was purified using affinity chromatography on a 1 mL Nickel Sepharose column by charging CFE on the column for at least three consecutive rounds using a peristaltic pump (Bio-Rad). The column was washed with binding buffer (20 mM Tris HCl, 500 mM NaCl, 20 mM imidazole, pH 8.0) until no protein eluted any longer. The attached mCherry *Tu*TreT was eluted using elution buffer (20 mM Tris HCl, 500 mM NaCl, 500 mM imidazole, pH 8.0) using a gradient over 10 column volumes. Protein samples were concentrated in a 12 mL Amicon Ultra Centrifugal filter (Merck, 30 kDa). Elution buffer was exchanged for HEPES (50 mM, pH 7.0) containing $MgCl_2$ (20 mM) by washing three consecutive rounds with 12 mL Amicon Ultra Centrifugal filters (Merck, 30 kDa), and analysed with SDS-PAGE and HPLC.

(iii) Purification of inclusion bodies The insoluble debris was homogenised in 20 mL Tris HCl buffer (50 mM, pH 8.5) containing 1% (w/w) deoxycholic acid (DOC). The solubilised trehalose transferase was separated from the inclusion bodies via centrifugation (20 000 × *g*, 15 min, 20°C). The resuspension and centrifugation were repeated twice. Subsequently, Tris HCl buffer (50 mM, pH 8.5) was utilised to remove remaining DOC in the cell pellet. The inclusion bodies were harvested *via* centrifugation (20 000 × *g*, 15 min, 20°C). The supernatant was decanted, resulting in the isolation of wet inclusion bodies. The wet inclusion bodies were frozen at -80°C. The purity with SDS-PAGE has been reported previously[15].

4.3.5 Lyophilization mCherry *Tu*TreT CatIBs

The frozen wet inclusion bodies of mChery *Tu*TreT (-80°C) were lyophilised (0.05 mbar, -72°C) within 12 hours. The resulting weight loss of 62% (w/w) a dry, purple powder was obtained. The lyophilised powder and wet inclusion bodies were solubilised in 2% SDS in Tris-HCl buffer (50 mM, pH 8.0) and protein

content was measured spectrophotometrically (λ_{587nm} = 0.9979 mL mg^{-1} cm^{-1}). The concentration of mCherry *Tu*TreT of wet CatIBs increased from 4% wt to 10% wt for the dry CatIBs. The protein did not denature with a GFP-like absorbance. The dry mCherry *Tu*TreT CatIBs were stored at -20°C in the dark and the activity was determined with HPLC. The 1.5 mL polypropylene Eppendorf vial containing 1.0 mL reagent mixture of D-glucose (10 mM), UDP-D-glucose (40 mM), HEPES (50 mM, pH 7.0), and $MgCl_2$ (20 mM) with either wet or freeze-dried mCherry *Tu*TreT IBs. The reaction was started and stirred at 1400 rpm at 60°C. After 15 minutes 100 µL of sample was quenched by the addition of 100 µL ice-cold HPLC-grade acetonitrile:water:formic acid (80:20:0.1) and incubated at -80°C for one hour. The samples were centrifuged at 24515 g for 10 min at 4°C. The supernatant was collected and analysed by HPLC (column: Imtakt UK-Amino 250 x 4.6 mm, 50 °C, ELSD, 80:20:0.1 acetonitrile:water:formic acid, 1.0 mL min^{-1}).

4.3.6 Protein quantification

Protein was quantified according to a method reported earlier, using the mass extinction coefficient of mCherry *Tu*TreT (ε_{587nm} = 0.9979 mg^{-1} mL cm^{-1}).[15] A protocol for the solubilisation of inclusion bodies in 2% SDS in Tris-HCl buffer (50 mM, pH 8.0) was utilised, as has been reported earlier.[15, 30]

4.3.7 Screening of Immobeads with mCherry *Tu*TreT

To 100.0 mg of carrier material 1.196 mL mCherry *Tu*TreT (5.0 mg, 4.18 mg mL^{-1} protein, 40 U) in HEPES buffer (50 mM, pH 7.0) was added and incubated overnight (4°C, 10 rpm, NeoLab rotator). The supernatant was transferred and residual protein content was measured spectrophotometrically (λ_{587nm} = 0.9979 mL mg^{-1} cm^{-1}). The immobilised enzymes were filtered and washed with 1 mL ice-cold MiliQ water. The immobilised enzymes were washed with ice-cold acetone, filtered, and dried with air. The activity of the immobilised enzymes was determined using a HPLC-based activity assay. A 1.5 mL polypropylene Eppendorf tube containing 1.0 mL reagent mixture of D-glucose (10 mM), UDP-D-glucose (40 mM), HEPES (50 mM, pH 7.0), and $MgCl_2$ (20 mM) with 5 mg carrier material containing immobilised

mCherry *Tu*TreT (max 2 U of soluble enzyme activity immobilised, maximum 0.25 mg of soluble protein) and stirred at 1400 rpm at 60 °C. After one hour of reaction time, 100 µL samples were quenched in by the addition of 100 µL of reaction solution to an equal volume of ice-cold HPLC-grade acetonitrile:water:formic acid (80:20:0.1). The samples were centrifuged at 14 000 rpm for 10 min at 4 °C. The supernatant was collected and analysed by HPLC (column: Imtakt UK-Amino 250 x 4.6 mm, 50 °C, ELSD, 80:20:0.1 acetonitrile:water:formic acid, 1.0 mL min^{-1}). Enzyme activity was calculated with external standards for trehalose using the slope of at least three different substrate concentrations. The enzyme activity was determined in duplicate. One unit (U) is defined as the conversion of 1 µmol of D-glucose per minute.

4.3.8 Rate and surface coverage of mCherry *Tu*TreT on COV-1

To a solution containing 1.0 mg mL^{-1} mCherry *Tu*TreT, $MgCl_2$ (20 mM), HEPES (50 mM, pH 7.0) were added 0, 1, 2, 5, 7.5, 10, 15, 20, and 25 mg of COV-1 material in a polystyrene cuvette (pathlength 1 cm) and shaken at 4°C at 1000 rpm. The absorbance of mCherry *Tu*TreT (ε_{587nm}) were measured within a time-course of 22 hours. The surface coverage was measured by the evaluation of unreacted amino groups of a fixed amount of enzyme (1.07 mg, 0.71 µmol amino groups) and a varying amount of epoxide groups (1.0 – 25.0 mg COV-1, 70 nmol epoxides mg^{-1} carrier) after 22 hours of reaction time.

4.3.9 Temperature stability of immobilized mCherry *Tu*TreT and CatIBs

The 1.5 mL polypropylene Eppendorf vial containing 1.0 mL containing: (i) soluble mCherry *Tu*TreT (0.02 mg mL^{-1}), HEPES (50 mM, pH 7.0), $MgCl_2$ (20 mM); (ii) CatIBs mCherry *Tu*TreT (40.0 mg wet CatIBs, 1.04 mg mCherry *Tu*TreT protein), HEPES (50 mM, pH 7.0), $MgCl_2$ (20 mM); (iii) COV-1 mCherry *Tu*TreT (5 mg mCherry *Tu*TreT COV-1, 0.25 mg protein mCherry *Tu*TreT), HEPES (50 mM, pH 7.0), $MgCl_2$ (20 mM) were incubated at 60 – 90 °C, 800 rpm, 2 hours of incubation time. After this, the enzyme activity was assayed for UDP-D-glucose (40 mM), D-glucose (10 mM), HEPES (50 mM, pH 7.0), $MgCl_2$ (20 mM), 800 rpm, 60

°C. The reaction was started by the addition of the biocatalyst and stirred at 1400 rpm at 60°C. Samples were quenched by the addition of 100 µL ice-cold HPLC-grade acetonitrile:water:formic acid (80:20:0.1) between 0 to 35 minutes and incubated at -80°C for one hour. The samples were centrifuged at 24515 g for 10 min at 4°C. The supernatant was collected and analysed by HPLC (column: Imtakt UK-Amino 250 x 4.6 mm, 50 °C, ELSD, 80:20:0.1 acetonitrile:water:formic acid, 1.0 mL min^{-1}). Enzyme activity was calculated with external standards for trehalose using the slope of at least three different substrate concentrations. The enzyme activity was determined in duplicate.

4.3.10 Recyclability of CatIBs and immobilised mCherry *Tu*TreT

The 1.5 mL polypropylene Eppendorf vial containing 1.0 mL reagent mixture of D-glucose (10 mM), UDP-D-glucose (40 mM), HEPES (50 mM, pH 7.0), and $MgCl_2$ (20 mM) with either wet 40.0 mg of mCherry *Tu*TreT IBs (1.04 mg protein), 5.3 mg COV-1 mCherry *Tu*TreT (0.44 U, 0.25 mg protein), or 5.3 mg COV-1 mCherry *Tu*TreT (0.44 U, 0.25 mg protein) which was heat-treated for 2 hours (50 mM HEPES, 20 mM $MgCl_2$, pH 7.0, 60 °C) before use. The reaction was started and stirred at 1400 rpm at 60°C. After 15 minutes 100 µL of sample was quenched by the addition of 100 µL ice-cold HPLC-grade acetonitrile:water:formic acid (80:20:0.1) and incubated at -80°C for one hour. The samples were centrifuged at 24515 g for 10 min at 4°C. The supernatant was collected and analysed by HPLC (column: Imtakt UK-Amino 250 x 4.6 mm, 50 °C, ELSD, 80:20:0.1 acetonitrile:water:formic acid, 1.0 mL min^{-1}). The IBs or COV-1 of mCherry *Tu*TreT were centrifuged (30 s, 24515 g, 4°C), washed with 1.0 mL HEPES (50 mM, pH 7.0) containing $MgCl_2$ (20 mM), and centrifuged (30 s, 24515 g, 4°C). Again, the IBs or COV-1 of mCherry *Tu*TreT were centrifuged (30 s, 24515 g, 4°C), washed with 1.0 mL HEPES (50 mM, pH 7.0) containing $MgCl_2$ (20 mM), and centrifuged (30 s, 24515 g, 4°C). The reaction was started again by the addition of the reagent mixture of D-glucose (10 mM), UDP-D-glucose (40 mM), HEPES (50 mM, pH 7.0), and $MgCl_2$ (20 mM) with a final reaction volume of 1.0 mL. The reaction-wash-reaction step was repeated until a total of ten cycles were performed.

4.3.11 Apparent enzyme kinetics of CatIBs and immobilised mCherry *Tu*TreT

The 1.5 mL polypropylene Eppendorf vial containing 1.0 mL reagent mixture of D-glucose (10 mM), UDP-D-glucose (40 mM), HEPES (50 mM, pH 7.0), and $MgCl_2$ (20 mM) with either 20.0 mg lyophilised mCherry *Tu*TreT IBs (0.39 U, 2 mg) or 5.0 mg COV-1 mCherry *Tu*TreT (0.44 U, 0.25 mg protein). The reaction was started and stirred at 1400 rpm at 60°C. Samples were quenched by the addition of 100 µL ice-cold HPLC-grade acetonitrile:water:formic acid (80:20:0.1) between 0 to 35 minutes and incubated at -80°C for one hour. The samples were centrifuged at 24515 g for 10 min at 4°C. The supernatant was collected and analysed by HPLC (column: Imtakt UK-Amino 250 x 4.6 mm, 50 °C, ELSD, 80:20:0.1 acetonitrile:water:formic acid, 1.0 mL min^{-1}). Enzyme activity was calculated with external standards for trehalose using the slope of at least three different substrate concentrations. The enzyme activity was determined in duplicate. The data was fitted (Gnuplot 5.2).

4.3.12 Quantification of D-trehalose with HPLC

Samples during activity assays were quenched by the addition of 100 µL of reaction solution to an equal volume of ice-cold HPLC-grade acetonitrile and incubated at -80°C for one hour. The samples were centrifuged at 24515 g for 10 min at 4°C. The supernatant was collected and analysed by HPLC (column: Imtakt UK-Amino 250 x 4.6 mm, 50 °C, ELSD, 80:20:0.1 acetonitrile:water:formic acid, 1.0 mL min^{-1}). Enzyme activity was calculated with external standards for trehalose using the slope of at least three different substrate concentrations. The enzyme activity was determined in duplicate.

References

[1] A. Basso, S. Serban, *Molecular Catalysis* **2019**, *479*, 110607.

[2] a) U. Hanefeld, L. Cao, E. Magner, *Chemical Society Reviews* **2013**, *42*, 6211-6212; b) G. Torrelo, N. van Midden, R. Stloukal, U. Hanefeld, *ChemCatChem* **2014**, *6*, 1096-1102; c) L. Mestrom, M. Przypis, D. Kowalczykiewicz, A. Pollender, A. Kumpf, S. R. Marsden, I. Bento, A. B. Jarzębski, K. Szymańska, A. Chruściel, D. Tischler, R. Schoevaart, U. Hanefeld, P.-L. Hagedoorn, *Internation Journal of Molecular Sciences* **2019**, *20*, 5263.

[3] L. Cao, L. v. Langen, R. A. Sheldon, *Current Opinion in Biotechnology* **2003**, *14*, 387-394.

[4] a) U. Krauss, V. D. Jäger, M. Diener, M. Pohl, K.-E. Jaeger, *Journal of Biotechnology* **2017**, *258*, 136-147; b) B. Nidetzky, in *Protein Aggregation in Bacteria*, **2014**, pp. 247-261.

[5] R. Kloss, T. Karmainski, V. D. Jäger, D. Hahn, A. Grünberger, M. Baumgart, U. Krauss, K.-E. Jaeger, W. Wiechert, M. Pohl, *Catalysis Science & Technology* **2018**, *8*, 5816-5826.

[6] a) S. Li, K. Lin, H. Pang, Y. Wu, J. Xu, *Applied Biochemistry and Biotechnology* **2013**, *169*, 612-623; b) Q. Dong, X. Yan, M. Zheng, Z. Yang, *Journal of Bioscience and Bioengineering* **2014**, *117*, 706-710; c) D. M. Worrall, N. H. Goss, *Australian Journal of Biotechnology* **1989**, *3*, 28-32; d) K. Tokatlidis, P. Dhurjati, J. Millet, P. Béguin, J.-P. Aubert, *FEBS Letters* **1991**, *282*, 205-208.

[7] a) E. García-Fruitós, N. González-Montalbán, M. Morell, A. Vera, R. M. Ferraz, A. Arís, S. Ventura, A. Villaverde, *Microbial Cell Factories* **2005**, *4*, 27; b) S.-Y. Park, S.-H. Park, S.-K. Choi, *Analytical Biochemistry* **2012**, *426*, 63-65.

[8] R. Kloss, M. H. Limberg, U. Mackfeld, D. Hahn, A. Grünberger, V. D. Jäger, U. Krauss, M. Oldiges, M. Pohl, *Scientific Reports UK* **2018**, *8*, 5856.

[9] a) A. S. Korovashkina, A. N. Rymko, S. V. Kvach, A. I. Zinchenko, *Journal of Biotechnology* **2013**, *164*, 276-280; b) J. Nahálka, A. Vikartovská, E. Hrabárová, *Journal of Biotechnology* **2008**, *134*, 146-153.

[10] L. Wang, S. K. Maji, M. R. Sawaya, D. Eisenberg, R. Riek, *PLOS Biology* **2008**, *6*, e195.

[11] J. Nahálka, A. Vikartovská, E. Hrabárová, *Journal of Biotechnology* **2008**, *134*, 146-153.

[12] a) T. Jesionowski, J. Zdarta, B. Krajewska, *Adsorption* **2014**, *20*, 801-821; b) M. Hoarau, S. Badieyan, E. N. G. Marsh, *Organic & Biomolecular Chemistry* **2017**, *15*, 9539-9551.

[13] a) R. C. Rodrigues, C. Ortiz, Á. Berenguer-Murcia, R. Torres, R. Fernández-Lafuente, *Chemical Society Reviews* **2013**, *42*, 6290-6307; b) J. C. S. d. Santos, O. Barbosa, C. Ortiz, A. Berenguer-Murcia, R. C. Rodrigues, R. Fernandez-Lafuente, *ChemCatChem* **2015**, *7*, 2413-2432; c) C. Garcia-Galan, Á. Berenguer-Murcia, R. Fernandez-Lafuente, R. C. Rodrigues, *Advanced Synthesis & Catalysis* **2011**, *353*, 2885-2904.

[14] a) C. Mateo, J. M. Palomo, G. Fernandez-Lorente, J. M. Guisan, R. Fernandez-Lafuente, *Enzyme and Microbial Technology* **2007**, *40*, 1451-1463; b) C. Mateo, R. Monti, B. C. C. Pessela, M. Fuentes, R. Torres, J. Manuel Guisán, R. Fernández-Lafuente, *Biotechnology Progress* **2004**, *20*, 1259-1262; c) B. C. C. Pessela, C. Mateo, M. Fuentes, A. Vian, J. L.

García, A. V. Carrascosa, J. M. Guisán, R. Fernández-Lafuente, *Enzyme and Microbial Technology* **2003**, *33*, 199-205.

[15] a) L. Mestrom, S. R. Marsden, M. Dieters, P. Achterberg, L. Stolk, I. Bento, U. Hanefeld, P.-L. Hagedoorn, *Applied and Environmental Microbiology* **2019**, *85*, e00942-00919; b) L. Mestrom, S. R. Marsden, M. Dieters, P. Achterberg, L. Stolk, I. Bento, U. Hanefeld, P.-L. Hagedoorn, *Applied and Environmental Microbiology* **2019**, *85*, e03084-03018.

[16] B. L. Urbanek, D. C. Wing, K. S. Haislop, C. J. Hamel, R. Kalscheuer, P. J. Woodruff, B. M. Swarts, *ChemBioChem* **2014**, *15*, 2066-2070.

[17] T. Kouril, M. Zaparty, J. Marrero, H. Brinkmann, B. Siebers, *Archives of Microbiology* **2008**, *190*, 355.

[18] a) E. García-Fruitós, N. González-Montalbán, M. Morell, A. Vera, R. M. Ferraz, A. Arís, S. Ventura, A. Villaverde, *Microbial Cell Factories* **2005**, *4*, 27; b) L. A. Gross, G. S. Baird, R. C. Hoffman, K. K. Baldridge, R. Y. Tsien, *PNAS* **2000**, *97*, 11990-11995; c) B. J. Bevis, B. S. Glick, *Nature Biotechnology* **2002**, *20*, 83-87; d) I. I. Shemiakina, G. V. Ermakova, P. J. Cranfill, M. A. Baird, R. A. Evans, E. A. Souslova, D. B. Staroverov, A. Y. Gorokhovatsky, E. V. Putintseva, T. V. Gorodnicheva, T. V. Chepurnykh, L. Strukova, S. Lukyanov, A. G. Zaraisky, M. W. Davidson, D. M. Chudakov, D. Shcherbo, *Nature Communications.* **2012**, *3*, 1204.

[19] U. Hanefeld, L. Gardossi, E. Magner, *Chemical Society Reviews* **2009**, *38*, 453-468.

[20] J. Andrade, C. G. Pereira, J. C. d. Almeida Junior, C. C. R. Viana, L. N. d. O. Neves, P. H. F. d. Silva, M. J. V. Bell, V. d. C. d. Anjos, *LWT - Food Science and Technology* **2019**, *99*, 166-172.

[21] R. Kloss, M. H. Limberg, U. Mackfeld, D. Hahn, A. Grünberger, V. D. Jäger, U. Krauss, M. Oldiges, M. Pohl, *Scientific Reports* **2018**, *8*, 5856.

[22] C. Mateo, O. Abian, R. Fernandez–Lafuente, J. M. Guisan, *Enzyme Microbial Technology* **2000**, *26*, 509-515.

[23] a) J.-P. Arié, M. Miot, N. Sassoon, J.-M. Betton, *Molecular Microbiology* **2006**, *62*, 427-437; b) M. Diener, B. Kopka, M. Pohl, K.-E. Jaeger, U. Krauss, *ChemCatChem* **2016**, *8*, 142-152.

[24] C. A. Schneider, W. S. Rasband, K. W. Eliceiri, *Nature Methods* **2012**, *9*, 671.

[25] S. Pletnev, D. M. Shcherbakova, O. M. Subach, N. V. Pletneva, V. N. Malashkevich, S. C. Almo, Z. Dauter, V. V. Verkhusha, *PLOS ONE* **2014**, *9*, e99136.

[26] E.-J. Woo, S.-I. Ryu, H.-N. Song, T.-Y. Jung, S.-M. Yeon, H.-A. Lee, B. C. Park, K.-H. Park, S.-B. Lee, *Journal of Molecular Biology* **2010**, *404*, 247-259.

[27] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Research* **2000**, *28*, 235-242.

[28] E. Jurrus, D. Engel, K. Star, K. Monson, J. Brandi, L. E. Felberg, D. H. Brookes, L. Wilson, J. Chen, K. Liles, M. Chun, P. Li, D. W. Gohara, T. Dolinsky, R. Konecny, D. R. Koes, J. E.

Nielsen, T. Head-Gordon, W. Geng, R. Krasny, G.-W. Wei, M. J. Holst, J. A. McCammon, N. A. Baker, *Protein Science* **2018**, *27*, 112-128.

[29] W. L. DeLano, *CCP4 Newsletter On Protein Crystallography* **2002**, *40*, 82-92.

[30] A. Singh, V. Upadhyay, A. K. Panda, in *Insoluble Proteins: Methods and Protocols* (Ed.: E. García-Fruitós), Springer New York, New York, NY, **2015**, pp. 283-291.

5

Anomeric selectivity of Trehalose Transferase with Rare L-Sugars

This chapter is based on:

Luuk Mestrom, Stefan R. Marsden, Hessel van der Eijk, Jesper U. Laustsen, Cy M. Jeffries, Dmitri I. Svergun, Peter-Leon Hagedoorn, Isabel Bento*, Ulf Hanefeld*, *accepted, ACS Catal.* **2020**, 10, 8835–8839. DOI: 10.1021/acscatal.0c02117

Purple rain, purple rain
I only want to see you
only want to see you
in the purple rain
Prince and the Revolution, Purple Rain, 1984

5.1 Introduction

The synthesis of a glycosidic bond is one of the most important reactions within glycochemistry. Enzymes couple sugars to afford oligosaccharides with high selectivity under mild reaction conditions. As enzymes are regarded to have evolved toward the selective conversion of the naturally more abundant D-sugars, their L-sugar enantiomers are often not considered as suitable substrates for enzymatic conversions. For this reason, the coupling of both D- and L-glycopyranose acceptors has rarely been compared for a single enzyme. However, the incorporation of L-sugars offers a broad spectrum of diametrically opposed glycosides or oligosaccharides, which might display new biological activities.

In one example, a retaining non-LeLoir glycosyltransferase (GT) coupled L-glycopyranose acceptors with sucrose as sugar donor in an a,β-(1→2)-fashion, while α,α-(1→2)-glycosidic bonds were observed with D-glycopyranose acceptors.[1] The switch of anomeric selectivity for the sugar acceptor was attributed to the $^4C_1$ and $^1C_4$ chair configuration for D- and L-glycopyranoses, which affect the position of the nucleophilic hydroxyl group at the anomeric position. In general, α-D and β-L anomers of the same sugar are structurally more alike (Fig. 1a), than the corresponding α-D and α-L anomers (Fig. 1b).[2] This structural similarity allows the conversion of both α-D and β-L configured substrates by an (*S*)-selective enzyme.

For retaining glycosyltransferases (GTs) with an internal nucleophilic substitution (S_Ni) mechanism the anomeric selectivity can be expected to invert when (*S*)-selectivity is retained (Fig. 1c,d).[3] The “same-face” attack of the nucleophile (i.e. sugar acceptor) on the leaving group (i.e. sugar donor) is guided by hydrogen bonding, and proceeds with high anomeric selectivity for the sugar donor and acceptor.[3b] The position of the anomeric hydroxyl of the sugar acceptor might affect the type of glycosidic bond formed.

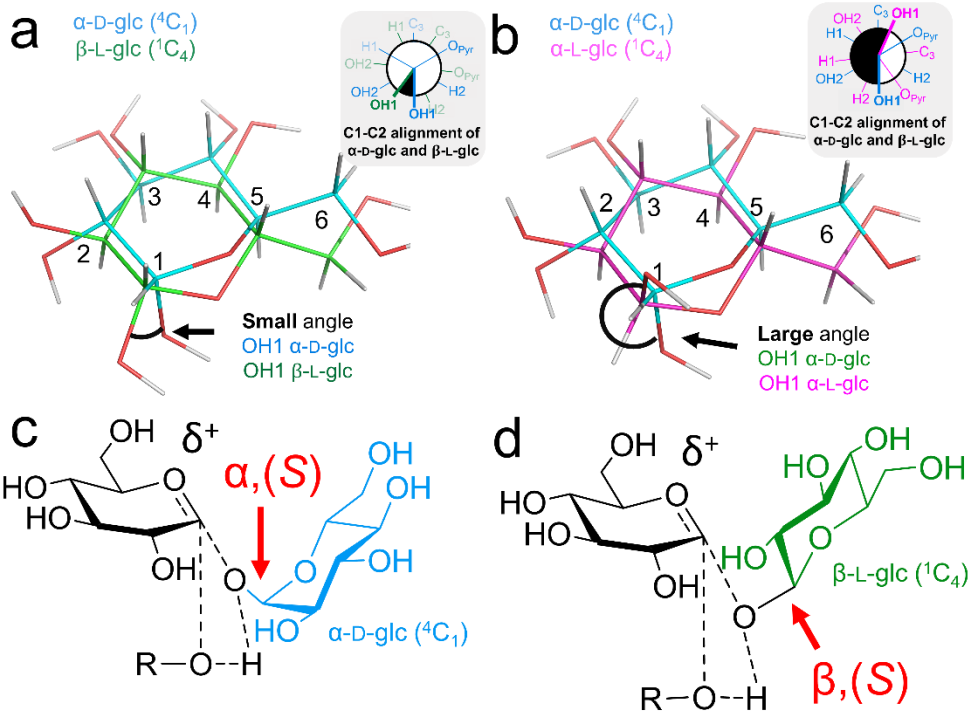


Figure 1: The position of the anomeric hydroxyl of α-D-glucopyranose ($^4C_1$, cyan) when superimposed to β-L-glucopyranose ($^1C_4$, green) in **a** is more similar than α-L-glucopyranose ($^1C_4$, purple) in **b**. The inset in **a, b** shows the representative overlay of the Newman projection of the anomeric OH1. The S_Ni-like reaction mechanism with an oxocarbenium transition state ($^4H_3$) that allows the approach of the anomeric hydroxyl, α-OH of D-glucose ($^4C_1$) in **c**, or the β-OH of L-glucose ($^1C_4$) in **d**, is guided by hydrogen bonding from the same face as the NDP leaving group. R = NDP.

Here, the anomeric selectivity of the retaining LeLoir GT trehalose transferase (TreT) was investigated. TreT is particularly suitable for the screening of L-glycopyranoses, as it couples nucleotide diphosphate (NDP) sugar donors to a wide spectrum of non-phosphorylated D-sugar acceptors, resulting in an α,α-(1→1)-glycosidic linkage.[3c] We focused on the recently described TreT from *Thermoproteus uzoniensis* (*Tu*TreT) fused to mCherry for the systematic screening

of D- and L-glycopyranoses as sugar acceptors.[4] mCherry *Tu*TreT is an interesting enzyme due to a high thermostability, activity, the possibility of fluorometric detection due to, mCherry and performance as an immobilized catalyst.[5]

5.2 Results and discussion

Initially, the reaction conditions were optimized to exclude any possible side reactions or promiscuous activities. *Tu*TreT did not display any phosphorylase or hydrolase activity. The use of glucose-1-phosphate as sugar donor did not result in the formation of trehalose (excluding phosphorylase activity), and no hydrolase activity was observed when the enzyme was incubated solely with trehalose. However, slow hydrolysis of UDP-glucose to UDP and glucose by *Tu*TreT was observed, resulting in the subsequent formation of trehalose from glucose and UDP-glucose. To minimize the undesired formation of trehalose as a side product *via* UDP glucose hydrolysis during the screening of other sugar acceptors, the reaction time was limited to 60 minutes using 1.0 mg mL^{-1} of *Tu*TreT.[6]

Using these optimized conditions, the substrate tolerance of *Tu*TreT was probed in a HPLC based screening of D- and L-sugars (Fig. 2). Conversion of L-glycopyranoses resulted in the hypothesized β-selectivity for *Tu*TreT. Successful enzymatic conversions were repeated on preparative scale and the obtained trehalose analogues were analyzed by NMR and HR-MS. D-glucose, D-mannose, D-galactose, D-xylose exclusively led to the formation of α,α-(1→1)-linked trehalose derivatives, while L-glucose, L-galactose, and L-gulose led to the formation of α,β-(1→1)-linked trehalose derivatives. The long-range C-H coupling over the glycosidic linkage confirmed the direct coupling of the $C1_{acceptor}$ with the $H1'_{donor}$ and *vice versa* in gHMBC experiments. The $^4C_1$ configuration of α-D-α-D-glycopyranosides was confirmed by $J_{1,2}$ coupling (~4 Hz) of the anomeric protons, which are *gauche* configured. The anomeric protons of β-L-glycopyranosides with a $^1C_4$ chair conformation are *anti* configured, resulting in larger $J_{1,2}$ coupling constants (~8 Hz).[1]

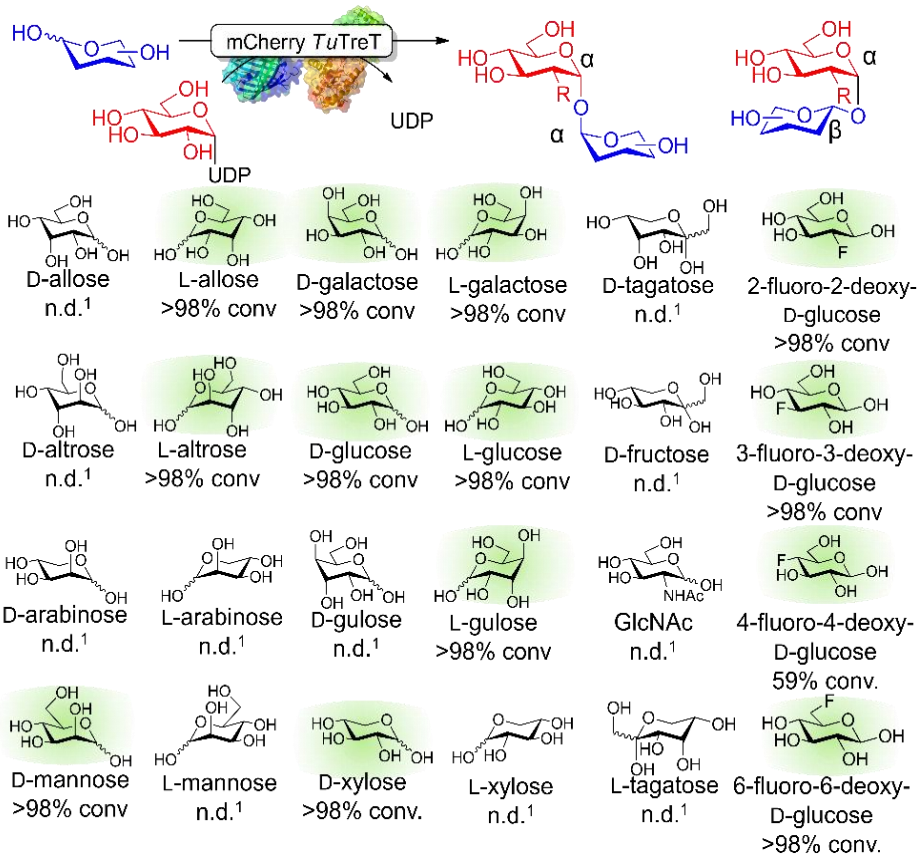


Figure 2: mCherry *Tu*TreT catalyzed conversion of D- and L-sugar acceptors with UDP-D-glucose as donor for the screening for the formation of product. The conversion was determined using external calibration curves of sugar acceptor with HPLC. [1]n.d: not detected including a trehalose analogue product. **Reaction conditions**: substrate (10 mM), UDP-D-glucose (40 mM), HEPES (50 mM), $MgCl_2$ (20 mM), mCherry *Tu*TreT (13.5 nM), pH 7.0, 60 °C, 1 hour of reaction time.

Further analysis of the HPLC screening demonstrates, that D- and L-enantiomers of glucose and galactose were accepted, but L-mannose was not. For L-gulose, L-allose, and L-altrose their D-glycopyranoses enantiomers were not accepted. Anomeric selectivity is dictated by more than the anomeric configuration, and the overall conformation of the sugar acceptor is important as well. The

structural variants of D-glycopyranoses with a $^4C_1$ configuration were readily converted, such as D-glucose, D-mannose, and D-galactose. This is in line with results for other TreTs.[7] C5 sugars were generally less well accepted, with the exception of D-xylose, which lacks a CH_2OH group in comparison to D-glucose. The ketohexopyranoses D-fructose, and D- and L-tagatose were not converted under the conditions provided, which display dissimilar overall structural conformation of the cyclic ring structure as well as the anomeric configuration. Interactions between active site residues and the carbohydrate substrate were investigated by extending the substrate screening to fluorodeoxy-carbohydrates. Unlike hydroxyl groups, fluorine can exclusively function as a hydrogen bond acceptor. All fluoro-deoxy-D-glucopyranoses were quantitatively converted as acceptor substrates, with the exception of 4-deoxy-4-fluoro-D-glucose.[7a] The interaction of hydrogen bond donor 4-OH of the sugar acceptor with the deprotonated Asp254 of *Tu*TreT is possibly important for acceptor substrate recognition

The substrate tolerance toward the glycopyranose moiety of the sugar donor and the sugar acceptor is distinct from one another.[8] For instance, the coupling of UDP-D-glucose with *N*-acetyl-D-glucosamine (GlcNAc) did not result in any observable conversion. The sugar donor UDP-D-GlcNAc and D-glucose is readily converted by *Tu*TreT with >98% conversion and α,α-(1→1)-selectivity.

As this is the first report that the α-D-selective retaining glycosyltransferase TreT catalyzes the glycosidic bond formation with β-L-glycopyranose acceptors, the protein crystal structure of TuTreT was studied. As the mCherry *Tu*TreT fusion construct did not crystallize satisfactory, the glycine-rich linker of the fusion protein was cleaved using "*in situ"* proteolysis with retention of enzyme activity (Fig. 8), and the protein was purified (Fig. 9). The protein without the mCherry tag subsequently crystallized as *apo* (PDB: 6ZJ4, 2.1Å resolution), co-crystallized with magnesium(II) (PDB: 6ZJ7, 2.15Å resolution), or soaked with D-trehalose (PDB: 6ZJH, 2.1Å resolution), D-glucopyranosyl-L-galactopyranose (PDB: 6ZN1, 1.75Å resolution), and UDP-α-D-glucose (PDB: 6ZMZ, 1.9Å resolution). The latter three are shown in Fig. 3a-c.

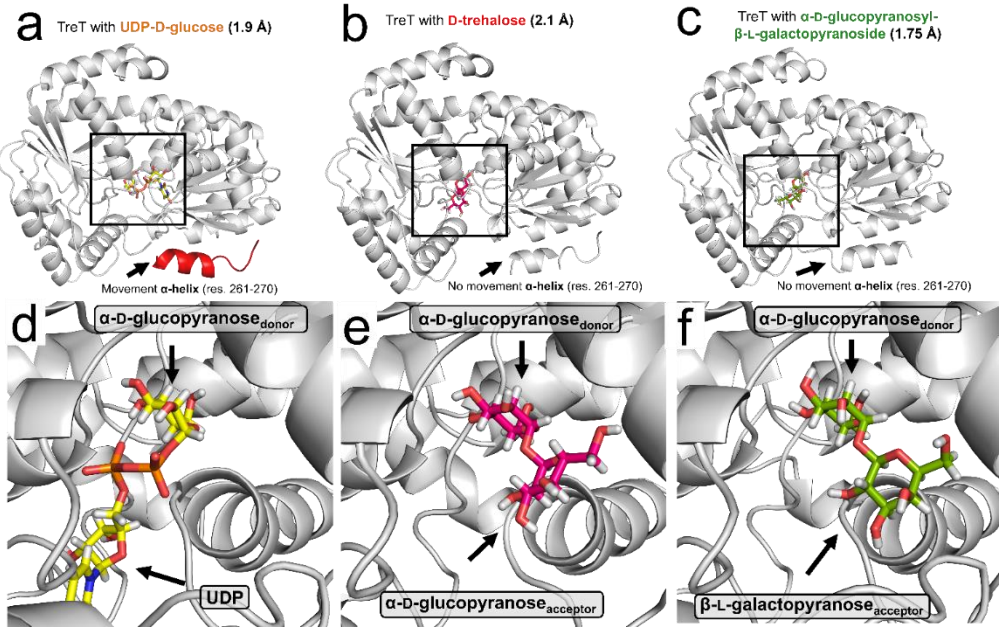


Figure 3: Protein crystal structure of *Tu*TreT containing UDP-D-glucose in **a** (PDB: 6ZMZ, inset in **d**), *Tu*TreT bound with α,α-trehalose in **b** (PDB: 6ZJH, inset shown in **e**), and *Tu*TreT in complex with enzymatically synthesized α-D-glucopyranosyl-β-L-galactopyranose in **c** (PDB: 6ZN1, inset shown in **f**). The nature of the glycosidic bond is demonstrated showing a high overlap for the sugar donor and the glycosidic linkages between **d** and **e**, or **d** and **f**, while the orientation of the sugar acceptor changes slightly between **e** and **f**.

The overall three-dimensional fold observed in all determined crystal structures are similar to the one found in trehalose phosphate synthase (OtsA, PDB: 1GZ5, RMSD of 2.0 Å for 304 Cα)[9] from *E. coli* and in trehalose transferase from *Pyrococcus horikoshii* (PDB: 2X6Q, RMSD of 2.5Å for 363 Cα),[10] showing in each domain a characteristic Rossmann fold (Fig. 4).

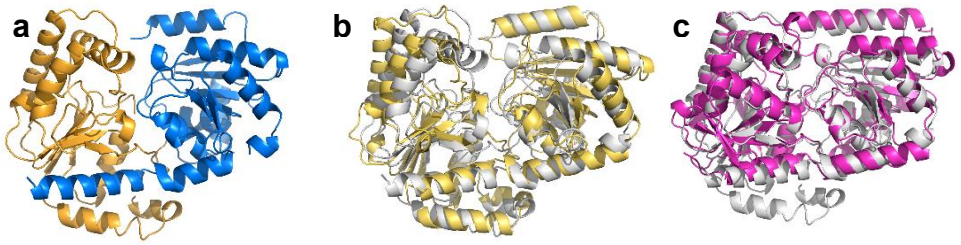


Figure 4: *Tu*TreT monomer with the *N*-terminal (sugar acceptor) domain colored in orange and the *C*-terminal (donor) domain coloured in blue in ***a***. Structural alignement between *Tu*TreT (in gray) and *Ph*TreT (in yellow; rmsd of 2.5Å for 363 Cα) *in* ***b****,* and structural alignement between TuTreT (in gray) and OtsA from *E. coli* (in magenta; rmsd of 2.0 Å for 315 Cα) *in* ***c****.*

Furthermore, all *Tu*TreT structures show a monomer in the asymmetric unit, and this state was confirmed by SEC-SAXS measurements in aqueous solution.[11] We conclude that the functional unit of *Tu*TreT is a monomer, whereas a dimer has been described for *Ph*TreT[10] and a tetramer for OtsA from *E. coli*.[9] The overall protein conformation remained unchanged when bound to ligands *in aqua*, according to SEC-SAXS. The protein crystal structures also show high structural similarity, however a minor conformational change was observed for the sugar donor binding site of *Tu*TreT when it was soaked with UDP-D-glucose. Hydrogen bond interactions between the uracil moiety and a disordered loop region of the protein (Fig. 3), were inducing a shift of an α-helix by 2.0 Å (Fig. 5). This finding is not in agreement for what has been found for *Ph*TreT, where a larger conformational change was observed for the whole domain after soaking the crystals with trehalose.[10]

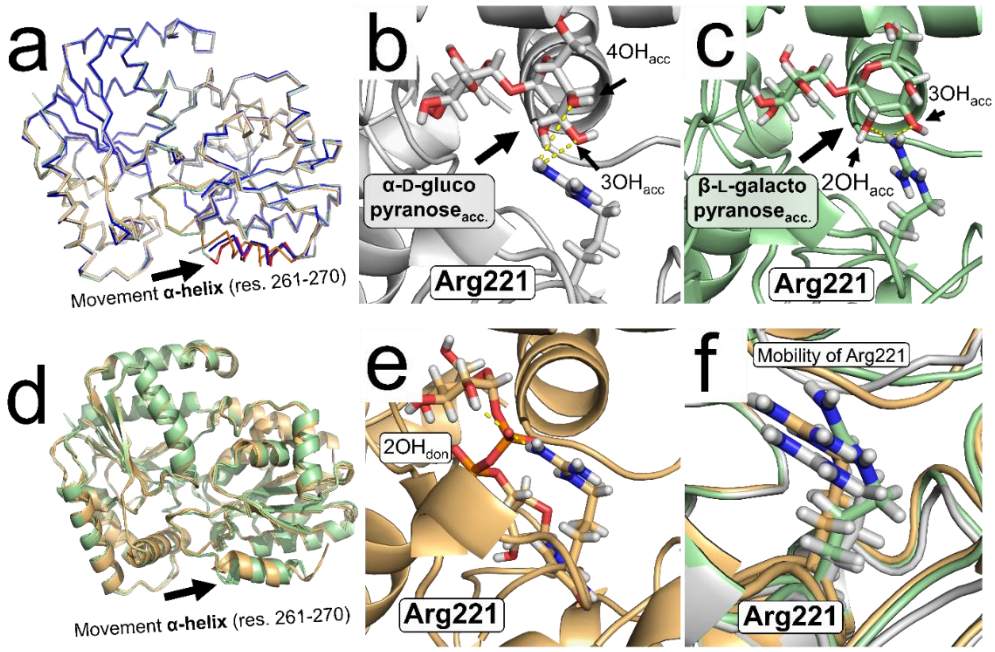


Fig. 5: Comparison of the protein crystal structure containing the movement of the α-helix (res 261-270) of *Tu*TreT in complex with UDP-D-glucose (orange, PDB: 6ZMZ), D-trehalose (white, PDB: 6ZJH), and α-D-glucopyranose-β-L-galactopyranose (green, PDB: 6ZN1) in **a** (ribbon) and **d** (cartoon). The movement of Arg221 with *Tu*TreT is shown in complex D-trehalose (white, PDB: 6ZJH) in **b**, α-D-glucopyranose-β-L-galactopyranose (green, PDB: 6ZN1) in **c**, and UDP-D-glucose (orange, PDB: 6ZMZ) in **e**, and overlay of the mobility of Arg221.

The active site of *Tu*TreT is located between the *N*- or *C*-terminal domains of the the acceptor and donor binding sites (Fig. 4). Substrate-bound structures demonstrated clear electron densities at the active site. The active site residues for the sugar donor and acceptor binding sites are conserved for *Tu*TreT, *Ph*TreT,[10] and OtsA (Fig. 6).[9] In *Tu*TreT, the pyrophosphate moiety of the nucleotide sugar donor interacts with Arg221 and Lys226, and the active site can accommodate pyrimidine or purine nucleobases (Fig. 3a,d). This leads to the ability of the enzyme to convert nucleotide sugar donors with different nucleotides, which holds for TreTs in general,[12] as was observed with UDP- and ADP-D-glucose with *Tu*TreT previously.[3,4]

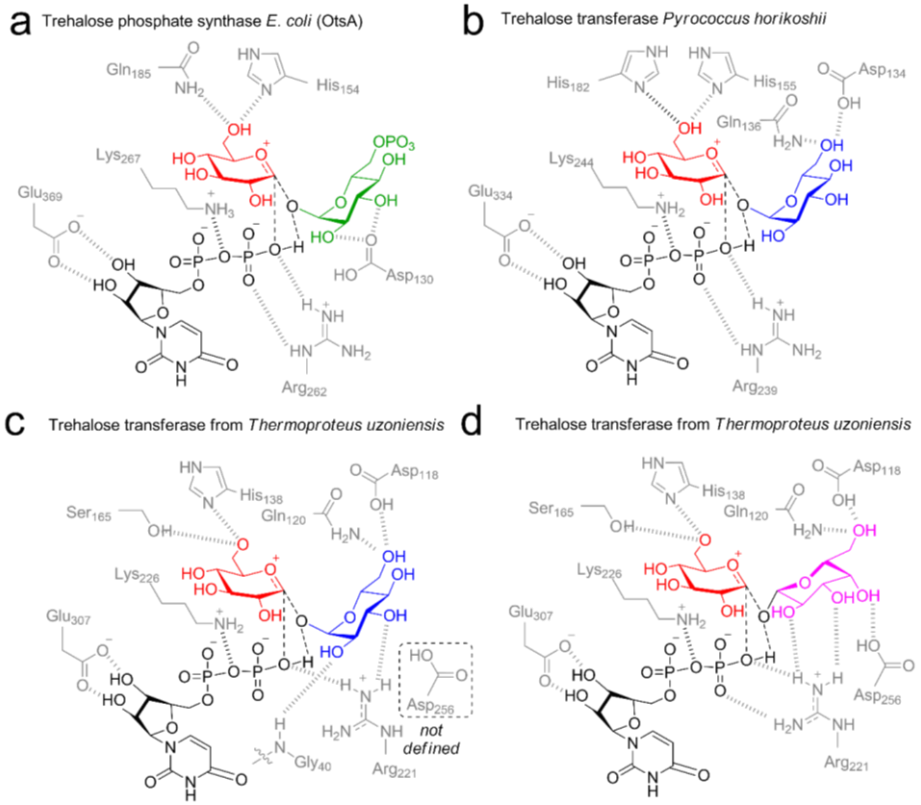


Figure 6: Similairities in amino acid active site residues demonstrating the $^4H_3$ chair conformation of UDP-D-glucose and the glucopyranose acceptor. For trehalose phosphate synthase in *Escherichia coli* (OtsA, PDB: 1GZ5) the sugar acceptor is α-D-glucose-6-phosphate (green) in **a**, trehalose transferase from *Pyrococcus horikoshii* (*Ph*TreT[10]) with α-D-glucose (blue) and **b**, and trehalose transferase from *Thermoproteus uzoniensis* (PDB: 6ZJH) with α-D-glucose (blue) in **c** or β-L-galactose (purple) in **d** (PDB: 6ZN1).

The natural product α,α-D-trehalose shows an α,α-(1→1)-glycosidic bond when bound to *Tu*TreT Fig. (3b,e). In trehalose, the α-D-glucopyranose$_{donor}$ moiety binds at the same sugar binding site as the UDP-α-D-glucopyranose donor (Fig. S7a-c). The sugar donor binding site of UDP-D-glucose of TuTreT (shown in Fig. S7d) is similar to what has been reported for TreT from *Pyrococcus horikoshii*. Interestingly, the α,α- and

α,β-trehalose derivatives the α-D-glucopyranose$_{acceptor}$ interacts with *Tu*TreT in a different binding mode than the β-L-galactopyranose$_{acceptor}$ (S7e,f). Arg221 might bind with the 2OH and 3OH of the β-L-galactopyranose$_{acceptor}$ (Fig. S7e), while the 3OH and 4OH of α-D-glucopyranose$_{acceptor}$ is at closer distance to Arg221 (Fig. S7f). The movement of the highly conserved Arg221 when UDP-D-glucose is bound was notable (Fig. 5), which has been postulated to play a role in substrate recognition in *Ph*TreT.[10] Also, Asp256 is at a hydrogen bonding distance to 4OH of α-D-galactopyranose$_{acceptor}$ (Fig. S8e), while no clear electron density for this loop could be found for the natural α-D-glucopyranose$_{acceptor}$. The α-D-glucopyranose$_{acceptor}$ moiety of trehalose is directed into the wide cavity of *Tu*TreT.

Within the protein crystal structure, the conformation of the carbohydrates demonstrated for the enzymatically synthesized α,β-(1→1)-L-galactotrehalose (Fig. 3c,f) a $^1C_4$ conformation of the β-L-galactopyranose moiety. As was shown in Fig. 1, the anomeric β-OH$_{acceptor}$ hydroxyl group points into the same direction as the α-OH$_{acceptor}$ hydroxyl group in the natural substrate D-trehalose (Fig. 3f). Hence, the overall geometry of α-D-glucopyranose with a $^4C_1$ conformation and β-L-galactopyranose with a $^1C_4$ conformation are highly similar (Fig. 7g,h).

Based on these findings, we hypothesized that highly (S)-selective *Tu*TreT guides the anomeric hydroxyl of the sugar acceptor according to a SNi-like mechanism. More specifically, the α-OH of D-glucose ($^4C_1$), or the β-OH of L-galactose ($^1C_4$) are guided by hydrogen bonding from the same face as the NDP leaving group. As the S_Ni-like mechanism requires the same-face participation with the anomeric hydroxyl of the nucleotide phosphate of the glycosyl donor, the sugar coupling does not readily proceed for the equatorial β-OH of D-glucose ($^4C_1$) or the axial α-OH of L-galactose ($^1C_4$). This allows *Tu*TreT to retain its (*S*)-selectivity for the anomeric hydroxyl group, while the anomeric configuration inverts. This mechanistic rationale explains the inversion of anomeric selectivity of *Tu*TreT with L-glycopyranose acceptors, emphasizing that understanding the structural conformations of hexopyranoses is important for understanding enzyme selectivity when glycosidic bonds are formed. This inversion of anomeric selectivity might not be limited to *Tu*TreT, and could occur in other GTs as well.

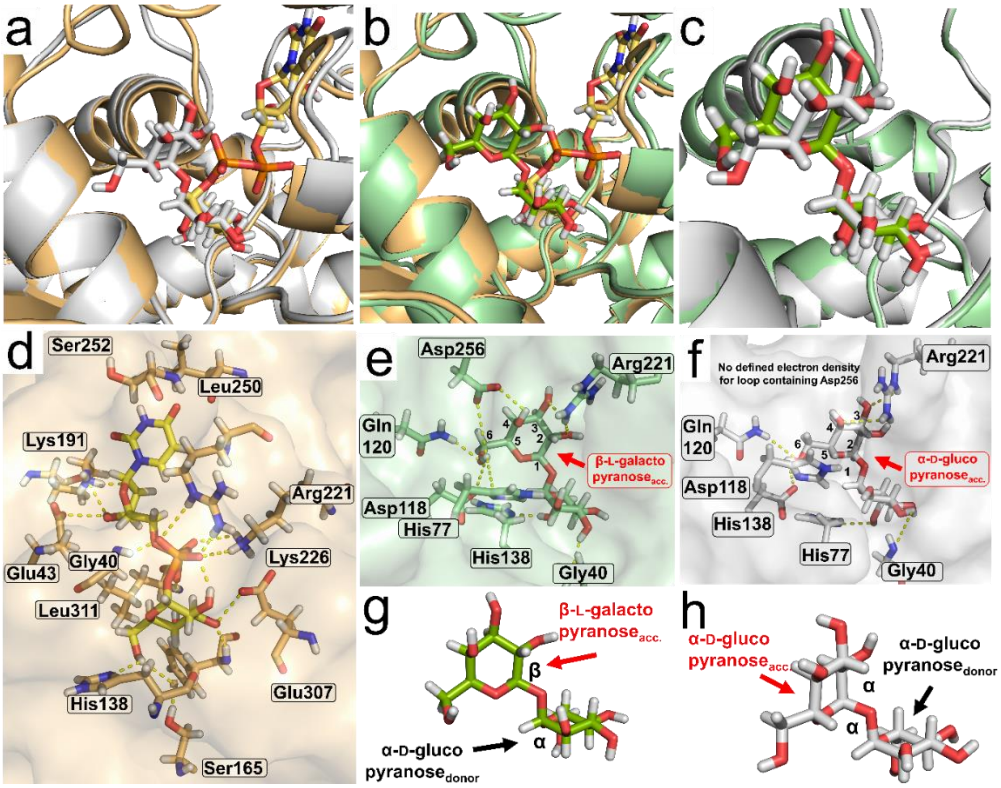


Figure 7: Protein crystal structure alignment of TreT from *Thermoproteus uzoniensis* containing UDP-α-D-glucose (yellow, PDB: 6ZMZ) and D-trehalose (white, PDB: 6ZJH) in **a**, UDP-α-D-glucose (yellow, PDB: 6ZMZ) and α-D-pyranosyl-β-L-galactopyranose (green, PDB: 6ZN1) in **b**. In **a** and **b**, the α-D-glucopyranoses are occupied at the same sugar donor binding site. In **c**, a structural alignment of D-trehalose (white) with α-D-pyranosyl-β-L-galactopyranose (green) is shown, demonstrating the different anomeric and structural configuration of the sugars. The active site residues of *Tu*TreT with UDP-D-glucose (PDB: 6ZMZ) is shown in **d**, α-D-glucopyranosyl-β-L-galactopyranoside (PDB: 6ZN1) in **e**, and D-trehalose (PDB: 6ZJH) are shown in **f**. The stereochemical configuration of the resolved protein crystal structures of β-L-galactopyranose-α-D-glucopyranose in **g** and α,α-D-trehalose in **h**.

Based on the observed inversion of anomeric selectivity with retention of (*S*)-selectivity we hypothesize, that *Tu*TreT also follows an S_Ni-like mechanism. In this mechanism, the α-OH of D-glucose ($^4C_1$), or the β-OH of L-galactose ($^1C_4$) are guided by hydrogen bonding from the same face as the NDP leaving group. As the S_Ni-like mechanism requires the same-face participation with the anomeric hydroxyl of the nucleotide phosphate of the glycosyl donor, the sugar coupling does not readily proceed for the equatorial β-OH of D-glucose ($^4C_1$) or the axial α-OH of L-galactose ($^1C_4$). This allows *Tu*TreT to retain its (*S*)-selectivity for the anomeric hydroxyl group, while the anomeric configuration inverts. This mechanistic rationale explains the inversion of anomeric selectivity of *Tu*TreT with L-glycopyranose acceptors, emphasizing that understanding the structural conformations of hexopyranoses is important for understanding enzyme selectivity when glycosidic bonds are formed.

5.3 Conclusion

mCherry *Tu*TreT catalyzes the formation of trehalose derivatives with a large substrate spectrum. The switch in anomeric selectivity for D- and L-sugar acceptors can be explained due to structural conformations of carbohydrates, leading to the formation of distinctive α-D-α-D- or α-D-β-L-glycosidic linkages. This paves the way for further studies of utilizing rare L-glycopyranoses with retaining LeLoir glycosyltransferases, which are especially interesting for the production of oligosacharides and glycans with unnatural glycosidic linkages.

5.4 Materials and Methods

5.4.1 Chemicals and strains

D-Allose (Carbosynth), L-allose (Carbosynth D-glucose monohydrate (>99.0%, Sigma Aldrich), L-glucose (Carbosynth), D-galactose (>99%, Sigma-Aldrich), L-galactose (Carbosynth), D-mannose (Carbosynth), L-mannose (Carbosynth), D-arabinose (Carbosynth), L-arabinose (Carbosynth), D-xylose (>99%, Sigma-Aldrich), L-xylose (Carbosynth), D-gulose (Carbosynth), L-gulose

(Carbosynth), D-altrose (Carbosynth), L-altrose (Carbosynth), D-tagatose (Carbosynth), L-tagatose (Carbosynth), D-fructose (Carbosynth), D-talose (Carbosynth), *N*-Acetyl-D-glucosamine (≥99%, Sigma-Aldrich), 2-fluoro-2-deoxy-D-glucose (Carbosynth), 3-fluoro-3-deoxy-D-glucose (Carbosynth), 4-fluoro-4-deoxy-D-glucose (Carbosynth), 6-fluoro-6-deoxy-D-glucose (Carbosynth), ampicillin sodium salt (Sigma-Aldrich), Bacto tryptone (Brunschwig), Bacto yeast extract (Brunschwig), deoxyribonuclease I from bovine pancreas (protein 85%, 400 000 U mg^{-1}, Sigma-Aldrich), disodium hydrogen phosphate dodecahydrate (VWR), glycerol (>99.5%, Sigma-Aldrich), HEPES (>99.5%m Sigma-Aldrich), lysozyme from chicken egg white (Crystalline powder, ~7000 U mg^{-1}, Sigma Aldrich), sodium chloride (99.9%, Sigma-Aldrich), sodium dihydrogen phosphate dihydrate (VWR), Tris(hydroxymethyl)aminomethane (Tris, 99%, Sigma-Aldrich), UDP-D-glucose disodium salt (Carbosynth), UDP-D-GlcNAc (Carbosynth).

Terrific broth medium consisting of 1.20% (wt/wt) tryptone, 2.40% (wt/wt) yeast extract, 53 mM K_2HPO_4, 16 mM KH_2PO_4, and 4% (wt/wt) glycerol was autoclaved at 121°C for 20 min. The phosphate buffer in TB medium was autoclaved separately. LB medium consisting of 1.00% (wt/wt) tryptone, 0.5% (wt/wt) yeast extract, and 1% NaCl was autoclaved at 121°C for 20 min. All media were supplemented with 100 µg mL^{-1} ampicillin. The pBAD/His A mCherry *Tu*TreT vector was used as described previously.[4b, 13] *E. coli* Top10 with the genotype F^- *mcrA* Δ*(mrr-hsdRMS-mcrBC) ϕ80lacZ* ΔM15 Δ*lacX74 recA1 araD139* Δ(*ara-leu*)7697 *galU galK rpsL* (Str^r) *endA1 nupG* was ordered from a commercial supplier (Invitrogen).

5.4.2 Materials

A stainless-steel bioreactor (Applikon) with a total volume of 20 L, inner diameter of 211 mm, inner height of 631 mm, liquid height of 440 mm, with a height of 2100 mm, and working volume 15 L was utilized. The bioreactor was equipped with a pH electrode (Mettler Toledo), a pO_2 electrode (Mettler Toledo), a thermometer (Mettler Toledo), a liquid level detector connected to a peristaltic pump (Masterflex) coupled to a bottle of Foam Clear Escaferm antifoam (KCC

Basildon). Temperature, stirrer speed and pressure were regulated. Stirrer speed pH, temperature, pO_2 and pressure were evaluated using MFCS/win V3 (Satorius). Chromatographic analysis of reaction products was performed using a Shimadzu high-performance liquid chromatography (HPLC) system equipped with an Imtakt Unison-UK amino column (0.4x25 cm, 60°C), an evaporative light-scattering detector (Shimadzu ELSD-LTII), a UV detector (SPD-20A), and acetonitrile-water-formic acid (80:20:0.1) as mobile phase (1 mL min^{-1}). ^{1}H-NMR and ^{13}C-NMR spectra were recorded on an Agilent 400 MHz (9.4 Tesla) spectrometer operating at 399.67 MHz for ^{1}H at 298 K. Spectra were interpreted using MestReNova. Data for NMR is reported as following: chemical shift (δ ppm), multiplicity (s = singlet, d = duplet, dd = double duplet, t = triplet, q = quartet, m = multiplet). Purified enzyme concentrations were determined with a Cary 60 UV-VIS spectrometer (Agilent) using the previously determined molar extinction coefficient of mCherry *Tu*TreT[4b, 13].

5.4.3 SDS-PAGE analysis

Protein samples were denatured using XT sample buffer (Bio-Rad) supplemented with XT reducing agent (Bio-Rad) at 95°C for 1 minute. Gel electrophoresis was performed with Criterion XT 4-12% bis-Tris precast gels (Bio-Rad) using XT MES Running buffer (Bio-Rad). The gels were run at 150 V for 30 to 45 minutes and stained with SimplyBlue SafeStain (Novex). A Precision Plus Protein Unstained Standard (Bio-Rad) was used to determine the relative molecular mass of the protein.

5.4.4 Stainless steel batch reactor fermentation

An inoculum culture of mCherry *Tu*TreT *E. coli* Top10 pBAD/His A was grown in 5 mL LB-medium containing 100 µg mL^{-1} ampicillin in a 15 mL Cellstar Greiner tube. The culture was grown overnight in an incubator at 37°C and 180 rpm. After 12 hours, the culture was transferred into a 2L Erlenmeyer flask containing 395 mL LB medium and 100 µg mL^{-1} ampicillin and was grown overnight as inoculum at 37°C and 180 rpm. The stainless-steel bioreactor (Applikon) was

sterilized with steam before use. The bioreactor was charged with 13.3 L TB-medium (12 g L^{-1} tryptone, 24 g L^{-1} yeast extract, 5 g L^{-1} glycerol, 23.1 g L^{-1} KH_2PO_4, 125.4 g L^{-1} K_2HPO_4) containing 100 µg mL^{-1} ampicillin. The pH, pressure, volume level, temperature, and pO_2 were controlled and measured. The addition of antifoam was controlled by the degree of foaming. Air was sparged through the medium at 5 L min^{-1}. Every hour samples were taken for OD_{600} measurements and SDS page analysis. After 2 to 3 hours when OD_{600} reached > 0.5 the culture was induced with a final concentration of 0.20 g L^{-1} L-arabinose. The fermentation was stopped after 18 h of fermentation. The cells were harvested by centrifugation on a RC 6+ (Sorvall) at 10 000 rpm (rotor: F12S 6X500, 10°C), washed with Tris buffer (20 mM, pH 8.0) followed by centrifugation at 10 000 rpm (rotor: F12S 6X500, 10°C), and stored at -20 °C.

5.4.5 Cell disruption of *E. coli* Top10 cells

The cell-pellets were thawed and homogenized in a Tris-HCl disruption buffer (20 mM Tris, 20 mM imidazole, 500 mM NaCl, 2 mg mL^{-1} lysozyme, 0.1 mg mL^{-1} DNase I, pH = 7.4) with a ratio of 4 to 5 mL buffer per gram of wet cell pellet. The suspended cells were lysed with a cell-disruptor (Constant systems Ltd., 1.8 kbar) during one to three rounds. The disrupted cells were centrifuged at 10 000 rpm (rotor: F10S 6X500Y or F12S 6X500, 10°C) for 10 minutes. The supernatant was transferred to polypropylene 50 mL Falcon tubes and incubated in a water bath (60°C for 20 minutes) for a heat treatment. After the heat treatment, the tubes were cooled down to room temperature and centrifuged (5810R, Eppendorf, 10 min, 10 000 rpm, 10 °C). The clear supernatant was collected for affinity chromatography and analyzed by SDS-PAGE.

5.4.6 Protein purification of mCherry *Tu*TreT

The heat-treated supernatant was purified by affinity chromatography using a XK 16/40 cartridge (GE Healthcare) containing 41 mL of Ni Sepharose 6Fast Flow (GE Healthcare) was utilized. After binding, the column was washed with

binding buffer (20 mM Tris, 20 mM imidazole, 500 mM NaCl, pH 7.4) and eluted using elution buffer (Tris 20 mM, 500 mM imidazole, 500 mM NaCl, pH 7.4) over a gradient of 10 column volumes. Fractions were collected in 8 mL polypropylene tubes. Protein samples were concentrated in a 12 mL Amicon Ultra Centrifugal filter (Merck) with a 30 kDa cutoff. The purified protein was transferred to 1-2 mL Eppendorf tubes and shock-frozen in liquid nitrogen. Samples were analyzed by SDS-PAGE and HPLC analysis.

5.4.7 Screening of sugar acceptors with mCherry *Tu*TreT using HPLC analysis

Reaction solutions containing UDP-D-glucose, HEPES (50 mM, pH 7.0), $MgCl_2$ (20 mM), 1.0 mg mL^{-1} mCherry *Tu*TreT were prepared with different sugar acceptors (10 mM) to afford a final volume of 1.0 mL in 1.5 mL Eppendorf tubes. The reaction was incubated in an Eppendorf thermomixer C (500 rpm and 60°C). Samples were quenched by mixing 60 µL of reaction volume with 60 µL of 80:20:0.1 ACN:H_2O:formic acid and incubated for at least 60 minutes at -80 °C. After thawing, the samples were centrifuged at 14 000 rpm (5417R Eppendorf, 4°C) for 1 min. The supernatant was analyzed by HPLC (column: Imtakt UK-Amino 250 x 4.6 mm, 50 °C, ELSD, 80:20:0.1 ACN:H_2O:formic acid, 1.0 mL min^{-1}). The substrate conversion was quantified by using external standards for D-glucose, D-mannose, D-galactose, D-xylose, L-arabinose, D-allose, D-altrose, D-trehalose, L-gulose, D-tagatose, D-fructose, 2-fluoro-2-deoxy-D-glucose, 3-fluoro-3-deoxy-D-glucose, 4-fluoro-4-deoxy-D-glucose, 6-fluoro-6-deoxy-D-glucose, and D-talose.

5.4.8 HRMS analysis

The purified trehalose analogues were dissolved in 50:50:0.1 acetonitrile:water:formic acid and transferred to glass HPLC vials. A HPLC 1100 system (Agilent) was equipped with a GlycoSep R, 3 µm, 4.6 x 150 mm, C18, HPLC column (Glyko, cat# GKI-4727) using 84% H_2O, 9% acetonitrile, 7% methanol plus 0.1% formic acid as solvent A and acetonitrile plus 0.1% formic acid as solvent B. The injection volume was 5 µL. An isocratic elution of 10% B was

maintained at a flowrate of 0.3 mL/min over 10 minutes, followed by two injections of mobile phase A to prevent carry over between sample injections. Full MS scanning mode at resolution 70.000 and PRM scanning mode at resolution 17.500 for the selected masses 343.1240, 313.1135, and 384.1506 using 26 eV collision energy were performed using a Q-Exactive Focus Orbitrap mass spectrometer (Thermo Scientific) operated in ES+ mode.

5.4.9 *Tu*TreT maintains catalytic activity after proteolysis.

To 100 a µL solution containing 10 mg mL^{-1} mCherry *Tu*TreT in HEPES buffer (50 mM, pH 7.0) with $MgCl_2$ (20 mM) was either 0, 10, 20, or 40 µg elastase (type IV, porcine pancrease) added and incubated for 18 hours. The reaction mixture was analyzed by SDS-PAGE after cleavage of the GGSGGGGSGG-linker (Fig. 8). A 1.0 mL reaction solution containing UDP-D-glucose, D-glucose (10 mM), HEPES (50 mM, pH 7.0), $MgCl_2$ (20 mM), and 0.25 mg of mCherry *Tu*TreT pretreated with elastase was prepared. The reaction was incubated at 60 °C in an Eppendorf thermomixer C (500 rpm and 60°C). The reaction was monitored by quenching 60 µL of reaction volume in 60 µL 80:20:0.1 ACN:H_2O:formic acid and incubated for at least 60 minutes at -80 °C.After thawing, the samples were centrifuged at 14 000 rpm (5417R Eppendorf, 4°C) for 1 min. The supernatant was collected and analyzed by HPLC (column: Imtakt UK-Amino 250 x 4.6 mm, 50 °C, ELSD, 80:20:0.1 ACN:H_2O:formic acid, 1.0 mL min^{-1}). The substrate conversion was quantified using external standards for D-trehalose (Fig. 8b)

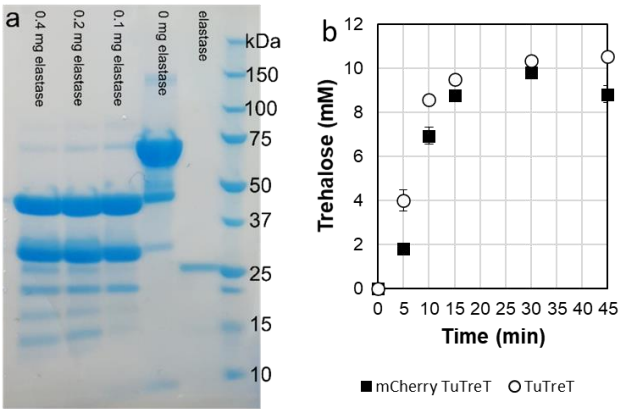


Figure 8: SDS-PAGE analysis of mCherry *Tu*TreT after incubation for 72 hours with 0 – 0.4 mg of elastase shown in **a**. The enzyme catalyzed production of trehalose according to HPLC analysis was followed before (grey) and after cleavage (black) in **b**. Reaction conditions: D-glucose (10 mM), UDP-D-glucose (40 mM), HEPES (50 mM, pH 7), $MgCl_2$ (20 mM), 60 °C, 800 rpm. The activity assay contained 0.25 mg for mCherry *Tu*TreT (no elastase present) or *Tu*TreT (0.2 mg elastase present). A Precision Plus Protein Standard (Bio-Rad) was used with Criterion XT 4 to 12% bis-Tris precast gels (Bio-Rad) using XT MES Running buffer (Bio-Rad) to determine the molecular weight of *Tu*TreT.

5.4.10 *"In situ"* proteolysis and enzyme purification of *Tu*TreT

0.1 mL of purified fusion protein mCherry *Tu*TreT with the glycine-rich GGSGGGGSGG-linker (10 mg mL^{-1}) was incubated with the protease elastase (1.0 mg mL^{-1}, porcine pancreas, ≥ 4 U mg^{-1}) at 400 rpm for 72 hours at 20 °C, in a 1.5 mL polypropylene Eppendorf tube in HEPES buffer (50 mM, pH 7.0) with NaCl (100 mM). Next, the protein was purified by affinity chromatography using 1 ml of Ni-Sepharose Excel beads (GE Healthcare) by washing the beads with 20 mL of binding buffer (50 mM HEPES, 100 mM NaCl, pH 7.0). The protein was eluted with elution buffer (50 mM HEPES, 500 mM imidazole, 100 mM NaCl, pH 7.0). Protein samples were concentrated using an Amplicon Ultra Centrifugal filter (Merck) with a 30 kDa cutoff in HEPES (50 mM, pH 7.0) containing NaCl (100 mM). All samples were analyzed before, during, and after purification by SDS-PAGE (Fig. 9).

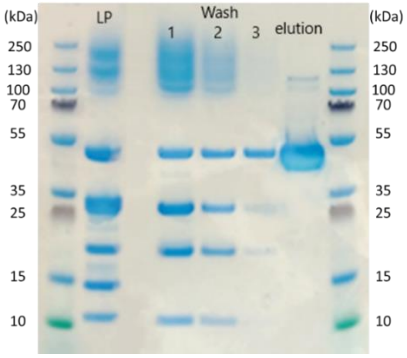


Figure 9: SDS-PAGE analysis of mCherry *Tu*TreT after proteolysis (LP), after the first, second, and third washing round of 5 ml binding buffer (wash 1, 2, 3), and purified *Tu*TreT after elution. A PageRuler Plus Pre-stained Protein Ladder (ThermoFischer) was used with Criterion XT 4 to 12% Bis-Tris precast gels (Bio-Rad) using XT MES Running buffer (Bio-Rad) to determine the molecular weight of *Tu*TreT.

5.4.11 *Tu*Tret crystallization and crystal soaking

Crystals of *Tu*TreT were obtained at 292K using the vapour diffusion method and a crystallization solution containing 0.1 M Bis-Tris-Propane pH 6.5, 15-20% PEG3350, 0.3 M KSCN, and 1-4% PEG400. Crystals appeared between 1-3 days, and grew to their final size in to 7-14 days. Crystals were soaked with different ligands by slowly exchanging, over the course of 60 minutes, the drop solution with a new solution containing the crystallization solution plus 25% of PEG400 and 50 mM of the desired ligand . The crystals were flashed freezed and stored in liquid N_2 until X-ray data collection was performed.

5.4.12 SEC-SAXS data collection

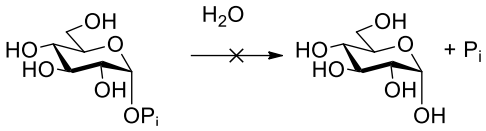
The SAXS data from *Tu*TreT, measured as $I(s)$ vs. s, where $s = 4\pi\sin\theta/\lambda$, 2θ is the scattering angle and λ the X-ray wavelength (0.124 nm, 10 keV), were collected at the EMBL P12-bioSAXS beam line at the PETRAIII storage ring (DESY, Hamburg) equipped with a Pilatus 6M 2D area detector (Dectris, Switzerland).[11] In-line size exclusion chromatography SAXS (SEC-SAXS) was employed whereby the SEC eluates were passed directly from the column outlet

and continuously flowed through the beam line during the measurements. Details of the SEC parameters, running buffer composition as well as the concentrations of $MgCl_2$, D-trehalose and UDP-α-D-glucose additive used in the running buffers for each successive SEC-SAXS measurement are provided in Table 1. SAXS data were acquired every 0.5 s spanning one column volume per experiment (1320 individual data frames). The 2D-data underwent azimuthal averaging using the P12 *SASFLOW*[14] pipeline to generate 1D-scattering profiles that underwent further data reduction using *CHROMIXS*[15] to yield the final background-corrected, averaged sample scattering profiles derived from each respective *Tu*TreT sample SEC elution peak. Additional data processing and evaluation steps were performed using the ATSAS 2.8 software suite,[16] that included determining the R_g from the Guinier approximation[17] ($\ln I(s)$ vs s^2, for $sR_g < 1.3$) and from the calculated probable frequency of real-space distances, or $p(r)$ profile.[18] In addition, all subsequent data-data comparisons and data-model fits were assessed using the reduced χ^2 test and the Correlation Map (CorMap) *P*-value, set at a significance threshold, α, of 0.01 (where a χ^2 in the range 0.9-1.1 or CorMap $P > 0.01$ are deemed statistically significant).[19]

Table 1: SEC-SAXS data reporting for *Tu*TreT

Sample details				
Organism	Thermoproteus uzoniensis			
Uniprot ID	F2L613			
SEC-SAXS buffer	50 mM HEPES 100 NaCl 4 mM DTT pH 7 +/-additive			
Sample injection volume	40 µl			
Sample injection conc.	6 mg/ml			
SEC column	GE S200 Increase 5/150			
SEC flow rate	0.3 ml/min			
Temperature	20 ^{o}C			
Instrument details				
Instrument	EMBL-P12 BioSAXS, PETRAIII DESY			
Exposure time/# frames	0.5 s/(1320 – full column elution)			
X-ray wavelength/energy	0.124 nm/10 keV			
Sample-to-detector distance	3 m			
Scattering intensity scale	arbitrary unit, a.u.			
SEC-SAXS primary data processing	EMBL-P12 SASFLOW pipeline and CHROMIXS			
SAMPLE	**Tret-apo**	**Tret-Mg**	**Tret-Mg-Trehalose**	**Tret-Mg-UDP-glucose**
Additive concentration	none	20 mM $MgCl_2$	20 mM $MgCl_2$ + 1 mM Trehalose	20 mM $MgCl_2$ + 1 mM UDP-glucose
# frames used for averaging	27(13.5 s)	30(15 s)	30(15 s)	29(14.5 s)
Working s-range (nm^{-1})	0.092-4	0.081-4	0.097-4	0.108-4
Guinier analysis:				
Primary data analysis software	PRIMUSQT			
Guinier I(0) (s)	10724(10)	10456(10)	10819(8)	10937(8)
R_g (Guinier, nm) (s)	2.46(0.05)	2.43(0.05)	2.44(0.05)	2.44(0.05)
sR_g range/(points used)	0.23-1.3(1-158)	0.2-1.3(1-164)	0.24-1.3(1-157)	0.26-1.3(1-153)
p(r) analysis:				
Method	GNOM5			
I(0), POR (s)	10740(11)	10450(10)	10820(10)	10930(10)
R_g (POR, nm) (s)	2.47(0.01)	2.43(0.01)	2.44(0.01)	2.44(0.01)
D_{max} (nm)	8.8	8.5	8.3	8.5
Quality of fit, CorMap P /c^2	0.94/1.01	0.29/0.99	0.08/1.04	0.94/1.05
Porod volume (nm^3)	77	76	76	73
Shape classification	compact	compact	compact	compact
MW analysis:				
Calculated MW, from amino acid sequence – inc. His tag	46.7 kDa			
MW kDa (Bayesian)/probability	47.7/0.4	43.8/0.3	45.7/0.3	46.7/0.4
MW Credibility Interval, kDa	44-49	41-46	43-47	44-48
Ab initio modeling:				
Method	DAMMIN			
Symmetry	P1			
Quality-of-fit, CorMap P /c^2	0.87/1.02	0.40/0.99	0.40/1.11	0.87/1.05
Average NSD (10 models)	0.58+/-0.01	0.60+/-0.02	0.57+/-0.01	0.59+/-0.01
Method	GASBOR			
Symmetry	P1			
Quality-of-fit, CorMap P /c^2	0.74-0.94/1.02	0.29/1.0	0.08/1.05	0.49-0.75/1.06
Average NSD (10 models)	0.93+/-0.03	0.95+/-0.04	0.92+/-0.04	0.92+/-0.05
Rigid-body modelling:				
Refinement Method	SREFLEX			
Symmetry	P1			
Atomistic template model	PDB 6ZJ4 Tret-apo			
Template model R_g, nm	2.3			
Template model, CorMap P /c^2	0/3.23	0/2.58	0/2.78	0/3.05
SREFLEX model R_g, nm	2.45	2.41	2.43	2.41
SREFLEX model, CorMap P /c^2	0.08/1.07	0.29/1.00	0.08/1.03	0.49/1.12
Fitting method	CRYSOL			
#harmonics/points/s-max (nm^{-1})	30/200/4.0			
Constant adjustment	Yes			
Data Accession code(s)				
SASBDB	SASDHG9	SASDHH9	SASDHJ9	SASDHK9

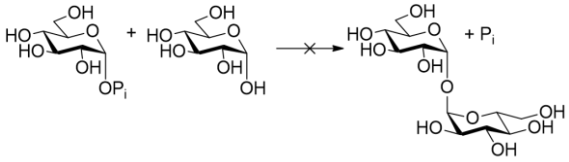
5.4.13 Control experiment for hydrolysis of α-D-glucose-1-phosphate



Scheme S1: Control experiment for mCherry *Tu*TreT catalyzed hydrolysis of α-D-glucose-1-phosphate. **Reaction conditions**: α-D-glucose-1-phosphate (10 or 100 mM), $MgCl_2$ (20 mM), HEPES (50 mM, pH 7.0), mCherry *Tu*TreT (3.0 mg mL^{-1}), reaction time 60 minutes, 800 rpm, 60 °C.

A 1.0 mL reaction solution containing α-D-glucose-1-phosphate (10 or 100 mM), HEPES (50 mM, pH 7.0), $MgCl_2$ (20 mM), mCherry *Tu*TreT (3.0 mg mL^{-1}) were prepared. The reaction was incubated at 60 °C in an Eppendorf thermomixer C at 800 rpm and 60°C. The enzymatic reaction was quenched in 60 µL 80:20:0.1 ACN:H_2O:formic acid and incubated for at least 30 minutes at -80 °C. The samples were centrifuged at 14 000 rpm (5417R Eppendorf, 4°C) for 1 min. The supernatant was collected and analyzed by HPLC (column: Imtakt UK-Amino 250 x 4.6 mm, 50 °C, ELSD, 80:20:0.1 ACN:H_2O:formic acid, 1.0 mL min^{-1}). Substrate concentration was quantified using the external calibration curves.

5.4.14 Control experiment for phosphorylase activity

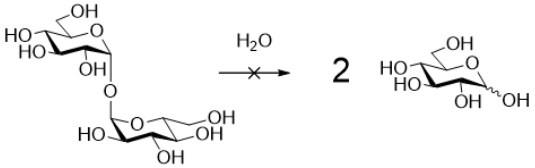


Scheme S2: Control experiment for the mCherry *Tu*TreT catalyzed synthesis of trehalose using α-D-glucose-1-phosphate and D-glucose. **Reaction conditions**: α-D-glucose-1-phosphate (10 or 100 mM), D-glucose (10 mM), $MgCl_2$ (20 mM), HEPES (50 mM, pH 7.0), mCherry *Tu*TreT (3.0 mg mL^{-1}), reaction time 60 minutes, 800 rpm, 60 °C.

A 1.0 mL reaction solution containing α-D-glucose-1-phosphate (10 or 100 mM), D-glucose (10 mM), HEPES (50 mM, pH 7.0), $MgCl_2$ (20 mM), 3.0 mg mL^{-1}

mCherry *Tu*TreT were prepared. The reaction was incubated in an Eppendorf thermomixer C at 800 rpm and 60°C. The enzymatic reaction was quenched in 60 µL 80:20:0.1 ACN:H_2O:formic acid and incubated for at least 30 minutes at -80 °C. The samples were centrifuged at 14 000 rpm (5417R Eppendorf, 4°C) for 1 min. The supernatant was collected and analyzed by HPLC (column: Imtakt UK-Amino 250 x 4.6 mm, 50 °C, ELSD, 80:20:0.1 ACN:H_2O:formic acid, 1.0 mL min^{-1}).

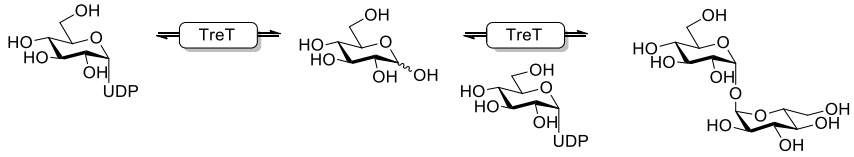
5.4.153 Control experiment for hydrolase activity



Scheme S3: Control experiment for the mCherry *Tu*TreT catalyzed hydrolysis of trehalose. **Reaction conditions**: D-trehalose (10 mM), $MgCl_2$ (20 mM), HEPES (50 mM, pH 7.0), mCherry *Tu*TreT (3.0 mg mL^{-1}), reaction time 60 minutes, 800 rpm, 60 °C.

A 1.0 mL reaction solution containing D-trehalose (10 mM), HEPES (50 mM, pH 7.0), $MgCl_2$ (20 mM), mCherry *Tu*TreT (3.0 mg mL^{-1}) were prepared. The reaction was incubated in an Eppendorf thermomixer C at 800 rpm and 60°C. The enzymatic reaction was quenched in 60 µL 80:20:0.1 ACN:H_2O:formic acid and incubated for at least 30 minutes at -80 °C. The samples were centrifuged at 14 000 rpm (5417R Eppendorf, 4°C) for 1 min. The supernatant was collected and analyzed by HPLC (column: Imtakt UK-Amino 250 x 4.6 mm, 50 °C, ELSD, 80:20:0.1 ACN:H_2O:formic acid, 1.0 mL min^{-1}).

5.4.16 Control experiment for hydrolase activity of UDP-D-glucose

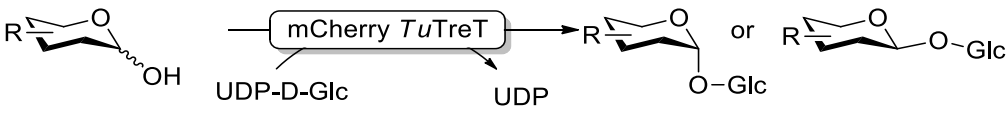


Scheme S4: Control experiment for the mCherry *Tu*TreT catalyzed hydrolysis of UDP-glucose towards D-glucose and UDP. **Reaction conditions**: UDP-D-glucose (40 mM), $MgCl_2$

(20 mM), HEPES (50 mM, pH 7.0), mCherry *Tu*TreT (3.0 mg mL^{-1}), reaction time 60 minutes, 800 rpm, 60 °C.

A 1.0 mL reaction solution containing D-trehalose (10 mM), D-glucose (10 mM), HEPES (50 mM, pH 7.0), $MgCl_2$ (20 mM), 3.0 mg mL^{-1} mCherry *Tu*TreT were prepared. The reaction was incubated in an Eppendorf thermomixer C at 800 rpm and 60°C. The enzymatic reaction was quenched in 60 µL 80:20:0.1 ACN:H_2O:formic acid and incubated for at least 30 minutes at -80 °C. The samples were centrifuged at 14 000 rpm (5417R Eppendorf, 4°C) for 1 min. The supernatant was collected and analyzed by HPLC (column: Imtakt UK-Amino 250 x 4.6 mm, 50 °C, ELSD, 80:20:0.1 ACN:H_2O:formic acid, 1.0 mL min^{-1}. The substrate conversion was quantified by using external standards for D-trehalose.

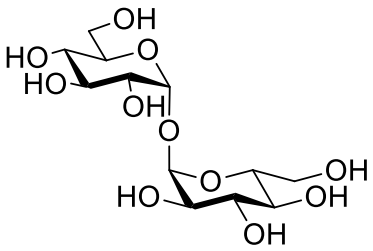
5.4.17 Semi-preparative enzymatic synthesis of trehalose derivatives



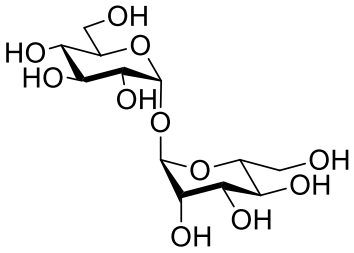
Scheme 5: Enzymatic sugar coupling for preparative scale reactions.

The reaction was started by the addition of enzyme resulting in a reaction mixture containing a monosaccharide (10 mM), UDP-D-glucose (40 mM) or UDP-GlcNAc (40 mM), HEPES (50 mM, pH 7.0), mCherry *Tu*TreT (variable amounts), and $MgCl_2$ (20 mM) with a reaction volume of 15 mL. Reaction was incubated at 60 °C with gentle shaking and snap-frozen in liquid nitrogen and stored at -20 °C. The trehalose derivatives were purified using ultra-filtration (Amicon Ultra centrifugal 3 kDa filter, Merck) and the residue was washed with two times 7.5 mL MilliQ water. The combined filtrate was desalted with stepwise addition of a mixed-bed ion exchange resin (Bio-Rex RG 501-X8, Bio-Rad). The desalted filtrated was lyophilized, dissolved in 0.6 mL MilliQ water, and purified by SEC chromatography

using P-2 Bio-Gel resin (Bio-Rad) in a XK 16/100 column (GE Healthcare) attached to NGC medium-pressure liquid chromatography system (Bio-Rad). The final product was lyophilized and analyzed by NMR and MS.

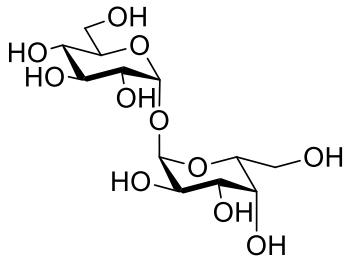


α-D-glucopyranosyl-α-D-glucopyranose: 0.03 mg mL^{-1} mCherry *Tu*TreT, isolated yield 27.1 mg (53%) of white crystals. ^{1}H-NMR (400 MHz, D_2O) = δ 3.46 (t, $J_{3,4\text{-α-D-glc}} = J_{4,5\text{-α-D-glc}}$ = 9.6 Hz, H4$_{\text{α-D-glc}}$, 2H), 3.66 (dd, $J_{1,2\text{-α-D-glc}}$ = 3.6 Hz, $J_{2,3\text{-α-D-glc}}$ = 10.0 Hz, H2$_{\text{α-D-glc}}$, 2H), 3.75 ($J_{5,6b\text{-α-D-glc}}$ = 5.2 Hz, $J_{6a,6b\text{-α-D-glc}}$ = 12.0 Hz, H6b$_{\text{α-D-glc}}$, 2H), 3.81 – 3.89 (m, H3$_{\text{α-D-glc}}$, H5$_{\text{α-D-glc}}$, H6a$_{\text{α-D-glc}}$, 6H), 5.20 (d, $J_{1,2\text{-α-D-glc}}$ = 3.6 Hz, H1$_{\text{α-D-glc}}$, 2H).**^{13}C-NMR** (101 MHz, D_2O) = δ 92.0 (C1$_{\text{α-D-glc}}$, C*H*), 71.3 (C3$_{\text{α-D-glc}}$, C*H*), 70.92 (C5$_{\text{α-D-glc}}$, C*H*CH$_2$), 69.81 (C2$_{\text{α-D-glc}}$, C*H*), 68.47 (C4$_{\text{α-D-glc}}$, C*H*), 59.3 (C6$_{\text{α-D-glc}}$, CHC*H$_2$*). **HRMS** (TOF): 365.1048 (calc. 365.1054 [M+Na]).
These ^{13}C- and ^{1}H-NMR signals matched the literature values:[20] **^{1}H-NMR** (800 MHz, D_2O) = δ 3.43 ($J_{3,4}$ = 9.8 Hz, $J_{4,5}$ = 9.8 Hz, H4$_{\text{α-glc}}$, 2H), 3.64 ($J_{1,2}$ = 3.8 Hz, $J_{2,3}$ = 9.8 Hz, H2$_{\text{α-glc}}$, 2H), 3.75 ($J_{5,6b}$ = 5.1 Hz, $J_{6a,6b}$ = 11.9 Hz, H6b$_{\text{α-glc}}$, 2H), 3.81 ($J_{4,5}$ = 9.8 Hz, $J_{5,6b}$ = 5.1 Hz, $J_{5,6a}$ = 2.0 Hz, H5$_{\text{α-glc}}$, 2H), 3.84 ($J_{2,3}$ = 9.8 Hz, $J_{3,4}$ = 9.8 Hz, H3$_{\text{α-glc}}$, 2H), 3.85 ($J_{5,6a}$ = 2.0 Hz, $J_{6a,6b}$ = 11.9 Hz, H6a$_{\text{α-glc}}$, 2H), 5.18 ($J_{1,2}$ = 3.8 Hz, H5$_{\text{α-glc}}$, 2H). **^{13}C-NMR** (125 MHz, D_2O) = δ 93.84 (C1$_{\text{α-glc}}$, C*H*), 73.14 (C3$_{\text{α-glc}}$, C*H*), 72.76 (C5$_{\text{α-glc}}$, C*H*CH$_2$), 71.66 (C2$_{\text{α-glc}}$, C*H*), 70.32 (C4$_{\text{α-glc}}$, C*H*), 61.16 (C6$_{\text{α-glc}}$, CHC*H$_2$*).

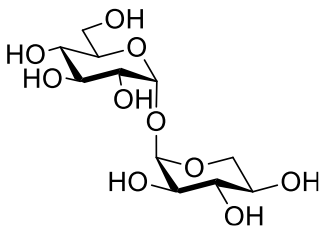


α-D-glucopyranosyl-α-D-mannopyranose: 0.33 mg mL^{-1} mCherry *Tu*TreT, isolated yield 27.1 mg (53%) of white crystals. **^{1}H-NMR** (400 MHz, D_2O) = δ 3.45 (t, $J_{3,4\text{-α-D-glc}} = J_{4,5\text{-α-D-glc}}$ = 9.6 Hz, H4$_{\text{α-glc}}$, 1H), 3.64 (dd, $J_{1,2\text{-α-D-glc}}$ = 4.0 Hz, $J_{2,3\text{-α-D-glc}}$ = 10.0 Hz, H2$_{\text{α-glc}}$, 1H), 3.69 (t, $J_{3,4\text{-α-D-man}} = J_{4,5\text{-α-D-man}}$ = 9.6 Hz, H4$_{\text{α-D-man}}$, 1H), 3.74 - 3.89 (m, H3$_{\text{α-D-glc}}$, H5$_{\text{α-D-glc}}$, H5$_{\text{α-D-man}}$, H6a$_{\text{α-D-man}}$, H6b$_{\text{α-D-man}}$ H6a$_{\text{α-D-glc}}$, H6b$_{\text{α-D-glc}}$, 7H), 3.96 (dd, $J_{2,3\text{-α-D-man}}$ = 4.0 Hz, $J_{3,4\text{-α-D-man}}$ = 9.6 Hz, H3$_{\text{α-D-man}}$, 1H), 4.01 (dd, $J_{1,2\text{-α-D-man}}$ = 4.4 Hz, $J_{2,3\text{-α-D-man}}$= 1.6 Hz, H2$_{\text{α-D-man}}$, 1H), 5.14 ($J_{1,2\text{-α-D-man}}$ = 1.6 Hz, H1$_{\text{α-D-man}}$, 1H), 5.20 (d, $J_{1,2\text{-α-D-glc}}$ = 3.6 Hz, H1$_{\text{α-D-glc}}$, 1H). **^{13}C-NMR** (101 MHz, D_2O) = δ 94.9 (C1$_{\text{α-D-man}}$, C*H*), 93.4 (C1$_{\text{α-D-glc}}$, C*H*), 73.2 (C3$_{\text{α-D-glc}}$, C*H*), 72.5 (C5$_{\text{α-D-glc}}$, C*H*CH$_2$), 72.4 (C5$_{\text{α-D-man}}$, C*H*CH$_2$), 70.8 (C2$_{\text{α-D-glc}}$, C*H*), 70.2 (C3$_{\text{α-D-man}}$, C*H*), 69.9 (C2$_{\text{α-D-man}}$, C*H*), 69.5 (C4$_{\text{α-D-glc}}$, C*H*), 66.6 (C4$_{\text{α-D-man}}$, C*H*), 60.8 (C6$_{\text{α-D-man}}$, CHC*H$_2$*),

60.5 ($C6_{\alpha\text{-D-glc}}$, CHCH_2). **HRMS** (ESI): 365.1049 (calc. 365.1054 [M+Na]).
These ^{13}C- and ^{1}H-NMR signals matched the literature values:[21] **^{1}H-NMR** (400 MHz, D_2O) = δ 5.08 (*J* = 3.6 Hz, $H1_{\alpha\text{-D-glc}}$, 1H), 5.02 (*J* = 1.5 Hz, $H1_{\alpha\text{-D-man}}$, 1H), 3.89 (*J* = 3.6 Hz, $H2_{\alpha\text{-D-man}}$, 1H), 3.84 (*J* = 9.9 Hz, $H3_{\alpha\text{-D-man}}$, 1H), 3.76-3.66 (m, $H6a_{\alpha\text{-D-man}}$, $H6a_{\alpha\text{-glc}}$, $H5_{\alpha\text{-D-man}}$, $H6b_{\alpha\text{-D-glc}}$, 4H), 3.65 (*J* = 9.5 Hz, $H3_{\alpha\text{-D-glc}}$, 1H), 3.64-3.61 (m, $H6b_{\alpha\text{-D-man}}$, $H5_{\alpha\text{-glc}}$, 2H), 3.61 (*J* = 9.5 Hz, $H4_{\alpha\text{-D-man}}$, 1H), 3.52 (*J* = 9.5 Hz, $H2_{\alpha\text{-D-glc}}$, 1H), 3.34 (*J* = 9.5 Hz, $H4_{\alpha\text{-D-glc}}$, 1H). **^{13}C-NMR** (67.89 MHz, D_2O) = δ 96.0 ($C1_{\alpha\text{-D-man}}$, C*H*), 94.4 ($C1_{\alpha\text{-D-glc}}$, C*H*), 74.2 ($C3_{\alpha\text{-D-glc}}$, C*H*), 73.6 ($C5_{\alpha\text{-D-man}}$, C*H*CH_2), 73.4 ($C5_{\alpha\text{-D-glc}}$, C*H*CH_2), 71.8 ($C2_{\alpha\text{-D-glc}}$, C*H*), 71.2($C3_{\alpha\text{-D-man}}$, C*H*), 71.0 ($C2_{\alpha\text{-D-man}}$, C*H*), 70.6 ($C4_{\alpha\text{-D-glc}}$, C*H*), 67.7 ($C4_{\alpha\text{-D-man}}$, C*H*), 61.9 ($C6_{\alpha\text{-D-man}}$, CHCH_2), 61.6 ($C6_{\alpha\text{-D-glc}}$, CHCH_2).



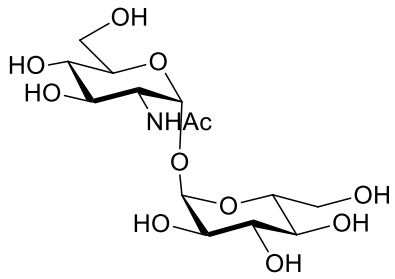
α-D-glucopyranosyl-α-D-galactopyranose: 0.67 mg mL^{-1} mCherry *Tu*TreT, isolated yield 19.3 mg (37%) of white crystals. **^{1}H-NMR** (400 MHz, D_2O) = δ 3.45 (t, $J_{3,4\text{-}\alpha\text{-}D\text{-}glc}$ = $J_{4,5\text{-}\alpha\text{-}D\text{-}glc}$ = 9.2 Hz, $H4_{\alpha\text{-D-gluc}}$, 1H), 3.65 (dd, $J_{H1\text{-}2\text{-}\alpha\text{-}D\text{-}glc}$ = 4.0 Hz, $J_{H2\text{-}3\text{-}\alpha\text{-}D\text{-}glc}$ = 10.4 Hz, $H2_{\alpha\text{-D-glc}}$, 1H), 3.73 – 3.93 (m, $H2_{\alpha\text{-D-gal}}$, $H5_{\alpha\text{-D-gal}}$, $H3_{\alpha\text{-D-glc}}$, $H6a_{\alpha\text{-D-glc}}$, $H6b_{\alpha\text{-D-glc}}$, $H6a_{\alpha\text{-glc}}$, $H6b_{\alpha\text{-D-glc}}$, 7H), 3.99 – 4.04 (m, $H3_{\alpha\text{-D-gal}}$, $H4_{\alpha\text{-D-gal}}$, 2H), 4.08 (t, $J_{4,5\text{-}\alpha\text{-}D\text{-}gal}$ = $J_{5,6a\text{-}\alpha\text{-}D\text{-}gal}$ = 6.0 Hz, $H5_{\alpha\text{-D-gal}}$, 1H), 5.20 (d, $J_{1,2\text{-}\alpha\text{-}D\text{-}glc}$ = 4.0 Hz, 1H), 5.22 (d, $J_{1,2\text{-}\alpha\text{-}D\text{-}gal}$ = 5.6 Hz, 1H). **^{13}C-NMR** (101 MHz, D_2O) = δ 93.2 ($C1_{\alpha\text{-D-gal}}$, C*H*), 93.1 ($C1_{\alpha\text{-D-glc}}$, C*H*), 72.5 ($C3_{\alpha\text{-D-glc}}$, C*H*), 72.0 ($C5_{\alpha\text{-D-glc}}$, C*H*CH_2), 71.2 ($C5_{\alpha\text{-D-gal}}$, C*H*CH_2), 71.0 ($C2_{\alpha\text{-D-glc}}$, C*H*), 69.5 ($C4_{\alpha\text{-D-glc}}$,C*H*), 69.2 ($C4_{\alpha\text{-D-gal}}$, C*H*), 68.9 ($C3_{\alpha\text{-D-gal}}$, C*H*), 67.9 ($C2_{\alpha\text{-D-gal}}$, C*H*), 61.1 ($C6_{\alpha\text{-D-gal}}$, CHCH_2), 60.5 ($C6_{\alpha\text{-D-glc}}$, CHCH_2). **HRMS** (ESI): 365.1049 (calc. 365.1054 [M+Na]).
These ^{13}C- and ^{1}H-NMR signals matched the literature values:[22] **^{1}H-NMR** (500 MHz, D_2O) δ 3.31 (app t, *J* = 9.3 Hz, 1H), 3.51 (dd, *J* = 3.5, 9.9 Hz, 1H), 3.60-3.64 (m, 3H), 3.67-3.78 (m, 4 H), 3.87 (m, 2H), 3.93 (app t, *J* = 6.0 Hz, 1H), 5.07 (d, *J* = 3.8 Hz, 1H), 5.08 (d, *J* = 3.9 Hz, 1H). **^{13}C-NMR** (125 MHz, D_2O) δ 93.18, 92.98, 72.38, 71.97, 71.20, 70.93, 69.54, 69.13, 68.81, 67.79, 61.08, 60.37.



α-D-glucopyranosyl-α-D-xylopyranoside: 0.33 mg mL^{-1} mCherry *Tu*TreT, isolated yield 17.2 mg (37%) of white crystals. **^{1}H-NMR** (400 MHz, D_2O) = δ 3.31 (t, *J* = 9.4 Hz, $H4_{\alpha\text{-D-gluc}}$, 1H), 3.48-3.74 (m, $H2_{\alpha\text{-D-xyl}}$, $H2_{\alpha\text{-D-glc}}$, $H4_{\alpha\text{-D-xyl}}$, $H5_{\alpha\text{-D-xyl}}$, 2H$6_{\alpha\text{-D-glc}}$, $H5_{\alpha\text{-D-glc}}$, $H3_{\alpha\text{-D-glc}}$, $H3_{\alpha\text{-D-xyl}}$, 9H), 5.02 (d, *J* = 3.6 Hz, $H1_{\alpha\text{-D-xyl}}$, $H1_{\alpha\text{-D-glc}}$, 2H). **^{13}C-NMR** (101 MHz, D_2O) = δ 93.4 ($C1_{\alpha\text{-D-glc}}$, C*H*), 93.3 ($C1_{\alpha\text{-D-xyl}}$, C*H*), 72.6 ($C3_{\alpha\text{-D-xyl}}$, C*H*), 72.5 ($C3_{\alpha\text{-D-glc}}$, C*H*), 72.1 ($C5_{\alpha\text{-D-glc}}$, C*H*CH_2), 71.0

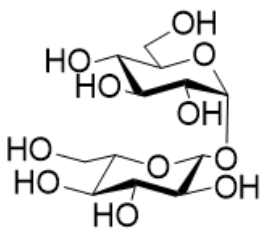
(C2$_{\alpha\text{-D-glc}}$, C*H*), 70.98 (C2$_{\alpha\text{-D-xyl}}$, C*H*), 69.6 (C4$_{\alpha\text{-D-glc}}$, C*H*), 69.4 (C4$_{\alpha\text{-D-xyl}}$, C*H*), 61.5 (C5$_{\alpha\text{-D-xyl}}$, CH_2), 60.4 (C6$_{\alpha\text{-D-glc}}$, CHCH_2). **HRMS** (ESI): 335.0944 (calc. 335.0949 [M+Na]).

These ^{13}C- and ^{1}H-NMR signals matched the literature values:[22] **^{1}H-NMR** (500 MHz, D_2O) δ 3.32 (app t, *J* = 9.0 Hz, 1H), 3.52-3.74 (m, 10H), 5.02 (m, 2H). **^{13}C-NMR** (125 MHz, D_2O) δ 93.33, 93.19, 72.53, 72.38, 72.01, 70.94, 70.91, 69.53, 69.45, 69.35, 61.41, 60.34.



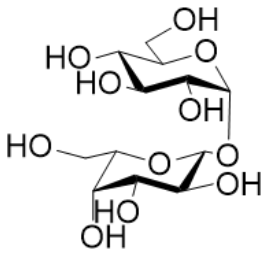
α-D-glucopyranosyl-*N*-Ac-α-D-glucopyranoside. 0.50 mg mL^{-1} mCherry *Tu*TreT, isolated yield 18.7 mg (33%) of white crystals. **^{1}H-NMR** (400 MHz, D_2O) = δ 2.01 (s, 3H, CH_3), 3.44 (t, $J_{3,4\text{-}\alpha\text{-D-glc}}$ = $J_{4,5\text{-}\alpha\text{-D-glc}}$ = 8.8 Hz, H4$_{\alpha\text{-D-glc}}$), 3.50-3.54 (m, 2H, H4$_{\alpha\text{-D-glcNAc}}$, H5$_{\alpha\text{-D-glc}}$), 3.62 (dd, $J_{1,2\text{-}\alpha\text{-D-glc}}$ = 4.8 Hz, $J_{2,3\text{-}\alpha\text{-D-glc}}$ = 10 Hz, H2$_{\alpha\text{-D-glc}}$), 3.73 - 3.77 (m, 4H, H3$_{\alpha\text{-D-glc}}$, H6$_{\alpha\text{-D-glcNAc}}$, H6a$_{\alpha\text{-D-glc}}$, H6b$_{\alpha\text{-D-glc}}$), 3.80 – 3.87 (m, 3H, H3$_{\alpha\text{-D-glcNAc}}$, H5$_{\alpha\text{-D-glcNAc}}$, H6$_{\alpha\text{-D-glcNAc}}$), 3.93 (dd, $J_{1,2\text{-}\alpha\text{-D-glcNAc}}$ = 4.0 Hz, $J_{1,2\text{-}\alpha\text{-D-glcNAc}}$ 11.0 Hz, H2$_{\alpha\text{-D-GlcNAc}}$), 5.13 (d, $J_{1,2\text{-}\alpha\text{-}D\text{-}glc}$ = 4.0 Hz, H1$_{\alpha\text{-D-glc}}$),5.17 (d, $J_{1,2\text{-}\alpha\text{-}D\text{-}glcNAc}$ = 3.6 Hz, H1$_{\alpha\text{-D-GlcNAc}}$). **^{13}C-NMR** (101 MHz, D_2O) = δ 173.6 (C7$_{\alpha\text{-D-glcNAc}}$, *C*=O), 92.8 (C1$_{\alpha\text{-D-glc}}$, *C*H), 91.2 (C1$_{\alpha\text{-D-glcNAc}}$, *C*H), 72.0 (C5$_{\alpha\text{-D-glc}}$, *C*H), 71.9 (C3$_{\alpha\text{-D-glc}}$, *C*H), 71.6 (C5$_{\alpha\text{-D-glcNAc}}$, *C*H), 70.2 (C2$_{\alpha\text{-D-glc}}$, *C*H), 69.4 (2C, C3$_{\alpha\text{-D-glcNAc}}$, C4$_{\alpha\text{-D-glcNAc}}$), 68.7 (C4$_{\alpha\text{-D-glcNAc}}$, *C*H), 59.8 (C6$_{\alpha\text{-D-glcNAc}}$, *C*$H_2$), 59.5 (C6$_{\alpha\text{-D-glc}}$, *C*$H_2$), 52.8 (C2$_{\alpha\text{-D-glcNAc}}$, *C*HNH),), 21.1 (C8$_{\alpha\text{-D-glcNAc}}$, *C*$H_3$). **HRMS** (ESI): 384.1493 (calc. 384.1500 [M+H]).

These ^{13}C- and ^{1}H-NMR signals matched the literature values:[8] **^{1}H-NMR** (500 MHz, D_2O) δ 5.16 (d, *J* = 2.5 Hz, 1 H, H1′), 5.13 (d, *J* = 4.0 Hz, 1 H, H1), 3.93 (dd, *J* = 4.0, 11 Hz, 1H, H2′) 3.89–3.80 (m, 3 H, H3′, H5′, 1 × H6′), 3.77–3.73 (m, 4 H, 3H, 1 × H6′, 2 × H6), 3.62 (dd *J* = 4.0, 10.5 Hz, 1 H, H2), 3.54–3.51 (m, 1 H, H5), 3.49 (t, *J* = 10 Hz, 1 H, H4′), 3.44 (t, *J* = 9.5 Hz, 1 H, H4), 2.01 (s, 3 H, CH_3). **^{13}C NMR** (125 MHz, D_2O): δ 174.2 (C═O), 93.4 (C1), 91.7 (C1′), 72.6 (C5), 72.5 (C3), 72.2 (C5′), 70.8 (C2), 70.00 (C3′ and C4′ overlapping), 69.2 (C4), 60.4 (C6′), 60.0 (C6), 53.4 (C2′), 21.6 (CH_3).

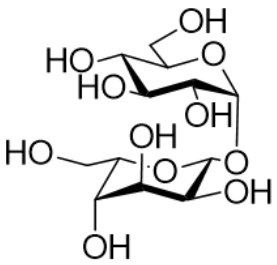


α-D-glucopyranosyl-β-L-glucopyranose: 0.20 mg mL^{-1} mCherry *Tu*TreT, isolated yield 23.3 mg (45%) of white crystals. **^{1}H-NMR** (400 MHz, D_2O): δ 3.39 (t, H2$_{\beta\text{-L-glc}}$, $J_{1,2}$ = $J_{2,3}$ = 8.6 Hz, 1H), 3.56 – 3.44 (m, H4$_{\alpha\text{-D-glc}}$, H5$_{\beta\text{-L-glc}}$, H4$_{\beta\text{-L-glc}}$, H3$_{\beta\text{-L-glc}}$, H4$_{\alpha\text{-D-glc}}$, 5H), 3.63 (dd, H2$_{\alpha\text{-D-glc}}$, $J_{1,2\text{-}\alpha\text{-}D\text{-}glc}$ = 3.6 Hz, $J_{2,3\text{-}\alpha\text{-}D\text{-}glc}$ = 10.0 Hz, 1H), 3.88 – 3.73 (m, H6a$_{\alpha\text{-D-glc}}$, H6b$_{\alpha\text{-D-glc}}$, H6a$_{\beta\text{-L-glc}}$, H5$_{\alpha\text{-D-glc}}$, H3$_{\alpha\text{-D-glc}}$, 5H), 3.93 (dd, H6b$_{\beta\text{-}}$

L-glc, $J_{6a,6b\text{-}\beta\text{-}L\text{-}glc}$ = 1.0 Hz, $J_{6a,5\text{-}\beta\text{-}L\text{-}glc}$ = 11.6 Hz, 1H), 4.69 (d, $J_{1,2\text{-}\beta\text{-}L\text{-}glc}$ = 8.0 Hz, , $H1_{\beta\text{-}L\text{-}glc}$, 1H), 5.35 (d, $J_{1,2\text{-}\alpha\text{-}D\text{-}Glc}$ = 3.6 Hz, $H1_{\alpha\text{-}D\text{-}glc}$, 1H). **^{13}C-NMR** (101 MHz, D_2O) = δ 97.2 ($C1_{\beta\text{-}L\text{-}glc}$), 94.8 ($C1_{\alpha\text{-}D\text{-}glc}$), 75.9 ($C3_{\beta\text{-}L\text{-}glc}$), 75.6 ($C5_{\beta\text{-}L\text{-}glc}$), 72.7 ($C2_{\beta\text{-}L\text{-}glc}$), 72.6 ($C3_{\alpha\text{-}D\text{-}glc}$), 72.1 ($C5_{\alpha\text{-}D\text{-}glc}$), 70.7 ($C2_{\alpha\text{-}D\text{-}glc}$), 69.6 ($C4_{\alpha\text{-}D\text{-}glc}$), 69.3 ($C4_{\beta\text{-}L\text{-}glc}$), 60.5 ($C6_{\beta\text{-}L\text{-}glc}$), 60.3 ($C6_{\alpha\text{-}D\text{-}glc}$). **MS** (ESI-TOF): 365.1051 (calc. 365.1054 [M+Na]).
^{1}H- and ^{13}C-NMR annotation of reported β-L-glucopyranose moiety was in agreement.[23]



α-D-glucopyranosyl-β-L-galactopyranose: 0.33 mg mL^{-1} mCherry *Tu*TreT, isolated yield 30.6 mg (60%) of white crystals**. ^{1}H-NMR** (400 MHz, D_2O) = δ 3.48 (t, $J_{3,4}$ =$J_{4,5}$ = 9.4 Hz, $H4_{\alpha\text{-}D\text{-}glc}$, 1H), 3.59 – 3.64 (m, $H2_{\beta\text{-}L\text{-}gal}$, $H2_{\alpha\text{-}D\text{-}glc}$, 2H),), 3.73 – 3.67 (m, $H5_{\beta\text{-}L\text{-}gal}$, $H3_{\beta\text{-}L\text{-}gal}$, $H2_{\beta\text{-}L\text{-}gal}$, $H2_{\alpha\text{-}D\text{-}glc}$, 4H), 3.89 – 3.75 (m, $H6ab_{\alpha\text{-}D\text{-}glc}$, $H6ab_{\beta\text{-}L\text{-}gal}$, $H5_{\alpha\text{-}D\text{-}glc}$, $H3_{\alpha\text{-}D\text{-}glc}$, 6H3.95 (app d, J = 3.2 Hz, $H4_{\beta\text{-}L\text{-}gal}$, 1H), 4.63 (d, $J_{1,2}$ = 8.0 Hz, $H1_{\beta\text{-}L\text{-}gal}$, 1H), 5.37 (d, $J_{1,2}$ = 3.6, $H1_{\alpha\text{-}D\text{-}glc}$, 1H). **^{13}C-NMR** (101 MHz, D_2O) = δ 97.8 ($C1_{\beta\text{-}L\text{-}gal}$, C*H*), 94.7 ($C1_{\alpha\text{-}D\text{-}glc}$, C*H*), 75.3 ($C5_{\beta\text{-}L\text{-}gal}$, C*H*$CH_2$), 72.7 ($C3_{\beta\text{-}L\text{-}gal}$, C*H*), 72.6 ($C3_{\alpha\text{-}D\text{-}glc}$, C*H*$CH_2$), 72.0 ($C5_{\alpha\text{-}D\text{-}glc}$, C*H*$CH_2$), 70.8 ($C2_{\beta\text{-}L\text{-}gal}$, C*H*), 70.3 ($C2_{\alpha\text{-}D\text{-}glc}$, C*H*), 69.3 ($C4_{\alpha\text{-}D\text{-}glc}$, C*H*, 68.6 ($C4_{\beta\text{-}D\text{-}gal}$, C*H*), 61.0 ($C6_{\beta\text{-}L\text{-}gal}$, CHC$H_2$), 60.4 ($C6_{\alpha\text{-}L\text{-}glc}$, CHC$H_2$). **HRMS** (ESI): 365.1048 (calc. 365.1054 [M+Na]).
^{1}H- and ^{13}C-NMR annotation of reported β-L-galactopyranose moiety was in agreement[24].



α-D-glucopyranosyl-β-L-gulopyranose: 0.67 mg mL^{-1} mCherry *Tu*TreT, isolated yield 20.6 mg (40%) of white crystals. **^{1}H-NMR** (400 MHz, D_2O) = δ 3.47 (app t, $J_{3,4\text{-}\alpha\text{-}D\text{-}glc}$=$J_{4,5\text{-}\alpha\text{-}D\text{-}glc}$= 9.6 Hz, $H4_{\alpha\text{-}D\text{-}glc}$, 1H), 3.63 (dd, $J_{1,2\text{-}\alpha\text{-}D\text{-}glc}$ = 3.6 Hz, $J_{2,3\text{-}\alpha\text{-}D\text{-}glc}$ = 10.0 Hz, $H2_{\alpha\text{-}D\text{-}glc}$, 1H), 3.76-3.90 (m, $2H6_{\beta\text{-}L\text{-}gul}$, $H2_{\beta\text{-}L\text{-}gul}$, $H5_{\alpha\text{-}D\text{-}glc}$, $H4_{\beta\text{-}L\text{-}gul}$, $H3_{\alpha\text{-}D\text{-}glc}$, $2H6_{\alpha\text{-}D\text{-}glc}$, 8H), 4.01 (dd, $J_{5,6a\text{-}\beta\text{-}L\text{-}gul}$ = 4.8 Hz, $J_{5,6b\text{-}\beta\text{-}L\text{-}gul}$ = 6.8 Hz, $H5_{\beta\text{-}L\text{-}gul}$, 1H), 4.10 (app t, $J_{2,3\text{-}\beta\text{-}L\text{-}gul}$ = $J_{3,4\text{-}\beta\text{-}L\text{-}gul}$ = 3.6 Hz, $H3_{\beta\text{-}L\text{-}gul}$, 1H), 4.94 (d, $J_{1,2\text{-}\beta\text{-}L\text{-}gul}$ = 8.4 Hz, $H1_{\beta\text{-}L\text{-}gul}$, 1H), 5.38 (d, $J_{1,2\text{-}\alpha\text{-}D\text{-}glc}$ = 3.6 Hz, $H1_{\alpha\text{-}D\text{-}glc}$, 1H). **^{13}C-NMR** (101 MHz, D_2O) = δ 95.6 ($C1_{\beta\text{-}L\text{-}gul}$, C*H*), 94.6 ($C1_{\alpha\text{-}D\text{-}glc}$, C*H*), 73.9 ($C5_{\beta\text{-}L\text{-}gul}$, C*H*$CH_2$), 72.7 ($C3_{\alpha\text{-}D\text{-}glc}$, C*H*), 72.0 ($C5_{\alpha\text{-}D\text{-}glc}$, C*H*$CH_2$), 71.2 ($C3_{\beta\text{-}L\text{-}gul}$, C*H*), 70.8 ($C2_{\alpha\text{-}D\text{-}glc}$, C*H*), 69.4 ($C4_{\alpha\text{-}D\text{-}glc}$, C*H*), 69.4 ($C4_{\beta\text{-}L\text{-}gul}$, C*H*), 67.4 ($C2_{\beta\text{-}L\text{-}gul}$, C*H*), 61.0 ($C6_{\beta\text{-}L\text{-}gul}$, CHC$H_2$), 60.4 ($C6_{\alpha\text{-}D\text{-}glc}$, CHC$H_2$). **HRMS** (ESI): 365.1049 (calc. 365.1054 [M+Na]).

5.4.18 Crystal X-ray data collection and refinement

X-ray diffraction data was collected at beamline P13 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany). *Tu*TreT crystallizes in the orthorhombic space group ($P2_12_12_1$), with one molecule per asymmetric unit and a solvent content of *ca.* 53%. Data were indexed and integrated with XDS,[25] scaled with AIMLESS and the space and group determined with POINTLESS[26] from CCP4 programs suite.[27] Data-collection statistics are given in Table 2. The structure of the apo-enzyme was solved by the Molecular Replacement Method using the program MOLREP,[28] and the apo-form of TreT from *Pyrococcus horikoshii* (PDB: 2X6Q)[10] as a search model. After a few cycles of refinement and model building, this structural model was used as a search model to solve the structure of the enzyme in complex with the different ligands. Model building and refinement of all structural models were performed with COOT,[29] and REFMAC5,[30] respectively. The positions of the soaked ligands, UDP-D-glucose, D-trehalose and α-D-glucopyranosyl-β-L-galactopyranoside, became evident after the first cycles of refinement, based on electron density difference maps, and were modeled and refined with full occupancy. Solvent molecules were included in the model after a few rounds of refinement as well as a couple of other molecules that belong to the crystallization solution and were located at the surface of the protein. Structure refinement statistics are listed in Table 2. Figures were drawn with PyMOL.[31]

Table 2: Data collection and refinement statistics.

Parameters	**Apo-*Tu*TreT (PDB: 6ZJ4)**	**Holo-*TuTret* with Mg^{2+} (PDB: 6ZJ7)**	***Tu*TreT with UDP-D-glucose (PDB: 6ZMZ)**	***Tu*TreT with D-trehalose (PDB: 6ZJH)**	***Tu*TreT α-D-glucopyranosyl-β-L-galactopyranoside (PDB: 6ZN1)**
Wavelength (Å)	0.976250	0.976300	0.976200	0.976200	0.97630
Resolution range	42.44 - 2.10 (2.17 - 2.1)	58.69 - 2.15 (2.22 - 2.15)	46.75-1.90 (1.96 - 1.90)	55.39 - 2.1 (2.17 - 2.1)	45.64 -1.75 (1.81-1.75)
Space group	P 2_1 2_1 2_1	P 2_1 2_1 2_1	P 2_1 2_1 2_1	P 2_1 2_1 2_1	P 2_1 2_1 2_1
Unit cell (Å)	61.5 68.7 112.2	60.66 9.3 110.5	53.0 69.3 98.9	60.5 68.4 110.8	61.0 68.8 110.9
Total reflections	99681 (6235)	200673 (20071)	189504 (18996)	170678 (16329)	592975 (53679)
Unique reflections	27351 (2121)	25646 (2477)	29172 (2860)	27410 (2709)	47789 (4717)
Multiplicity	3.6 (2.9)	7.8 (8.1)	6.5 (6.6)	6.2 (6.0)	12.4 (11.4)
Completeness (%)	96.8 (93.6)	98.75 (97.79)	99.03 (98.42)	99.54 (99.30)	99.97 (99.94)
Mean I/sigma (I)	8.5 (1.0)	17.05 (1.75)	15.93 (1.42)	13.45 (1.18)	24.36 (1.47)
Wilson B-factor	40.54	46.65	34.69	45.62	34.46
R-merge	0.065 (0.730)	0.071 (1.164)	0.068 (1.307)	0.078 (1.521)	0.054 (1.568)
CC1/2*	0.99 (0.45)	0.99 (0.70)	0.99 (0.57)	0.99 (0.53)	1.0 (0.69)
Reflections used in refinement	27269 (2537)	25638 (2476)	29170 (2860)	27411 (2705)	47780 (4714)
Reflections used for R-free	1335 (120)	1311 (123)	1445 (136)	1386 (129)	2329 (248)
R-work	0.208 (0.328)	0.203 (0.306)	0.182 (0.286)	0.206 (0.327)	0.207 (0.321)
Rfree	0.255 (0.331)	0.262 (0.337)	0.218 (0.315)	0.223 (0.307)	0.245 (0.364)
Number of non-hydrogen atoms	3224	3239	3407	3266	3486
Protein residues	392	391	398	396	400
ligands	38	12	36	29	72
solvent	140	138	128	142	209
RMS(bonds)	0.014	0.017	0.014	0.003	0.014
RMS(angles)	1.87	2.11	1.78	0.57	1.78
Ramachandran favored (%)	96.36	95.61	96.21	96.40	95.73
Ramachandran allowed (%)	3.38	3.62	3.28	3.08	3.77
Ramachandran outliers (%)	0.59	0.78	0.51	0.51	0.50
Average B-factor	51.1	56.9	42.3	55.7	42.59
macromolecules	51.2	57.1	42.1	55.8	42.37
ligands	42.5	57.8	43.1	59.4	43.08
solvent	52.1	48.6	44.7	51.9	45.87

* Values in parenthesis correspond to the highest resolution shell

References

[1] J. Seibel, R. Moraru, S. Götze, K. Buchholz, S. Na'amnieh, A. Pawlowski, H.-J. Hecht, *Carbohyd. Res.* **2006**, *341*, 2335-2349.

[2] W. A. Bubb, *Concepts in Magnetic Resonance Part A* **2003**, *19A*, 1-19.

[3] a) L. L. Lairson, B. Henrissat, G. J. Davies, S. G. Withers, *Annual Review of Biochemistry* **2008**, *77*, 521-555; b) S. S. Lee, S. Y. Hong, J. C. Errey, A. Izumi, G. J. Davies, B. G. Davis, *Nature Chemical Biology* **2011**, *7*, 631-638; c) A. Ardèvol, C. Rovira, *Angew. Chem. Int. Edit.* **2011**, *50*, 10897-10901.

[4] a) L. Mestrom, S. R. Marsden, M. Dieters, P. Achterberg, L. Stolk, I. Bento, U. Hanefeld, P.-L. Hagedoorn, *Appl. Environ. Microbiol.* **2019**, *85*, e03084-03018; b) L. Mestrom, S. R. Marsden, M. Dieters, P. Achterberg, L. Stolk, I. Bento, U. Hanefeld, P.-L. Hagedoorn, *Appl. Environ. Microbiol.* **2019**, *85*, e00942-00919.

[5] L. Mestrom, S. R. Marsden, D. G. G. McMillan, R. Schoevaart, P.-L. Hagedoorn, U. Hanefeld, *ChemCatChem* **2020**, *12*, 3249-3256.

[6] a) L. Mestrom, M. Przypis, D. Kowalczykiewicz, A. Pollender, A. Kumpf, S. R. Marsden, I. Bento, A. B. Jarzębski, K. Szymańska, A. Chruściel, D. Tischler, R. Schoevaart, U. Hanefeld, P.-L. Hagedoorn, *Int. J. Mol. Sci.* **2019**, *20*, 5263; b) S. R. Marsden, L. Mestrom, D. G. G. McMillan, U. Hanefeld, *ChemCatChem* **2020**, *12*, 426-437.

[7] a) B. L. Urbanek, D. C. Wing, K. S. Haislop, C. J. Hamel, R. Kalscheuer, P. J. Woodruff, B. M. Swarts, *ChemBioChem* **2014**, *15*, 2066-2070; b) H.-M. Kim, Y.-K. Chang, S.-I. Ryu, S.-G. Moon, S.-B. Lee, *J. Mol. Catal. B-Enzym.* **2007**, *49*, 98-103.

[8] J. M. Groenevelt, L. M. Meints, A. I. Stothard, A. W. Poston, T. J. Fiolek, D. H. Finocchietti, V. M. Mulholand, P. J. Woodruff, B. M. Swarts, *J. Org. Chem.* **2018**, *83*, 8662-8667.

[9] R. P. Gibson, J. P. Turkenburg, S. J. Charnock, R. Lloyd, G. J. Davies, *Chem. Biol.* **2002**, *9*, 1337-1346.

[10] E.-J. Woo, S.-I. Ryu, H.-N. Song, T.-Y. Jung, S.-M. Yeon, H.-A. Lee, B. C. Park, K.-H. Park, S.-B. Lee, *J. Mol. Biol.* **2010**, *404*, 247-259.

[11] C. E. Blanchet, A. Spilotros, F. Schwemmer, M. A. Graewert, A. Kikhney, C. M. Jeffries, D. Franke, D. Mark, R. Zengerle, F. Cipriani, S. Fiedler, M. Roessle, D. I. Svergun, *J. Appl. Crystallogr.* **2015**, *48*, 431-443.

[12] a) S.-I. Ryu, J.-E. Kim, E.-J. Kim, S.-K. Chung, S.-B. Lee, *Process Biochem.* **2011**, *46*, 128-134; b) Q. Qu, S.-J. Lee, W. Boos, *J. Biol. Chem.* **2004**, *279*, 47890-47897.

[13] L. Mestrom, S. R. Marsden, M. Dieters, P. Achterberg, L. Stolk, I. Bento, U. Hanefeld, P.-L. Hagedoorn, *Appl. Environ. Microbiol.* **2019**, AEM.03084-03018.

[14] a) M. A. Graewert, D. Franke, C. M. Jeffries, C. E. Blanchet, D. Ruskule, K. Kuhle, A. Flieger, B. Schäfer, B. Tartsch, R. Meijers, D. I. Svergun, *Sci. Rep.* **2015**, *5*, 10734; b) D. Franke, A. G. Kikhney, D. I. Svergun, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **2012**, *689*, 52-59.

[15] A. Panjkovich, D. I. Svergun, *Bioinformatics* **2017**, *34*, 1944-1946.

[16] D. Franke, M. V. Petoukhov, P. V. Konarev, A. Panjkovich, A. Tuukkanen, H. D. T. Mertens, A. G. Kikhney, N. R. Hajizadeh, J. M. Franklin, C. M. Jeffries, D. I. Svergun, *J. Appl. Crystallogr.* **2017**, *50*, 1212-1225.

[17] A. Guinier, *Ann. Phys.* **1939**, *11*, 161-237.

[18] D. Svergun, *J. Appl. Crystallogr.* **1992**, *25*, 495-503.

[19] D. Franke, C. M. Jeffries, D. I. Svergun, *Nat. Methods* **2015**, *12*, 419-422.

[20] F. L. Lin, *Carbohyd. Res.* **2007**, *342*, 2014.

[21] a) É. Bar-Guilloux, *Carbohyd. Res.* **1975**, *45*, 217; b) K. Bock, J. Defaye, H. Driguez, E. Barguilloux, *Eur. J. Biochem.* **1983**, *131*, 595-600.

[22] M. R. Pratt, *Org. Lett.* **2003**, *5*, 3185.

[23] A. Szekrenyi, X. Garrabou, T. Parella, J. Joglar, J. Bujons, P. Clapés, *Nat. Chem.* **2015**, *7*, 724-729.
[24] R. Orii, M. Izumi, Y. Kajihara, R. Okamoto, *J. Carbohyd. Chem.* **2015**, *34*, 560-566.
[25] W. Kabsch, *Acta Crystallogr. D* **2010**, *66*, 125-132.
[26] P. Evans, *Acta Crystallogr. D* **2011**, *67*, 282-292.
[27] Collaborative, *Acta Crystallogr. D* **1994**, *50*, 760-763.
[28] A. Vagin, A. Teplyakov, *J. Appl. Crystallogr.* **1997**, *30*, 1022-1025.
[29] P. Emsley, K. Cowtan, *Acta Crystallogr. D* **2004**, *60*, 2126-2132.
[30] G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallogr. D* **2011**, *67*, 355-367.
[31] Schrödinger, *LLC* **2015**.

6

Enzyme Catalyzed Synthesis of Esters in Water

This chapter is based on:

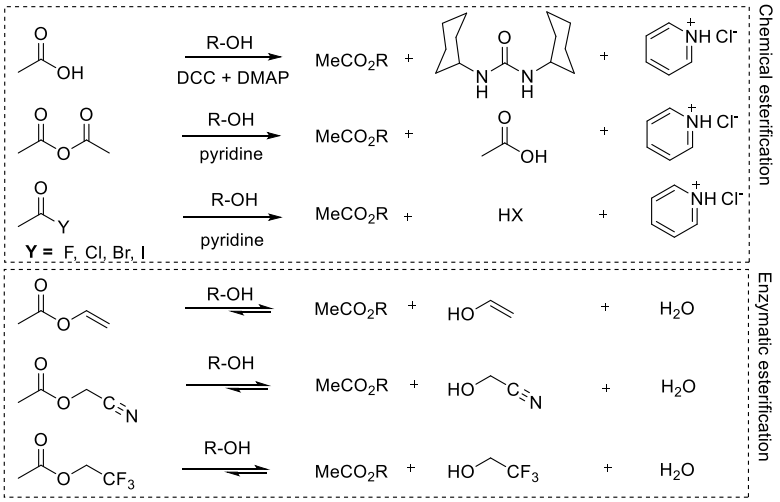
Luuk Mestrom, Jord G. R. Claessen, Ulf Hanefeld.

ChemCatChem, **2019**, *11*, 2004-20.

Pressure pushing down on me
Pressing down on you, no man ask for
Under pressure that burns a building down
Splits a family in two
Puts people on streets
Queen & David Bowie, “Hot Space”, Under Pressure, **1981**

6.1 Introduction

The synthesis of esters is textbook-knowledge and is performed according to standard protocols that are taught unaltered already for many decades.[1] An activated acid, i.e. an acid chloride or anhydride, reacts with an alcohol in the presence of a base, typically pyridine, often catalyzed by DMAP, in an organic solvent. Alternatively, the acid is activated *in situ* with reagents such as DCC. In all cases considerable amounts of waste are generated: the solvent and stoichiometric amount of salt from neutralization steps. Furthermore the activation, *in situ* or not, also generates considerable amounts of waste. The direct esterification catalyzed by an acid requires azeotropic removal of water to shift the equilibrium towards ester formation and the reaction conditions are so drastic that it is limited to stable starting materials and products. Moreover it still leaves the problem of the organic solvent and the neutralization steps generating salt waste.[1a-c] Similarly the enzyme catalyzed synthesis of esters is performed in dry organic solvents with activated acids (acyl donors), often generating much waste which needs to be removed in additional work up steps (Scheme 1).[2] Since the activated acids in the enzyme catalyzed ester synthesis are esters themselves, these reactions are transesterifications.



Scheme 1. Textbook chemical and enzymatic synthesis strategies for esters.

To address the problem the acyltransferase from *Mycobacterium smegmatis* (MsAcT) was employed as catalyst.[3] This enzyme was described to enable the synthesis the esters in water. With MsAcT it should thus be possible to avoid the use of organic solvents and to eradicate the need for neutralization steps commonly employed to remove the base utilized in ester syntheses. In all cases described to date, the ester synthesis catalyzed by MsAcT is strictly speaking a transesterification since the reactions all utilize ethyl acetate or more reactive acid derivatives.[3] While this is similar to the classical ester synthesis which utilizes acid chlorides or anhydrides it raises the question whether MsAcT catalyzes the direct synthesis of esters, too. This has been described for lipases and other catalysts;[4] but in all those examples the yield was determined by the thermodynamic stability of the product under reaction conditions. Therefore, the direct esterification and transesterification of alcohols catalyzed by MsAcT in water was studied in parallel. As a catalyst MsAcT should accelerate the reaction but not alter its overall equilibrium.[5] High transesterification yields with MsAcT can only be obtained under kinetic control using activated acids.

MsAcT is known to be an excellent catalyst for the transesterification of primary alcohols and amines with a variety of acyl donors[3d, 6]. Benzyl alcohol and isobutyl alcohol were chosen for this study in combination with acetic acid in the form of potassium acetate buffer or ethyl-, vinyl- and phenyl-acetate as activated acids. The kinetic properties of MsAcT were investigated for both the hydrolysis and transesterification rate of vinyl acetate. This acyl donor irreversibly tautomerizes upon hydrolysis, simplifying the reaction kinetics of the transesterification.

6.2 Results and Discussion

To establish the thermodynamic equilibrium of the reaction and to probe whether MsAcT had the potential to overcome the macroscopic thermodynamic forces of the reaction[3b] the direct esterification of benzyl alcohol with acetic acid at pH 7.5 in potassium phosphate buffer (KPi) was studied. Investigation of both the

synthesis (10 mM benzyl alcohol) and hydrolysis (10 mM benzyl acetate) reaction catalyzed by MsAcT revealed a final concentration of benzyl acetate of 0.12 mM. A K_{eq} of 0.00012 could thus be established. The uncatalyzed control reactions yielded essentially no product. MsAcT is thus a classical catalyst, catalysing a reaction but not altering it.

To prepare the desired ester it is therefore necessary to employ activated acids. As was demonstrated earlier this can be simple ethyl acetate, a benign solvent that can be used monophasic, i.e. dissolved in water, or as a separate layer.[3] Here, all studies were performed under monophasic conditions with the acyl donor dissolved in the pH 7.5 KPi buffer. In order to characterize the affinity and activity of activated acyl donor ethyl, vinyl and phenyl acetate to MsAcT, the optimal enzyme concentration was investigated. At less than 50 ng mL^{-1} MsAcT the conversion of benzyl alcohol to benzyl acetate was not complete. In line with earlier observations,[3d] a significant amount of benzyl acetate was hydrolyzed again at high enzyme concentrations (Fig. 1). The trend in conversion of vinyl acetate > phenyl > ethyl acetate after one hour of reaction time was in line with the reactivity of the acyl donor.[2a,b] Vinyl acetate clearly gave the highest conversion for ester synthesis. Moreover it is synthesized via an atom efficient catalytic process from acetic acid and ethylene making its production environmentally benign.[7] Additionally it is very readily available as it is a monomer for polymer synthesis. Also the side product acetaldehyde does on laboratory scale not require special work up steps. On industrial scale measures to handle it are well established.[8] Therefore, this activated acyl donor is relatively environmentally benign. It does not require neutralization steps in down-stream processing nor are toxic compounds produced at a large scale. The characterization of this acyl donor was pursued in more detail in water to also avoid organic solvents necessary when used in combination with lipases.

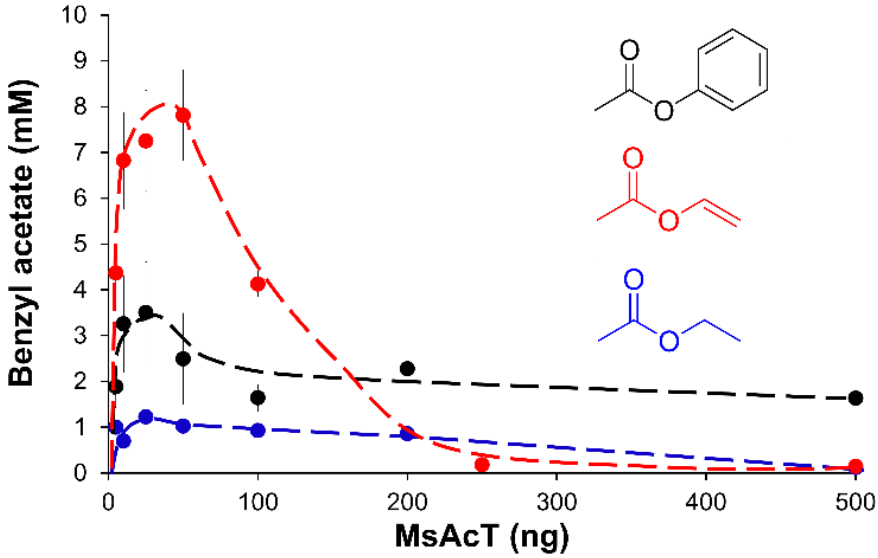


Figure 1. Transesterification of benzyl alcohol with the acyl donor vinyl acetate (red), phenyl acetate (black), and ethyl acetate (blue) using different MsAcT concentrations. **Reaction conditions**: KPi buffer (200 mM, pH 7.5), 10 mM benzyl alcohol, 100 mM acyl donor, 5 – 500 ng MsAcT, 1 mL, 1000 rpm, 1 hour, 21°C. The experiment was performed in triplicates and the error bars show the standard deviation.

To firmly establish all possible products and side products, such as hemiacetal esters, *in-situ* ^{1}H-NMR analysis was performed. Since transesterification of benzyl alcohol and vinyl acetate could not be followed via *in-situ* ^{1}H-NMR due to overlapping NMR signals, the acylation of 11 mM isobutanol with 212 mM vinyl acetate with less than a µg mL^{-1} of MsAcT was studied under monophasic conditions. To keep an excess of the acyl donor, a lower substrate concentration was applied. Controls with CAL-A as well as without catalyst were carried out under the same reaction conditions. Quantitative conversion was monitored by in situ NMR. While the purified MsAcT-catalysed reaction went to completion, no conversion could be observed with CAL-A nor for the negative control.

Next, the substrate concentration was increased to 97 mM isobutanol and 20-fold more MsAcT was added. Under these conditions the starting materials are all soluble but the product will be insoluble in water. During the course of the reaction

a second organic phase appeared due to the insolubility of isobutyl acetate (Fig. 2). No hemiacetal ester was observed. After prolonged reaction time the reaction mixture became clear again demonstrating the hydrolysis of isobutyl acetate. This occurred after all acyl donor had been consumed, indeed no isobutyl acetate was detectable after 18 hours of reaction time at this stage of the reaction. MsAcT catalyzed only the hydrolysis of the product ester, leading to the final reaction equilibrium; that of acid, alcohol, ester and water.

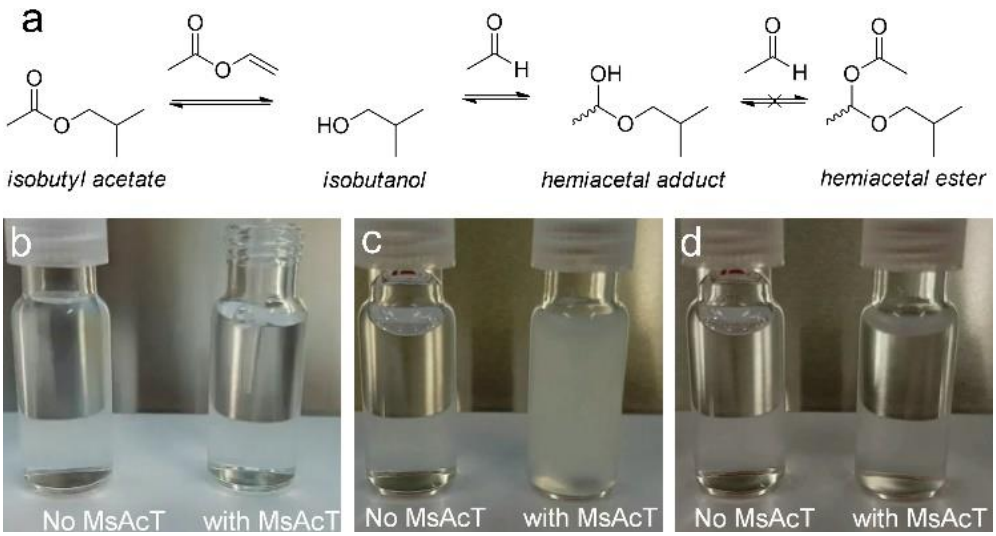
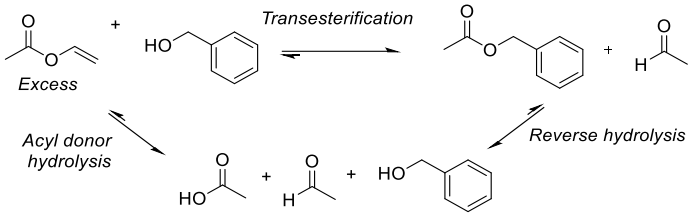


Figure 2. Transesterification of isobutyl alcohol with MsAcT resulting in the reversible formation of an emulsion. a) desired reaction and potential side reaction; b) initial reaction mixture at t=0; c) reaction mixture at t=5 min; d) reaction mixture at t=60min.

Having ruled out hemiacetate esters as side product and established the transient formation of the desired ester, optimal reaction conditions were established. The concentration of vinyl acetate was varied at constant benzyl alcohol and MsAcT concentration (Scheme 2 and Fig. 3). At a 1:1 ratio of benzyl alcohol:vinyl acetate the benzyl ester is quickly formed and hydrolyzed; full conversion is never reached (Fig. 3a). However, at 1:10 benzyl alcohol:vinyl

acetate excellent conversions are obtained with a prolonged stability of the synthesized ester (Fig. 3b). Indeed, by increasing the acyl donor concentration the yield is increased and hydrolysis is outcompeted until all acyl donor is consumed.



Scheme 2: MsAcT catalyzed transesterification of benzyl alcohol with vinyl acetate.

The transesterification of alcohols with activated acids catalyzed by MsAcT follows overall a synthesis/hydrolysis pattern as is well established for the amide bond synthesis.[9] Initially the alcohol is the preferred substrate of MsAcT and the desired ester, the kinetic product, is formed. In parallel both the activated acid and the ester are also subject to hydrolysis. Once all acyl donor is consumed the hydrolysis reaction leads to the final reaction equilibrium and thermodynamics dictate the ester yield (Scheme 2 and Fig. 5).

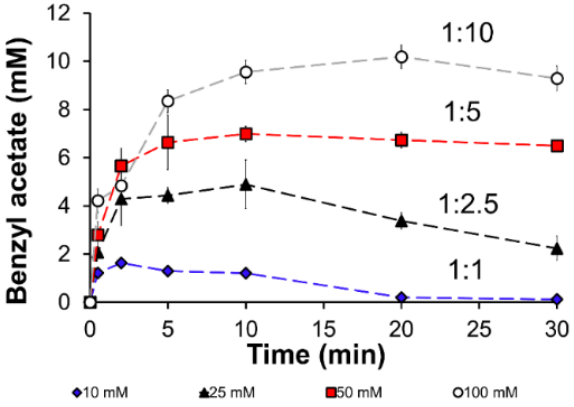


Figure 3. The synthesis of an ester in water. In a) the synthesis of an ester under kinetic control followed by reverse hydrolysis into the thermodynamic equilibrium; b) solely kinetic

control is demonstrated by a 10-fold excess of vinyl acetate; c) one can observe the decrease in reverse hydrolysis by 1, 2.5-, 5-, and 10-fold excess of acyl donor. Reaction conditions: KPi buffer (200 mM, pH 7.5), 10 mM benzyl alcohol, 0 - 100 mM acyl donor, 50 ng MsAcT, 1 mL, 1000 rpm, 21°C. The error bars show the standard deviation of triplicates.

After the synthesis of esters in water with different acyl donors or enzyme concentrations the kinetic parameters of MsAcT catalyzed transesterification in water was investigated. To firmly establish the transesterification of benzyl alcohol with vinyl acetate the ester formation was monitored using GC analysis for the determination of the Michaelis-Menten constants for two-substrate reactions[10]. The MsAcT-catalyzed transesterification of benzyl alcohol with vinyl acetate resulted for benzyl alcohol a K_m of 9.1 mM ± 2.8 mM and k_{cat} of 3.16 x 10^3 s^{-1}, demonstrating high transesterification efficiencies. More interestingly, the transesterification was revealing a K_M and k_{cat} of 30.1 mM ± 6.3 mM and 2.49 x 10^3 s^{-1} respectively.

The formation of acetaldehyde via the hydrolysis of vinyl acetate was monitored using a coupled spectrophotometric activity assay, as is shown in table 1. The hydrolysis of vinyl acetate with MsAcT showed high affinity to the acyl donor vinyl acetate, with moderate catalytic efficiency. The transesterification of benzyl alcohol with vinyl acetate was repeated using the spectrophotometric assay to ensure saturation conditions for two-substrate reactions of vinyl acetate with initial rate conditions[11], were a >100-fold higher k_{cat} was observed for the transesterification over hydrolysis. The ratio between the catalytic rate constants of the hydrolysis and synthesis has been utilized to classify acyltransferase efficiencies,[12] indicating that MsAcT classifies as an efficient acyltransferase in water. The affinity for vinyl acetate is high during the hydrolysis towards acetaldehyde, but during transesterification in the presence of benzyl alcohol the apparent affinity is significantly reduced indicating that competitive inhibition occurs. Taken together, MsAcT favors transesterification over hydrolysis towards benzyl acetate under kinetic control while hydrolyzing benzyl acetate to its alcohol under thermodynamic control.

Table 1. Kinetic parameters of MsAcT catalyzed hydrolysis and esterification in water

Compound[a]	K_i (mM)	K_M (mM)	k_{cat} (s^{-1})†
Vinyl acetate[a]	n.a.*	0.0123 ± 0.0012	182 ± 2.5
Benzyl acetate[b]	2.28 ± 0.42	0.0308 ± 0.0087	52.7 ± 3.5
Benzyl alcohol[c]	n.a.*	13.6 ± 1.7	2586 ± 2.4

†k_{cat} is based on the molecular weight of MsAcT (23.3 kDa).

**n.a.: not applicable

[a] Assay conditions: Vinyl acetate (0 – 10 mM), NADH (0.3 mM), KPi (200 mM, pH 7.5), MsAcT (24 ng mL^{-1}), *Sc*ADH (50 U mL^{-1}); [b] Assay conditions: Benzyl alcohol (0 – 100 mM), vinyl acetate (10 mM), NADH (0.3 mM), KPi (200 mM, pH 7.5), MsAcT (24 ng mL^{-1}), *Sc*ADH (50 U mL^{-1}); [c] Assay conditions: Benzyl acetate (0.00 – 9.45 mM), NAD (5 mM), KPi (200 mM, pH 7.5), MsAcT (24 ng mL^{-1}), HLADH-E (5 U mL^{-1}).

More interestingly, using a coupled spectrophotometric assay the hydrolysis of benzyl acetate was measured demonstrating high affinity and uncompetitive inhibition for MsAcT. When benzyl acetate was added to the hydrolysis of vinyl acetate the apparent affinity towards vinyl acetate was significantly lower representing competitive inhibition (Fig. 6). Recently, a computational study suggested that benzyl acetate would have high affinity towards the aromatic active site.[6b] In this study, we demonstrated that the kinetics of the enzymatic reaction are severely inhibited by the addition of benzyl acetate indicating that ligand exchange severely lowers the reaction rates.

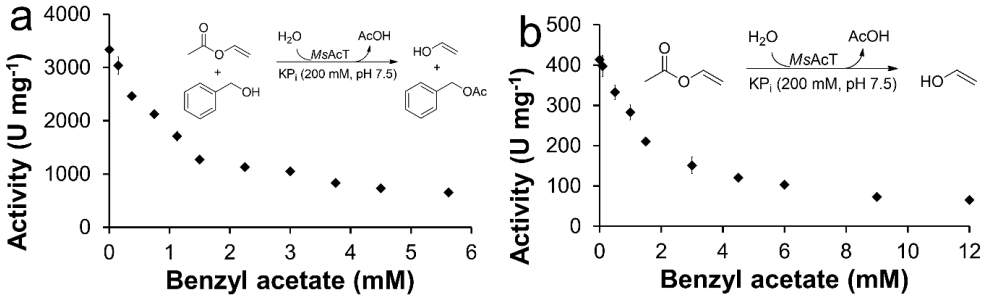


Figure 6. Product inhibition of the synthesis of benzyl acetate shown in figure a), while the competitive inhibition of the hydrolysis of vinyl acetate is shown in figure b). Reaction conditions: (i) product inhibition: benzyl acetate (0.0 – 12.0 mM), vinyl acetate (5.4 mM), benzyl alcohol (80 mM), NADH (0.25 mM), KPi (200 mM, pH 7.5), MsAcT (24 ng mL^{-1}), *Sc*ADH (50 U mL^{-1}); (ii) competitive hydrolysis: Benzyl acetate (0.00 – 5.6 mM), vinyl acetate (5.4 mM), NADH (0.25 mM), KPi (200 mM, pH 7.5), MsAcT (24 ng mL^{-1}), *Sc*ADH (50 U mL^{-1}). The error bars show the standard deviation of triplicates.

6.3 Conclusions

The synthesis of esters in high yields can be achieved in water with activated acids using MsAcT. The transesterification in water proceeds under kinetic control and careful monitoring of the reaction is necessary to avoid reverse hydrolysis. The kinetic parameters for MsAcT demonstrated a high synthesis to hydrolysis ratio. If the reaction is performed at high concentrations of benzyl alcohol, the product will form a separate layer leading to *in-situ* product removal essential for reducing enzyme inhibition under synthesis conditions. From a reaction engineering perspective, since the side product acetaldehyde is volatile, the laboratory scale reaction is straightforward to perform. On a larger scale standard measures to prevent acetaldehyde from escaping into the environment need to be taken.[8] These are well-established for industrial scale. Overall, organic solvents and wasteful neutralization steps can be avoided.

6.4 Experimental Section

6.4.1 Materials and methods

Ampicillin (Sigma-Aldrich), benzyl acetate (Sigma-Aldrich), benzyl alcohol (Acros Organics), bovine serum albumin (ThermoFischer), diethyl ether (VWR), DNAse I (bovine pancreas, Sigma-Aldrich), dodecane (Janssen), ethyl acetate (VWR), imidazole (VWR), isopropyl β-D-1-thiogalactopyranoside (ThermoFischer), lysozyme (chicken egg white, Sigma-Aldrich), magnesium sulfate (VWR), neopentyl glycol (Sigma-Aldrich), phenyl acetate (sigma-aldrich), potassium acetate (Sigma-Aldrich), potassium phosphate dibasic (Honeywell), potassium phosphate monobasic (Merck), sodium sulfate (Merck), vinyl acetate (Sigma-Aldrich). LB-medium consists of 1.0% (w/w) tryptone, 0.5% (w/w) yeast extract, 1.0% NaCl, and autoclaved at 121°C for 20 minutes. All media was supplemented with 100 µg mL^{-1} ampicillin before use. A multimode spectrophotometer plate reader (Synergy 2, BioTek) was used for spectrophotometric measurements. The pH electrode (Metrohm) was calibrated before use with standards of pH 4.00 and pH 7.00 with $R^2 > 0.98$. ^1H and ^{13}C NMR spectra were recorded on a Varian 400 (400 MHz) spectrometer in $CDCl_3$ (buffer for the NMR-monitored reaction). The chemical shifts are given in ppm relative to the solvent signal (^1H: δ ($CHCl_3$) = 7.26 ppm) and (13C: δ ($CDCl_3$) = 77.16 ppm, for centerline of $CDCl_3$ triplet).

6.4.2 Expression and purification of MsAcT

A previous expression protocol was followed with slight modifications[3d]. A synthetic gene of MsAcT (GenBank accession: ABK70783) from *Mycobacterium smegmatis* MC2 155 was cloned in pET16b-MsAcT containing a *C*-terminal HisTag and expressed in *E. coli* BL21(DE3). A baffled 5L shake flask containing 1 liter of TB-media was inoculated with 10 mL preculture and shaken until the OD600 reached 0.5 (25 °C, 180 rpm). Expression was induced with 0.1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and the culture was shaken overnight (25 °C, 180 rpm). The cells were harvested with centrifugation (10 min, 10 000 rpm, 4 °C, Sorvall RC 6+) and the supernatant was discarded. The cell pellet was washed with KPi buffer (20 mM, pH 7.5) and the supernatant was discarded again. The cell pellet was suspended in KPi buffer (20 mM, pH 7.5) containing 20 mM imidazole, lysozyme (0.5 mg mL^{-1}, chicken egg white, Sigma-Aldrich), DNase I (0.1 mg mL^{-1}, bovine pancreas, Sigma-Aldrich) were incubated at 0 °C for 30 minutes. The cells were lysed by cell disruption (Constant Cell disruption systems) for three

consecutive rounds (1.35 kbar). The cell-free extract was obtained after centrifugation (10 min, 10 000 rpm, 4 °C, Sorvall RC 6+). The cell free extract was filtered using a syringe filter (0.45 µm) before sample application on Ni Sepharose 6Fast Flow resin (12 mL, GE Healthcare) equipped in a XK 16/20 adapter (GE Healthcare) connected to a NGC system (Bio-Rad). Before sample application the system and resin were flushed with 20% ethanol, deionized water, and equilibration buffer containing KPi (20 mM, pH 7.5) with 20 mM imidazole. The protein was flushed through the column until all MsAcT protein was bound and eluted with a gradient of elution buffer KPi (20 mM, pH 7.5) and 500 mM imidazole. The eluted fractions containing MsAcT were desalted with a PD-10 desalting column and flash frozen with liquid nitrogen prior to storage at -80 °C.

6.4.3 SDS-PAGE analysis

The purity of MsAcT (24 kDa) was analyzed with SDS-PAGE analysis. The protein samples were denatured using a XT Sample Buffer (Bio-Rad) and XT Reducing Agent (Bio-Rad) at 95 °C for 5 to 10 minutes. A Criterion XT Bis-Tris MOPS 4-12% precast gel (Bio-Rad) equipped with a Precision Plus Protein Unstained Standard (Bio-Rad) and denatured protein samples was run at 180 V in MES buffer (Bio-Rad) for 40 minutes. The gel was stained with SimplyBlue SafeStain (ThermoFischer).

6.4.4 BCA assay

Protein content was determined with the bicinchoninic acid (BCA) protein quantitation kit (Thermo Scientific, Carlsbad, USA). Standard curves were prepared with bovine serum albumin (BSA) in the range of 0.003 to 1.38 mg mL^{-1} in a (poly)styrene 96-well plate. Samples were measured in triplicate and monitored at 562 nm utilizing a microtiter plate spectrophotometer (Synergy 2, BioTek).

6.4.5 MsAcT activity assay with neopentyl glycol (NPG)

The activity assay was started with the addition of 25 µL from a solution of MsAcT (0.02 mg mL^{-1}) dissolved in KPi buffer (200 mM, pH 8) to 975 µL ethyl acetate containing the external standard dodecane (10 mM) and NPG (100 mM). The biphasic reaction mixtures were shaken for 30 seconds at 2500 rpm at room temperature. Reactions were quenched after 0, 2, 4, 6 and 8 minutes by addition of excess sodium sulfate. 50 µL of each reaction mixture was pipetted into a GC vial

and 950 µL ethyl acetate was added prior to GC analysis. The amount of NPG monoacetate was quantified using external calibration curve. Activity is reported as micromoles of NPG monoacetate produced per minute per milligram of purified MsAcT.

6.4.6 MsAcT activity assay with *p*-nitrophenylbutyrate (pNPB)

The activity assay was started by adding 20 µL of MsAcT to 180 µL buffer containing KPi buffer (50 mM, pH 7.5) in transparent 96-well polystyrene plates. The rate of hydrolysis was monitored at 37 °C at 405 nm using a multimode spectrophotometer plate reader (Synergy 2, BioTek). Buffer and 96-well plate were pre-heated to 37 °C before the start of the reaction. An external calibration curve of *p*-nitrophenol was constructed between 0.00 – 1.00 mM.

6.4.7 Esterification of benzyl alcohol with potassium acetate under monophasic reaction conditions

For the synthesis of benzyl acetate from benzyl alcohol and potassium acetate the following reaction conditions were applied. MsAcT (12000 ng) was added to KPi buffer (200 mM, pH 7.5) containing 10 mM benzyl alcohol and 100 mM potassium acetate to yield a monophasic reaction mixture of 1 mL. For the hydrolysis of benzyl acetate to benzyl alcohol and potassium acetate the following reaction conditions were applied. MsAcT (12000 ng) was added to KPi buffer (200 mM, pH 7.5) containing 10 mM benzyl acetate and 100 mM potassium acetate to yield a monophasic reaction mixture of 1 mL. The reaction mixture was shaken at 1000 rpm at 21 °C. Samples were taken after 1, 2, 3, 4, 5 and 24 hours. Samples were quenched by addition of 500 µL diethyl ether containing dodecane (10 mM) as external standard. The mixture was rapidly vortexed for 30 seconds and centrifuged at 13 000 rpm to remove enzyme precipitates. The organic layer was collected, dried with an excess of dry $MgSO_4$, and analyzed by GC. The amount of benzyl alcohol and benzyl acetate was quantified using an external calibration curve.

6.4.8 Reversible hydrolysis and synthesis of isobutyl acetate by MsAcT

The reaction was initiated by the addition of MsAcT to a reaction solution resulting in the final concentration of 11 mM isobutanol, 212 mM vinyl acetate, KPi (200 mM, pH = 7.5), and 20 µg mL^{-1} MsAcT. The solution was resuspended vigorously during

addition for 1 - 3 seconds and measured immediately with ^{1}H-NMR (Agilent, 400 MHz). Also, the reaction was repeated and recorded with a dual pixel 12MP OIS (F1.7) camera to show the reversible synthesis and hydrolysis. For the synthesis of isobutyl acetate from isobutanol and vinyl acetate the following reaction conditions were applied. The reaction was initiated by the addition of MsAcT to a reaction solution resulting in the final concentration of 11 mM isobutanol, 212 mM vinyl acetate, 200 mM KPi (200 mM, pH = 7.5), and 0.5 µg mL^{-1} MsAcT. The solution was resuspended vigorously during addition for 1-3 seconds and measured immediately with ^{1}H-NMR (Agilent, 400 MHz).

6.4.9 Optimizing MsAcT concentration for the acylation of benzyl alcohol with vinyl-, phenyl-, and ethyl acetate

MsAcT (50 - 12000 ng) was added to KPi buffer (200 mM, pH 7.5) containing 10 mM benzyl alcohol, and 100 mM acyl donor being either vinyl-, phenyl-, and ethyl acetate to yield a monophasic reaction mixture of 1 mL. The mixture was shaken at 1000 rpm at 21 °C for 60 minutes. Product and substrate were extracted twice by addition of 500 µL diethyl ether containing dodecane (10 mM) as external standard. The mixture was rapidly vortexed for 30 seconds and centrifuged at 13 000 rpm to remove enzyme precipitates. The organic layer was collected, dried with an excess of dry $MgSO_4$, and analyzed by GC. The amount of benzyl alcohol and benzyl acetate was quantified using external calibration curve. Initial rates were calculated from the linear slope of benzyl acetate concentration over time.

6.4.10 Determination kinetic parameters for vinyl acetate and benzyl alcohol using GC analysis

For the kinetic parameters for benzyl alcohol the following reaction conditions were applied. MsAcT (50 ng) was added to KPi buffer (200 mM, pH 7.5) containing either 0, 1, 2.5, 5, 10, 15, and 20 mM benzyl alcohol, and 100 mM vinyl acetate to yield a monophasic reaction mixture of 1 mL. For the kinetic parameters for vinyl acetate the following reaction conditions were applied. MsAcT (50 ng) was added to KPi buffer (200 mM, pH 7.5), 10 mM benzyl alcohol, and 0, 10, 25, 50, 100, and 200 mM vinyl acetate to yield a monophasic reaction mixture of 1 mL. The mixture was shaken at 1000 rpm at 21 °C for 0, 2, 5, 10, 20, and 30 minutes. Product and substrate were extracted twice by addition of 500 µL diethyl ether containing dodecane (10 mM) as external standard. The mixture was rapidly vortexed for 30

seconds and centrifuged at 13 000 rpm to remove enzyme precipitates. The organic layer was collected, dried with an excess of dry $MgSO_4$, and analyzed by GC. The amount of benzyl alcohol and benzyl acetate was quantified using external calibration curves. Initial rates were calculated from the linear slope of benzyl acetate concentration over time. The curves were fitted to the Michaelis-Menten equation using the fit function of Gnuplot 5.2[13].

6.4.11 Coupled spectrophotometric activity assay with *Sc*ADH of vinyl acetate hydrolysis with MsAcT

The assay was performed in polyacrylate 1 cm cuvettes by monitoring the conversion of 0.25 mM NADH at 340 nm and 20 °C with an extinction coefficient of 6.221 mM^{-1} cm^{-1}. Before the addition of MsAcT, the reaction was monitored until stable. Enzymatic reactions were started by the addition of MsAcT with a final concentration of vinyl acetate (0 – 10 mM) in KPi (200 mM, pH 7.5), ScADH (50 U mL^{-1}), MsAcT (24 ng uL^{-1}). All measurements were performed in triplicates. The addition of additional *Sc*ADH did not result in a higher response for indirect acetaldehyde detection. The curves were fitted to the Michaelis-Menten equation using the fit function of Gnuplot 5.2[13].

6.4.12 Coupled spectrophotometric activity assay with *Sc*ADH of benzyl acetate synthesis with MsAcT

The assay was performed in polyacrylate 1 cm cuvettes by monitoring the conversion of 0.25 mM NADH at 340 nm and 20 °C by using an extinction coefficient of 6.221 mM^{-1} cm^{-1}. Before the addition of MsAcT, the reaction was monitored until stable. Enzymatic reactions were started by the addition of MsAcT with a final concentration of benzyl alcohol (1 – 100 mM), vinyl acetate (10 mM) in KPi (200 mM, pH 7.5), *Sc*ADH (50 U mL^{-1}), MsAcT (0.24 ng uL^{-1}). All measurements were performed in triplicates. The addition of additional *Sc*ADH did not result in a higher response for indirect acetaldehyde detection. The curves were fitted to the Michaelis-Menten equation using the fit function of Gnuplot 5.2[13].

6.4.13 Coupled spectrophotometric activity assay with HLADH-E of benzyl acetate hydrolysis with MsAcT

The assay was performed in polyacrylate 1 cm cuvettes by monitoring the conversion of 0.25 mM NAD+ at 340 nm and 20 °C by using an extinction

coefficient of 6.221 mM^{-1} cm^{-1}. Before the addition of MsAcT, the reaction was monitored until stable. Enzymatic reactions were started by the addition of enzyme with a final concentration of benzyl acetate (1 – 10 mM), in KPi (200 mM, pH 7.5), HLADH-E (5 U mL^{-1}), MsAcT (0.24 ng uL^{-1}). All measurements were performed in triplicates. The addition of additional HLADH-E did not result in a higher response for indirect benzyl alcohol detection. The curves were fitted to the non-competitive Michaelis-Menten equation using the fit function of Gnuplot 5.2[13].

6.4.14 Coupled spectrophotometric activity assay with ScADH of competitive vinyl- and benzyl acetate hydrolysis with MsAcT

The assay was performed in polyacrylate 1 cm cuvettes by monitoring the conversion of 0.25 mM NADH at 340 nm and 20 °C by using an extinction coefficient of 6.221 mM^{-1} cm^{-1}. Before the addition of MsAcT, the reaction was monitored until no chemical background hydrolysis occurred. Enzymatic reactions were started by the addition of MsAcT with a final concentration of benzyl acetate (0.00 – 5.6 mM), vinyl acetate (5.4 mM), NADH (0.25 mM), KPi (200 mM, pH 7.5), MsAcT (24 ng mL-1), ScADH (50 U mL-1). The error bars show the standard deviation of triplicates. All measurements were performed in triplicates. The addition of additional ScADH did not result in a higher response for indirect acetaldehyde detection.

6.4.15 Coupled spectrophotometric activity assay with ScADH of benzyl acetate synthesis with MsAcT in the presence of varying amounts of product

The assay was performed in polyacrylate 1 cm cuvettes by monitoring the conversion of 0.25 mM NADH at 340 nm and 20 °C by using an extinction coefficient of 6.221 mM^{-1} cm^{-1}. Before the addition of MsAcT, the reaction was monitored until no chemical background hydrolysis occurred. Enzymatic reactions were started by the addition of MsAcT with a final concentration of benzyl acetate (0.0 – 12.0 mM), vinyl acetate (5.4 mM), benzyl alcohol (80 mM), NADH (0.25 mM), KPi (200 mM, pH 7.5), MsAcT (24 ng mL^{-1}), *Sc*ADH (50 U mL^{-1}). All measurements were performed in triplicates. The addition of additional ScADH did not result in a higher response for indirect acetaldehyde detection.

Literature

[1] a) J. Clayden, N. Greeves and S. Warren, *Organic Chemistry*, Oxford University Press, 2nd edn., **2012,** 30; b) K. Schwetlick, *Organikum*, Wiley 23rd edn, **2009**; c) A. Miller, P. H. Solomon, *Writing Reaction Mechanisms in Organic Chemistry,* Harcourt Academic Press 2nd edn, **2000**.

[2] a) K. Faber, *Biotransformations in Organic Chemistry*, Springer, 7th edn, **2018**; b) U. Hanefeld, *Org. Biomol. Chem.* **2003**, *1*, 2405-2415; c) M. Paravidino, U. Hanefeld, *Green Chem.* **2011**, *13*, 2651-2657.

[3] a) I. Mathews, M. Soltis, M. Saldajeno, G. Ganshaw, R. Sala, W. Weyler, M. A. Cervin, G. Whited, R. Bott, *Biochemistry* **2007**, *46*, 8969-8979; b) L. Wiermans, S. Hofzumahaus, C. Schotten, L. Weigand, M. Schallmey, A. Schallmey, P. Domínguez de María, *ChemCatChem* **2013**, *5*, 3719-3724; c) K. Szymańska, K. Odrozek, A. Zniszczoł, G. Torrelo, V. Resch, U. Hanefeld, A. B. Jarzębski, *Catal. Sci. & Technol.* **2016**, *6*, 4882-4888; d) N. de Leeuw, G. Torrelo, C. Bisterfeld, V. Resch, L. Mestrom, E. Straulino, L. van der Weel, U. Hanefeld, *Adv. Synth. & Catal.* **2018**, *360*, 242-249.

[4] a) K. Manabe, S. Iimura, X.-M. Sun, S. Kobayashi, *J. Am. Chem. Soc.* **2002**, *124*, 11971-11978; b) K. Manabe, X.-M. Sun, S. Kobayashi, *J. Am. Chem. Soc.* **2001**, *123*, 10101-10102; c) S. Suda, H. Uyama, S. Kobayashi, *Proc. Japan Academy, Series B* **1999**, *75*, 201-206; d) O. Susumu, I. Mieko, T. Yoshio, *Biochim. Biophys. Acta Lipids Lipid Metabolism* **1979**, *575*, 156-165; e) A. H. Jan, E. Dubreucq, J. Drone, M. Subileau, *Biochim. Biophys. Acta Proteins Proteomics* **2017**, *1865*, 1105-1113.

[5] K. Buchholz, V. Kasche, U. T. Bornscheuer, *Biocatalysts and Enzyme Technology*. Wiley, 2nd edn., **2012**.

[6] a) M. L. Contente, A. Pinto, F. Molinari, F. Paradisi, *Adv. Synth. Cat.* **2018**, DOI: 10.1002/adsc.201801061; b) M. Kazemi, X. Sheng, W. Kroutil, F. Himo, *ACS Catal.* **2018**, *8*, 10698-10706.

[7] Y. F. Han, D. Kumar, C. Sivadinarayana, D. W. Goodman, *J. Catal.* **2004**, *224*, 60-68.

[8] G. F. Marc Eckert, Reinhard Jira, Hermann M. Bolt, Klaus Golka, *Acetaldehyde* in *Ullmann's Encyclopedia of Industrial Chemistry*, Wiley, **2012**, DOI: 10.1002/14356007.a01_031.pub2.

[9] a) A. Bruggink, E. C. Roos, E. de Vroom, *Org. Process Res. Develop.* **1998**, *2*, 128-133; b) *Catalysis: An Integrated Textbook for Students*, Edited by U. Hanefeld, L. Lefferts, Wiley-VCH, Weinheim **2017**. 370 pp. ISBN 978-3527341597; c) M. Y. Gololobov, E. V. Kozlova, I. L. Borisov, U. Schellenberger, V. Schellenberger, H.-D. Jakubke, *Biotechnol. Bioeng.* **1992**, *40*, 432-436; d) V. Kasche, *Enzyme Microb. Technol.* **1986**, *8*, 4-16.

[10] A. Wangler, R. Canales, C. Held, T.Q. Luong, R. Winter, D. H. Zaitsau, S.P. Verevkin, G. Sadwoski, *Phys. Chem. Chem. Phys.,* **2018**, *20***,** 11317-11326.

[11] A. Wangler, D. Böttcher, A. Hüser, G. Sadowki, and C. Held. *Chem. Eur. J.* **2018**, *24*, 16418-16425.

[12] a) M. Subileau, A. H. Jan, J. Drone, C. Rutyna, V. Perrier, E. Dubreucq, *Catal. Sci. Technol.* **2017**, *7*, 2566-2578; b) A.-H. Jan Deniau, M. Subileau, E. Dubreucq, *ChemBioChem* **2018**, *19*, 1839 – 1844; c) A.-H. Jan, E. Dubreucq, M. Subileau, *ChemBioChem* **2017**, *18*, 941 – 950.

[13] T. Williams, C. Kelley. Gnuplot 5.2. An Interactive Plotting Program. **2018**. Retrieved from http://www.gnuplot.info/docs_5.2/Gnuplot_5.2.pdf

7

Outlook and conclusions

They call me the breeze,

I keep blowing down the road

I ain't got nobody

I ain't carrying me no load

*JJ Cale, "*Naturally*",* Call me the Breeze, **1972**

The reaction conditions of (in)organic catalysts for transfer-type reactions rely on the use of (toxic) organic solvents, result often in poor yields, or demonstrate poor regio- or enantioselectivity. Enzymes, such as transferases, offer high enantioselectivity and high yield and are active in water-abundant media. Transferases are essential for life and have become an integral part of applied biocatalysis. Nature is dependent on activated donors and acceptors for transferases, often labile components that can degrade spontaneously when decoupled from their metabolic regeneration systems. Within this thesis, we have shown that glycosyl- and acyltransferases act as an additional toolbox for the synthesis of carbohydrates and esters. The societal relevance of these findings is that it provides a contribution to the development of sustainable enzymatic production technologies.

Within this thesis we described the use and potential of LeLoir glycosyltransferases (GTs) in applied biocatalysis. In **chapter 2**, the potential synthetic toolbox of LeLoir glycosyltransferases to couple, elongate, or branch glycoconjugates, oligosaccharides, and glycans with high regio- and stereoselectivity has been described. We highlighted the importance of separating kinetics from thermodynamics on glycosylation reactions to design nucleotide sugar regeneration systems efficiently. These insights can be pivotal to optimize the reaction conditions for the computer-controlled synthesis of large oligosaccharides using LeLoir GTs. Furthermore, such principles can be extended to transfer-type reactions where kinetic and thermodynamic control are catalytically interconnected.

In **chapter 3**, the use of the mCherry fusion with TreT of *Thermoproteus uzoniensis* (*Tu*TreT) to monitor protein aggregation and to increase enzyme solubility was demonstrated. mCherry is a versatile fluorescent protein tag, since upon denaturation it changes color from purple to green. It was successfully used to report the denaturation of the fusion protein mCherry *Tu*TreT. A large portion of TreT was produced as inclusion bodies and demonstrated catalytic activity (CatIBs).

The soluble expression of biocatalysts can be challenging to achieve. The fusion of fluorescent protein mCherry offers rapid insight in different formulations of immobilized enzymes. We highlighted the use of the mCherry TreT fusion construct

for monitoring the quality of the enzyme immobilized as CatIBs and as carrier attached immobilized enzymes in **chapter 4**. This allowed qualitative and quantitative assessment of native and denatured protein, and its distribution within an immobilization matrix. The performance of CatIBs and immobilized enzymes were compared, revealing that CatIBs can be applied for batch reactions with high total productivity of trehalose per liter of expression host culture. Nevertheless, immobilized enzymes on carrier materials exhibit a superior catalytic performance and ease of separation. If enzyme solubility and expression can be increased, higher STY and catalytic efficiency can be achieved using enzymes immobilized on carriers. Taken together, these parameters show that carrier-attached enzymes are more suitable for large-scale batch reactions. A judicious choice between CatIBs or enzyme immobilization for a particular batch process should be based on the required catalyst performance and the ease of enzyme production.

In **chapter 5,** mCherry *Tu*TreT catalyzes the formation of trehalose derivatives with a large substrate spectrum. Understanding the stereochemical conformation and anomeric configurations of sugars are important to understand their selectivities. For *Tu*TreT, the switch in anomeric selectivity for D- and L-sugar acceptors can be explained due to structural conformations of carbohydrates, resulting in α-D-α-D- or α-D-β-L-glycosidic linkages. These findings can be used for the enzymatic production of oligosaccharides and glycans with unnatural glycosidic linkages with rare L-glycopyranoses.

The synthesis of esters in high yield can be achieved in water with activated acids using an acyltransferase from *Mycobacterium smegmatis* as described in **chapter 6**. The intrinsic catalytic properties of MsAcT determine the selectivity between hydrolysis and transesterification of esters in water. The transesterification proceeds under kinetic control with a high synthesis over hydrolysis ratio. Under thermodynamic control complete hydrolysis was observed. Careful monitoring of the reaction in time is therefore necessary to avoid reverse hydrolysis.

Acknowledgements

When I walk out that door
Oh, how I want to be free, baby
Oh, how I want to be free
Oh, how I want to break free

Queen, "The Works", I Want to Break Free, **1984**

Everything that has beginning, has also an end. With this most *well-read* chapter of my thesis, I would like to express my deep sense of gratitude to all the people who I shared my PhD journey with. Looking back, this doctoral degree is an achievement that is not only mine to claim, but without doubt, the result of all those extraordinary persons who have supported me throughout these years. All the joyful memories, of curiosity and wonders, of moments of defeat and overcoming them, made this journey unforgettable. Through these experiences, I have grown into the person who I am today. Every end however, also has a new beginning. But before that endeavor starts, I would first like to take the opportunity to thank those who have contributed to this thesis.

First and foremost, I would like to sincerely thank my promotor **Ulf**. In 2015, when I walked into your office to persuade you of a research project, I was impressed by the quality of your scientific suggestions. In the next five years, I got to know you as direct, open, and reasonable person, who I deeply respect. Your genuine interest in the personal and professional development of PhD candidates, and the incredible pace and thoroughness of reviewing experiments, presentations, reports or manuscripts, is what distinguishes you from all other promotors. For this, I am filled with pride to say that I have completed my PhD under your supervision.

Of course, I would equally also like to thank my copromotor **Peter-Leon**. Your contribution as a scientist, supervisor, and colleague, has been invaluable to my personal and professional development. Without your sense of humor, the talks about life, and your scientific input, I have no doubt that the quality of this PhD journey would not have been the same. You have the rare qualities of being an excellent teacher, understanding mentor, and analytical scientist in one person. I am grateful for all your support, and I wish you a bright academic future.

I also would like to cordially thank the other members of the reading committee, **K. Szymańska**, **F. Hollmann**, **W.J.H. van Berkel**, **G. van der Marel**, **T. Desmet**, **W. R. Hagen** for your participation in the PhD ceremony.

The **European Research Area – Industrial Biotechnology (ERA-IB)** is gratefully acknowledged for the funding; and the partners for their participation in the project. **Arkadiusz** is gratefully acknowledged for the reactor design in the enzymatic glycosylation process. **Isabel**: your enthusiasm, perseverance, and quality of your scientific work was a great contribution to our final understanding of TreT. I would wholeheartedly like thank you for chapter 2, 3, and 5; your contributions were key to their success. **Rob**, thank you for all your useful suggestions regarding enzyme immobilization on carrier materials. I really enjoyed the talks we had about work, and life, and hope that our collaborations might find a continuation in the future. Thank you for chapter 2 and 4! Also, many thanks for my collaboration with **Andrzej** and **Katarzyna** for your hospitality and stay at the Silesian University of Technology in Gliwice. It was a great pleasure to work together on the immobilization of TreT, HNL, and MsAcT, and seeing your microreactors in action. **Dirk**, thank you for your input during the challenging protein expression and wish you a fruitful career in Bochum! To my fellow PhD candidates **Antje**, **André**, **Daria**, and **Marta**, thank you for the great collaboration and all the fun we had during all the ERA-IB project meetings. You are excellent scientists, and I hope we meet again in the future!

The appreciation for all the welcome input to my scientific work is extended to **Duncan**. There are not many people who have such passion and enthusiasm for science, and I enjoyed working together greatly! **Frank**, you have an incredible drive to find the next (big) research topic, and I wish you good luck in finding your next discovery. **Caroline**, it was great that I could share my passion for sugar chemistry, thank you for all the help during both MSc & PhD. There lays a very bright future ahead of you. **Fred**, there are not many scientists who can light up a room when talking about their own scientific curiosities like you do. I hope to see another lecture with your (dream) EPR machine in the future!

A word to my paranymphs. **Stefan**, a special word of gratitude goes to you. All your scientific input was invaluable to our PhD research, as chapters **2**, **3**, **4**, and **5** are also your effort. We supported each other through the thrills of science –

did nobody investigate this before!? – or the incredible depths – *somebody did investigate it before…* Without having somebody to go for a "borrel and bitterballen", the yoga and mindfulness sessions, the regular talks about life, I would not be quite sure if I would have endured my PhD journey. I can proudly say that this PhD thesis is also your achievement, so a huge "*dankjewel*" for that! **Hugo**, I am pretty sure that my PC would have ended outside the window if it weren't for our regular breaks, walks, coffee's, (brewing) beers, "*vasteloavend*", testing your synths, and many more. It is remarkable that our grandparents have met, and we meet a lifetime later. I am confident that, without you guys, my PhD would not have been half as much fun.

During my PhD, I had the pleasure to work with students from Life Science & Technology. As I learned incredibly much from you, I hope I have returned the favor, at least for a little. First, the BEP(ies). **Marit**, it was a great pleasure to work together. To this day, I am impressed how you managed all your extracurricular activities (1 day has 24 hours?). You are a highly independent researcher, and I am confident you will do well in the future! **Puck**, you have a sparkling curiosity to question everything. Never let anybody take this away from you! **Lysanne**, you are analytical and precise, which was a great contribution to this project. Then to my MEP(ies). **Jord**, you thanked me for my "excellent daily" supervision, but I am still questioning: for what exactly? You were incredibly autonomous, which reduced my supervision to problem-shooting now and then. I am proud of chapter 6, and I wish you all the best with your (*real*) job! At last, **Hessel**, your diligence and enthusiasm in the production of your purple protein made all the difference for chapter 5. Wish you all the best in the future!

Dear fellow (ex-)BOC/BT PhD colleagues, thank you for creating such a great atmosphere and memories. To **Hanna**, catch! Our talks about life, the good and the bad, helped me to stay on track (and reality sometimes). Especially, when I went through the eye of a *needle*. To **Georg**, thank you the good times; use your acquired "gabber" skills wisely! To **Morten**, our scientific discussions were great and I really enjoyed going now and then to Chateau de Jacques. To **Marine** and

Seb, thank you for all the lovely talks and wish you all the best starting up your new lives! To “almighty” **Albert**, thanks for all the laughs over the last few years! To **Sam,** thank you for demonstrating the “formalities” of ceremonies from student associations and good luck with your PhD! To **José**, it was great to share our time in Gliwice, and I really enjoyed our collaboration on cascade biocatalysis in flow. To **Alexandra**, thank you for the organization of “social activities” (euphemism?), and wish you the best with your magnetic nanoparticles! **Retna**, I wish you also the best as well! **Florence**, it was great to share a coffee/drinks/food to talk about the PhD journey. Your sanguinity is refreshing, and I wish all the best for the future. **Wijb**, all the moments of brewing beers, game nights, and drinks were great. Let us continue that in the future!

Fabio, whenever I look back on the shared memories at BOC, at conferences, or in the lab, I have a big smile on my face. Your passion for science – in a way only Italians can have - is contagious. I am confident that you have a very bright future ahead! **Linda**, also for you a very big thank you for being there during my time as a MSc and PhD candidate. Our regulars talk about research has been a great support throughout the last five years, and I would like to wholeheartedly like to thank you for that! I wish you all the best in continuing your own journey in the future. **Paula** and **Rosario**, you were there for me with all my initial PhD struggles. To this day, I am very grateful for your help and advices.

The backbone of the research group Biocatalysis are the technicians who make all the research performed of this PhD thesis possible. **Marc**, I am grateful for setting up the large fermenter and producing the large amounts of purple protein. Thank you for your support and involvement! **Lloyd**, I am glad you joined the BOC team. Without your technical support of the HPLCs, I am sure that I would not be able to obtain the results presented in this thesis. Besides this, I am looking forward to your *musical performance*(***s***?)! **Laura**, thank you for all your useful suggestions in the MB lab and the support throughout the years. **Remco** and **Stephen**, thank you for help with GC and NMR analysis. **Maarten**, thank you for the assistance with analytics and the organization relocation of the laboratories. It

meant a world of difference when we just got up and running. **Mieke van der Kooij,** I am pretty sure that BOC will fall in despair without your support and guidance of all the aspects of the TUD. A very big thank you for answering all the (countless) questions I had.

At last, to all the people who have contributed within BOC and the department of Biotechnology: **Dirk**, **Rob**, **Diego**, **Caroline**, **Natália, Eman, Kourosh, Simon, Sander, Michael, Jia Jia, Florian, Mieke, Nicolas, Kristina, Ashmiani, Kirsten, Tiago, Sabry, Sandy, Jonathan, Eva, Max, Meera, Elena, Wuyuan, Jia Jia, Milja, Florian, Mieke,** and all ***the fantastic others***! The list of people is endless, but every single person contributed to this PhD journey. I just loved to the daily coffee's, the (ridiculous) quantities of (cake, but preferably) vlaai, and the lunch breaks together. I could say a word about the (weekly) "borrels", but it is probably for the best, that I do not.

And then, to my family *Mestrom*. Groote broôr **Koen** en kleine broôr **Jules**, dét ig aaf en toe kós aafwejje van dét hoge Hollandj haet veur mig al 't versjil van de wereld gemaak. Hieël erg bedankt det geer dr väör mig zeentj, en det ozze broöderbandj dr altied, moog blieve. Dét geldj natuurlijk oug väör de (sjoône)zusse **Jule** en **Joël**, geêr zeent un aanwins binne ozze familie en ig zoê ug neet meer los van wille zeên. Leve **Pap** en **Mam**, geêr zeent mien veurbeeld in alles wat ig doon. Bedankt veur ugge onveurwaardelijke steun al mien gânse leave, en dét al welgetild 28 joar lang. Ig bun dr hertstikke trots op ugge zoôn te moôgge zeên. Zonger ug, bun ig der van euvertuûg, had dit proofsjrif nooit 't daagleêg gezeen. Doarom draag ig dit proofsjrif op aan ug, veur ôs **Pap** en **Mam**.

Curriculum vitae

Luuk Mestrom was born in (Sittard, NL) in 1992 and received his bachelor in biochemistry from the Zuyd University of Applied Sciences (Heerlen, NL). During his bachelor he joined the Materials and Interface Chemistry group under supervision of Prof. Nico Sommerdijk at the Technical University Eindhoven (NL) for the development of ferrofluids. During that time, he became the co-inventor of a PET-recycling process at Ioniqa Technologies B.V., now awarded as one of the National Icons of the Netherlands (Dutch Ministry of Economic Affairs and Climate Policy) with a running 10 kton pilot plant (Rotterdam, 2020). He joined Maastricht University (NL) as a synthetic organic chemist for the development of novel BioBased Building Blocks (BBBBs). Afterwards, he pursued a MSc and PhD in Biocatalysis at TU Delft (NL) under the supervision of Prof. dr. Ulf Hanefeld and dr. Peter-Leon Hagedoorn. Recent work was focussed on biocatalytic industrial applications of acyl- and LeLoir glycosyltransferases.

List of publications

[1] Vilaplana Artigas, M., **Mestrom, L**., De Groot, R., Philippi, V., Guerrero Sanchez, C., Hooghoudt, T., **2014**. Polymer Degradation. WO/2014/209117, Eindhoven.

[2] **L. Mestrom**, J. J. M. Lenders, R. de Groot, T. Hooghoudt, N. A. J. M. Sommerdijk, M. V. Artigas, *Nanotechnology* **2015**, *26*, 285602.

[3] K. Bernaerts, **L. Mestrom**, S. De Wildeman. Biocatalysis toward New Biobased Building Blocks for Polymeric Materials in *Applied Biocatalysis: From Fundamental Science to Industrial Applications*, **2016**, 405-428.

[4] **L. Mestrom**, P. Bracco, U. Hanefeld, *European Journal of Organic Chemistry* **2017**, *2017*, 7019-7025.

[5] N. de Leeuw, G. Torrelo, C. Bisterfeld, V. Resch, **L. Mestrom**, E. Straulino, L. van der Weel, U. Hanefeld, *Advanced Synthesis & Catalysis* **2018**, *360*, 242-249.

[6] a) **L. Mestrom**, S. R. Marsden, M. Dieters, P. Achterberg, L. Stolk, I. Bento, U. Hanefeld, P.-L. Hagedoorn, *Applied and Environmental Microbiology* **2019**, *85*, e03084-03018; b) **L. Mestrom**, S. R. Marsden, M. Dieters, P. Achterberg, L. Stolk, I. Bento, U. Hanefeld, P.-L. Hagedoorn, *Applied and Environmental Microbiology* **2019**, *85*, e00942-00919.

[7] **L. Mestrom**, J. G. R. Claessen, U. Hanefeld, *ChemCatChem* **2019**, *11*, 2004-2010.

[8] S. R. Marsden, **L. Mestrom**, I. Bento, P.-L. Hagedoorn, D. G. G. McMillan, U. Hanefeld, *Advanced Synthesis & Catalysis* **2019**, *361*, 2649-2658.

[9] **L. Mestrom**, M. Przypis, D. Kowalczykiewicz, A. Pollender, A. Kumpf, S. R. Marsden, I. Bento, A. B. Jarzębski, K. Szymańska, A. Chruściel, D. Tischler, R. Schoevaart, U. Hanefeld, P.-L. Hagedoorn, *International Journal of Molecular Sciences*, **2019**, *20*, 5263.

[10] S. R. Marsden, **L. Mestrom**, D. G. G. McMillan, U. Hanefeld, *ChemCatChem*, **2020**, *12,* 426-437.

[11] **L. Mestrom**, S. R. Marsden, D. McMillan, Rob Schoevaart, Peter-Leon Hagedoorn, Ulf Hanefeld, **2020**. *ChemCatChem*, **2020**, *12*, 3249-3256. DOI: 10.1002/cctc.202000241.

[12] **L. Mestrom**, S. R. Marsden, H. van der Eijk, J. U. Laustsen, C. M. Jeffries, D. I. Svergun, P.-L Hagedoorn, I. Bento, U. Hanefeld, *accepted, ACS Catal.* **2020**, 10, 8835–8839. DOI: 10.1021/acscatal.0c02117

[13] S.R. Marsden, **L. Mestrom**, H.J. Wijma, S. Noordam, D.G.G. McMillan, U. Hanefeld, *under review*.