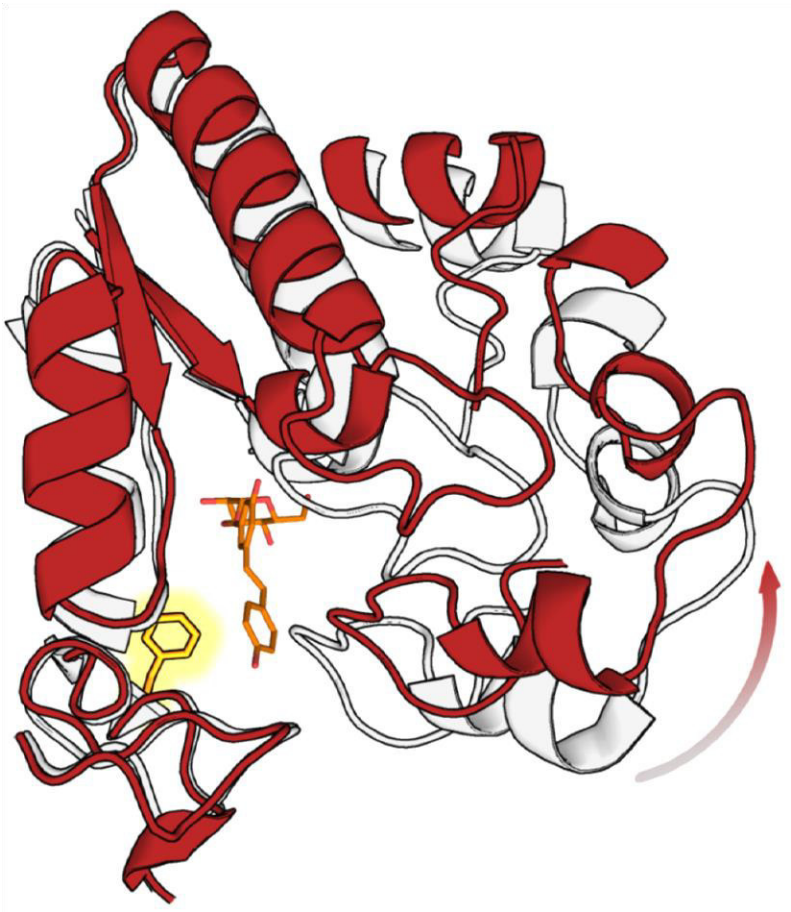
The Conversion of *Bifidobacterium adolescentis* Sucrose Phosphorylase into a Polyphenol Transglucosidase *via* Structure-based Enzyme Engineering



Michael Kraus

The Conversion of *Bifidobacterium adolescentis* Sucrose Phosphorylase into a Polyphenol Transglucosidase *via* Structure-based Enzyme Engineering

Dissertation zur Erlangung des naturwissenschaftlichen Doktorgrades
der Julius-Maximilians-Universität Würzburg

vorgelegt von

Dipl. Chem. Michael Kraus

aus Amberg in der Oberpfalz

Würzburg 2018

Eingereicht bei der Fakultät für Chemie und Pharmazie am

18.06.2018

Gutachter der schriftlichen Arbeit

1. Gutachter: Prof. Dr. Jürgen Seibel
2. Gutachter: Dr. Clemens Grimm

Prüfer des öffentlichen Promotionskolloquiums

1. Prüfer: Prof. Dr. Jürgen Seibel
2. Prüfer: Dr. Clemens Grimm
3. Prüfer: Jun.-Prof. Ann-Christin Pöppler

Datum des öffentlichen Promotionskolloquiums

29.05.2019

Doktorurkunde ausgehändigt am

“You can't always get what you want

But if you try sometimes well you might find

You get what you need”

Richards K. and Jagger M.

Experimental studies of the presented work were performed from June 2011 until June 2017 at the Institute of Organic Chemistry, Julius-Maximilians-Universität Würzburg under supervision of Prof. Dr. Jürgen Seibel.

Parts of this work have previously been published as:

- M. Kraus, C. Grimm, J. Seibel, Chem Comm, 2017, 53, 12181-12184
- M. Kraus, J. Görl, M. Timm, J. Seibel, Chem Comm, 2016, 52, 4625-4627
- M. Kraus, C. Grimm, J. Seibel, Chembiochem, 2016, 17, 33-36
- M. Kraus, C. Grimm, J. Seibel, Scientific Reports, 2018, 8, 10490

ACKNOWLEDGEMENTS

To paraphrase Newton: If I have come this far, it would not have been possible without the support of several great people.

First I want to thank Prof. Dr. Jürgen Seibel for giving me the opportunity to pursue my Ph.D. in his group and providing me with support and guidance throughout it, while allowing me to simultaneously follow my own ideas and perform independent research.

I also want to express my gratitude to Dr. Clemens Grimm from the Chair of Biochemistry for introducing me into protein crystallography and for solving the X-ray structures that form the foundation of important parts of this work.

Muchisimas gracias por Dr. Maria Ortiz Soto for so many things: I shall miss our discussion, on work, science in general and pretty much all other topics. You also possess the gift of providing me (and others) with always constructive criticism and thereby frequently helped along my work including this one. Also thanks for the unforgettable Mexican food you brought back here.

Likewise I want express my gratitude to Julian Görl for introducing me to Autodock and Photoshop (after flat out preparing way too many images for me), for his very productive cooperation on the research presented in chapter 3 and for the many work and non-work related discussion. Also thank you for many a delicious lunch.

Thank you Malte Timm for showing me the ropes in many of the biochemical techniques.

To Christina Plank, Aravindan Vishnavathan and Emilia Gärtner: Thank you for giving me such a warm welcome in the Biochemistry department.

To all of the previous and current members of the Seibel research group (Maria, Julian, Malte Elli, Nadja, Riaz, Bernd, Tim, Christian, Julian, Natalia, Julia and Sabine) thank you for the supportive and friendly atmosphere during worktime, cocktails evenings, and sailing- canoe- and skiing-trips.

I want to thank my parents for their love and support during this endeavour, my brother Johannes for checking and correcting my english punctuation and my brother Benedikt for continuously reminding me that all good things must come to an end.

Finally to my beloved Xiaoqiao: Thank you for picking me up after my failures and celebrating my successes with me and I also want you to know that I appreciate your great patience with me during this adventure.

Table of Contents

Chapter 1

INTRODUCTION

1.1. Background

1.1.1. Drug and Bioactive Compound Glycosylation

In 1964, it was said that a spoonful of sugar helps the medicine go down.[1] Over the past five decades it has become clear that sugar can improve more than just the taste of medicine and carbohydrates often are an integral part of the remedy in the first place. Several potent classes of naturally occurring drugs contain one or more carbohydrate moieties. Prominent examples are macrolide antibiotics like erythromycin[2-5], glyco-peptide antibiotics like vancomycin[6, 7], anthracycline cystostatica like doxorubicine[8] or enediyne antitumor antibiotics like calicheamicins.[9, 10] In addition the glycosylation of carbohydrate free (aglycon) drugs or bioactive molecules has emerged as a strategy to improve or fine-tune them.

The first example, Taxol (**Figure 1A**), is a powerful chemotherapeutic against breast cancer that is virtually insoluble in water (<0.004 mg/ml).[11] The attachment of a methyl-α-D-glucopyranoside via a malonyl ester increases the water solubility by a factor of 88. The glycosylated derivate further displayes a selectivity for breast adenocarcinoma cells over healthy kidney cells that improves by more than three orders of magnitude.[11]

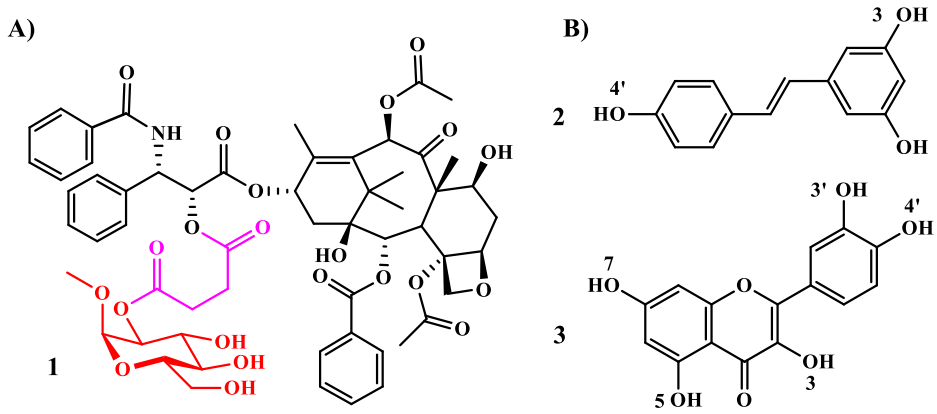


Figure 1 Glycosylation target: A) Glycosylated Taxol (1) Taxol moiety (black) linked by a malonyl ester (pink) to the OH-2 of methyl-α-D-glucopyranosid (red). B) Resveratrol (2) and quercetin (3) with numbering of their OH-groups

A further example for bioactive compounds that can be improved via glycosylation are dietary polyphenols like resveratrol and quercetin (**Figure 1B**). They act as antioxidants and have drawn considerable research interest because of their anti-carcinogenic activities[12, 13] and their potential for lifespan and health span extension.[14-16] Both compounds possess a limited bioavailability due to their poor water solubility. Glycosylation has been used to improve polyphenol uptake,[17, 18] install antiviral properties[19] or improve binding to anticancer targets.[20]

1.1.2. Chemical Glycosylation

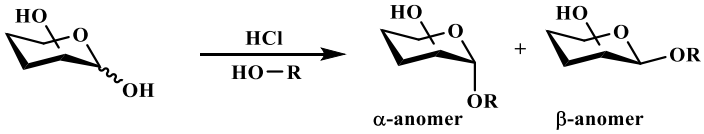


Figure 2 Fischer glycosylation

The most straightforward way to make glycosidic bonds is the Fischer glycosylation (**Figure 2**). A monosaccharide reacts with an excess of the acceptor alcohol under acid catalysis to give the sugar acetals as a mixture of α- and β- products (**Figure 2**).[21, 22] Naturally occurring glycosides also are linked via the anomeric carbon. In this respect the Fischer glycosylation is a suitable method for preparing these compounds in the laboratory. For more complex substrates it is however not a viable strategy. Quercetin (**Figure 1**) for example contains five potential glycosylation sites which are chemically and sterically almost equivalent. A simple Fischer glycosylation of quercetin results in five different regio-isomers as a mixture of α- and β-stereo isomers. Chemical synthesis of glycosides has a long history[21] and the ingenuity of the synthetic organic chemist has provided the ability to synthesize even complex oligomeric glycan structures.[23] Despite the existence of a wide range of chemo-synthetic tools[24] and their continued improvement[25] some fundamental problems persist. In order to solve the regio- and further stereo selectivity issues described above, chemical synthesis must rely heavily on protection group chemistry and sophisticated activation steps. The result is a low atom efficiency i. e. most of the material used during synthesis does not end up in the final product. This is undesirable from an ecologic and an economic standpoint alike.

1.1.3. Enzymatic Glycosylation

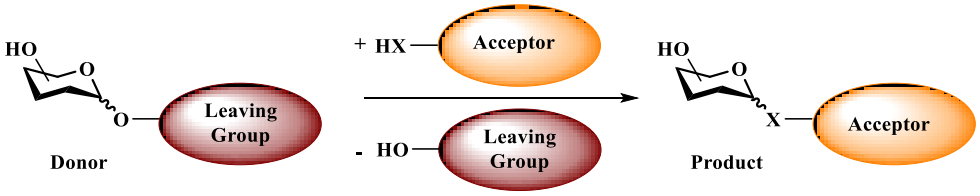


Figure 3 General enzymatic glycosylation reaction

The alternative to classic chemical synthesis is the application of nature's tools: enzymes. They forego the need for protection groups due to their built-in ability to form glycosidic bonds stereo- and regio- and chemo selectively. *In vivo*, glycosylation is achieved by two different enzyme classes: glycosyl transferases (GT's, E. C.: code 2.4.x.x) and glycosyl hydrolases (GH's, E. C.: code 3.2.x.x). Both enzyme classes require an activated donor substrate and transfer to a nucleophilic functional group in the acceptor usually an alcohol (**Figure 3**). The transfer to other nucleophilic functionalities like amines, thiols, phenols, carboxylic acids or amides has been observed as well. Both GT and GH enzymes are completely selective regarding the stereochemistry of the anomeric carbon of their donor substrates and their product molecules.

Glycosyl Transferases

In living systems most glycosylation reactions are performed by glycosyl transferases. Consequently GTs exist for almost all naturally occurring substance classes. They are involved in the synthesis of disaccharides (i. e. sucrose)[26] and polysaccharides (i.e. glycogen)[27] as well as protein N-[28, 29] and O-[30] glycosylation. Others catalyse the sugar transfer to antibiotics[31] or to polyphenolic compounds[32]. They usually display a very high affinity towards their donor and acceptor substrates, with K_M-values that can be as low as 1-10 µM,[32] and possess an excellent regio-selectivity towards their natural substrates.

Their main drawback is the use of nucleoside (di)-phosphate activated carbohydrates or other expensive and poorly available donor substrates[33] limiting the use of GT's in the synthesis of bulk chemicals.

Glycosyl Hydrolases and Transglycosidases

The most frequent reaction catalysed by GH enzymes is hydrolysis, the transfer of the sugar to water. Several enzymes within the glycosyl hydrolase families transfer to other nucleophiles like sugars, alcohols or phosphate. These enzymes are called transglycosidases, and are involved in the formation of small saccharides like sucrose[34], sugar phosphates[35, 36] or polymers like amylose[37], levan[38] or dextran[38]. The transglycosylation to hydrolysis ratio of these enzymes varies and in most cases hydrolysis remains the predominant reaction.[39] In this case transglycosylation products are obtained though manipulation of the reaction conditions. Albeit rare, some true transglycosidases exist. For example sucrose phosphorylases efficiently catalyse the sucrose or glucose-α-D-1-phosphate and hydrolysis of the donor substrates or the products is a slow side reaction.[40]

The key advantage transglycosidases over glycosyl transferases is their capability to consume relatively cheap donor substrates that are available in bulk amounts, like sucrose[35], maltose[36] or lactose[41]. Transglycosidases possess a far narrower natural acceptor spectrum

than GT enzymes. While some enzymes possess a degree of substrate promiscuity[40] and act on other acceptor substrates like polyphenols[42], these reactions are usually inefficient and result in low yields and difficult to separate reaction mixtures.

1.2. Thesis outline –goal of this work

As both GT and GH enzymes possess their own disadvantages as biocatalysts, we endeavour to create novel tools that combine the advantages of both enzyme classes. The focus of this work is the reinvention of a transglycosidase to enable the glucosylation of acceptor substrates that were previously unattainable for the enzyme.

The goal of this work is:

A) Enabling the glucosylation of aromatic substrates by a transglycosidase through structure based enzyme engineering

B) Evaluation of the generated enzyme variants and their products

C) Investigation of the structural and mechanistic properties of the novel variants for further structure based enzyme engineering

Resveratrol and quercetin were chosen as target acceptor substrates due to their promising bioactivities (see **chapter 1.1.1**).

The enzyme of choice is the Sucrose Phosphorylase (SP) from *Bifidobacterium adolescentis* (EC 2.4.1.7, GH13, BaSP), which utilizes the cheap and abundant donor substrate sucrose. BaSP was chosen as it is the only SP for which a crystal structure has been solved and thus mutagenesis targets can be chosen rationally.[43] [44] In addition BaSP is relatively thermostable and tolerant to organic solvents, especially DMSO which is required for solubilisation of the hydrophobic target acceptor substrates.[40]

1.3. Sucrose Phosphorylase

1.3.1. Native reaction of Sucrose Phosphorylase

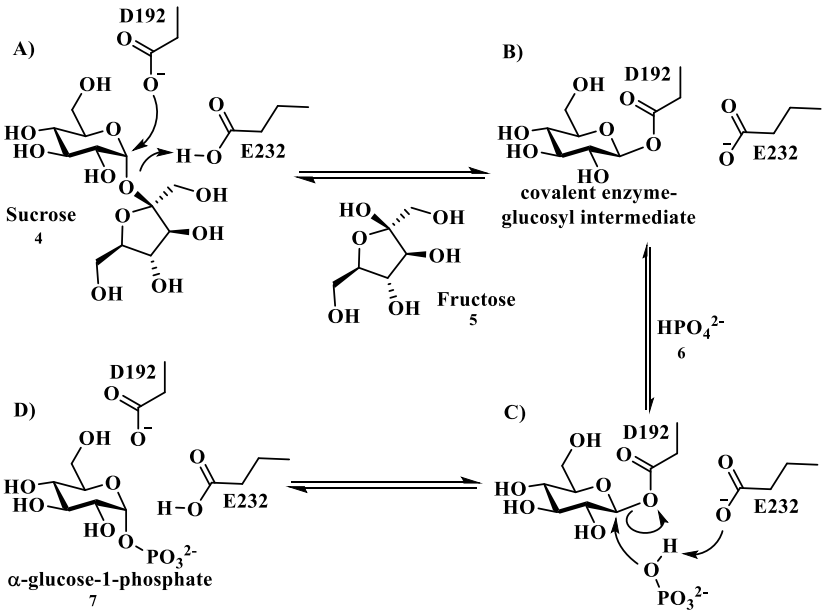


Figure 4 Reaction mechanism of SPs. The residue numbering is that of *Bifidobacterium adolescentis* sucrose phosphorylase (BaSP). A) Glu232 protonates the glycosidic O-atom to transform the fructosyl moiety of sucrose (4) into a better leaving group. Asp192, the catalytic nucleophile attacks from above and forms the covalent enzyme-glycosyl intermediate and fructose (5) is eliminated (B). C) Glu232 deprotonates the nucleophile hydrogen phosphate (6), which then displaces Asp192, to form α-D-glucose-1-phosphate (7) (D). The reaction is fully reversible and the synthesis of sucrose from α-D-glucose-1-phosphate and fructose is achieved via reversing the order of all steps.

Sucrose Phosphorylases (EC 2.4.1.7) belong to the CAZy family GH13, the α-amylase family.[45, 46] The native reaction catalysed by these enzymes is the interconversion of sucrose (**4**) and α-D-glucose-1-phosphate (**7**).[34, 35] Depending on the direction of the reaction, fructose (**5**) and phosphate (**6**) act as either acceptor substrate or leaving group.[34, 35] The reaction follows the ping-pong or Koshland double displacement mechanism (**Figure 4**),[47] during which a covalent enzyme substrate intermediate is formed. In BaSP Asp192 acts as the nucleophile and Glu232 as the acid/base catalyst.[44] This intermediate, a β-glucosyl-aspartic-acid ester, is remarkably stable and has been identified via radioactive labelling[48] and x-ray crystallography[44] (**Figure 5**). Due to the ping-pong mechanism the transferred glucosyl moiety, in the starting materials and in the final product always has an α-configuration. The reported affinities of various SPs towards the four natural substrates sucrose, phosphate, glucose-α-D-1-phosphate and fructose depends on the enzyme and the assay conditions. The affinity of SPs for their natural substrates is in the mM range. In general the K_M values are in the range of 5-20 mM.[40, 49] For *Leuconostoc mesenteroides* SP (LmSP) KM values for all natural substrates were determined (**Table 1**).[50] The K_M values of six different SP' variants for sucrose range from 0.8 to 14.1 mM, whereas the values for fructose range from 8.3 to 22.7 mM.[40]

Table 1 Affinity of LmSP for its natural substrates measured at 30 °C[49]

Substrate	Sucrose	Phosphate	Glucose-phosphate	Fructose
K_M values [mM]	*9.8*	*6.0*	*4.7*	*13*

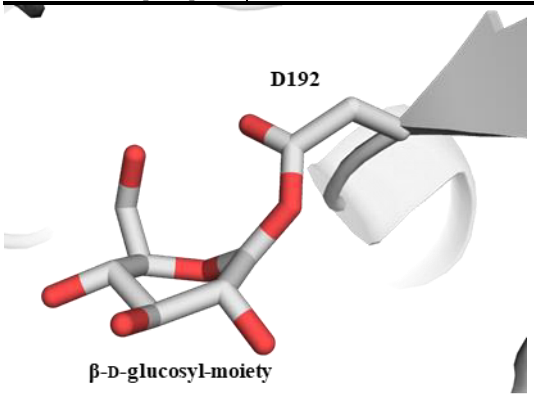


Figure 5 β-D-glucose covalently linked to Asp192 of BaSP. (PDB ID 2gdv, chain A)

1.3.2. General structure and domain architecture of BaSP

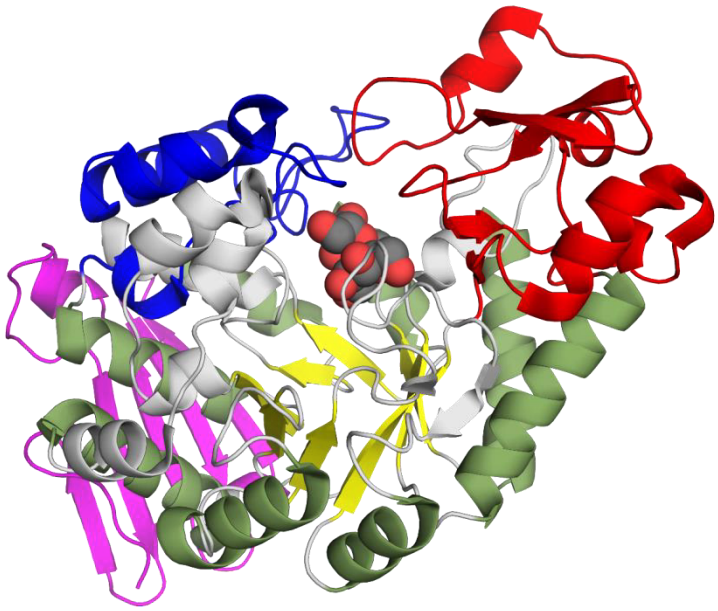


Figure 6 Structure of BaSP (PDB ID 2gdu) in complex with sucrose. The central β-barrel is shown in yellow, the surrounding α-helices in green. The active site is closed by domains B (red) and B' (blue). The C-terminal domain C (pink) is not catalytically important.

Sucrose Phosphorylase belongs to the GH-H clan of glycosyl hydrolases consisting of the glycoside hydrolase families GH13, GH70 and GH77.[51] SPs are members of the GH13 family and composed of four domains.[43] The core domain is the TIM-barrel, domain A (residues 1-85, 167-291, 356-435).[43] This $(\beta\alpha)_8$-barrel is made up of eight parallel β-sheets (**Figure 6**, yellow) which are surrounded by eight α-helizes featuring an antiparallel orientation in respect to the sheets(**Figure 6**, green).[43, 46] At the N-terminal side of the TIM-barrel short loops are inserted between the helices(αn) and sheets (βn+1). At the C-terminal side the

loops between the sheets βn and helices(αn) are longer and contain other secondary structures as well.[43, 44] In SPs two of these loops are in fact domains. Domain B (**Figure 6**, red, residues 86-166) is located between strand β3 and helix α3 and domain B' (**Figure 6**, blue, residues 292-355) is found between strand β7 and helix α7.[43] The TIM barrel and a domain B are common features in the GH13 family whereas domain B' is typical for the SP sub-family of the GH13 family. The active site is located in the cleft between the TIM barrel and domain B.[46] The GH-H clan shares a common active site and catalytic residues, although TIM-barrel elements of the different families are circularly permuted.[52, 53] Residues 436-504 form domain C which is common in the GH13 family and considered to stabilize the catalytic centre by shielding hydrophobic residues of the TIM-barrel from the solvent (**Figure 6**, pink).[43, 46 51]

1.3.3. Loop rearrangement during sucrose conversion

Prior to this work the x-ray structures of BaSP in two distinct conformations were solved.[43, 44] The active sites of both conformations differ in the acceptor binding subsite +1-site, while the donor binding subsite -1 remains identical.[44] One conformation is responsible for the accommodation of sucrose and fructose as demonstrated by the presence of sucrose in the crystal structure (PDB ID 2gdu). This conformation will be referred to as the *F*-conformation (short for fructose binding conformation).[44] The second conformation is the proposed

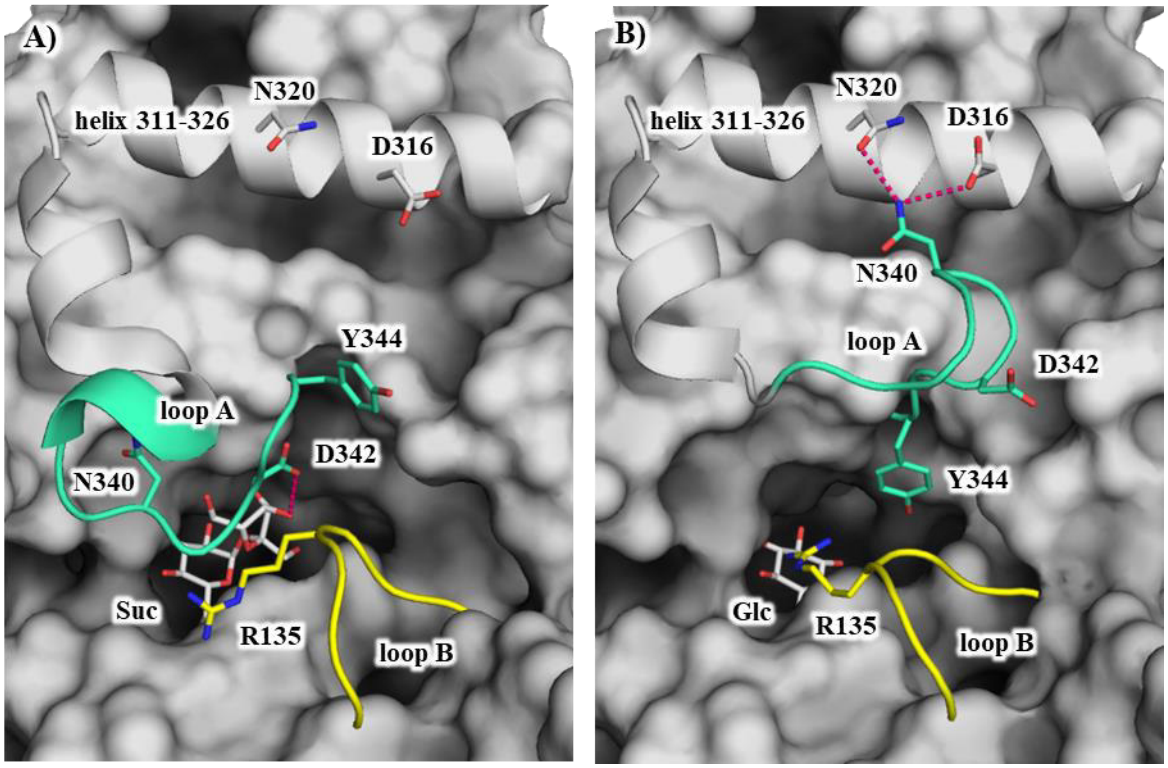


Figure 7 Loop rearrangement of BaSP and the orientation of the flexible residues Arg135, Asp342 and Tyr 344. Turqoise: loop A, yellow: loop B, A) *F*-conformation (PDB ID: 2dgu) pink dashes indicate the hydrogen bond of Asp342 to OH-4 of fructose B) *P*-conformation (PDB ID 2gdv, chain B) pink dashes indicate the hydrogen bond of Asn340 to Asp 316 and Asn320.

phosphate binding conformation (*P*-conformation).[44] *F*- and *P*-conformation differ mainly in the orientation of two loops. Loop A (^{336}AAASNLDLY^{344}) is in domain B' whereas loop B (^{132}YRPRP^{136}) is part of domain B.[43, 44]

Loop A rearranges drastically during the catalytic cycle. The C_α of Asn340 in its centre is moved by 14.9 Å. In the *F*-conformation loop A points towards the active site and Asp342 becomes a part of the acceptor binding site while the sidechain of Tyr344 is outside of it (**Figure** 7**A**). In the *P*-conformation loop A is oriented away from the active site and towards the α-helix formed by the residues 311-326. Tyr344 becomes a part of the acceptor binding site, while Asp342 leaves it.[44] In this conformation Asn340 exhibits hydrogen bonds with Asp316 and Asn320 of the helix (**Figure** 7**B**).

The rearrangement of loop B is less drastic than the one of loop A. The key residue in this loop is Arg135. Its sidechain is part of the active site in the *P*-conformation but not in the *F*-conformation. The C_α of Arg135 is shifted by 3.4 Å during the loop rearrangement (**Figure** 7).[44]

Both *F*- and *P*-conformation do not present an access channel. Substrates and products must enter and leave the enzyme either during the loop rearrangement or via a third open conformation that has not been described.

1.3.4. Active site and substrate recognition of BaSP

Subsites in GH enzymes

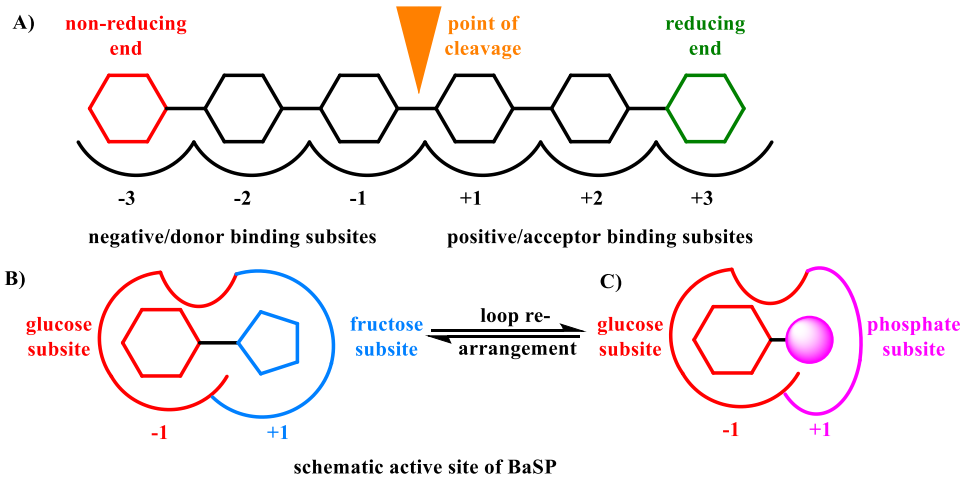


Figure 8 Schematic representation of subsites. A) general subsites in glycosyl hydrolases B) Fructose binding conformation (*F*-conformation) of BaSP C) Phosphate binding conformation (*P*-conformation) of BaSP

The subsites in GH enzymes are numbered in respect to the point of cleavage.[54] The subsites towards the non reducing end are called donor subsites and are numbered with negative integers denoting their distance to the point of cleavage. The subsites towards the reducing end are referred to as acceptor binding subsites and are numbered with positive

integers (**Figure 8A**).[54] Sucrose Phosphorylases possess a common -1-subsite but two different +1-subsites due to the above mentioned loop rearrangement (**Figure 8A, B**).

The donor binding site/-1-subsite

The glucosyl moiety is recognized in the donor or -1-subsite of BaSP and adopts the $^{4}C_{1}$- or chair conformation in all but one published crystal structure. (PDB ID 2gdu and 2gdv chain B). The single exception is the β-linked covalent intermediate which features a twisted $^{1,4}B$- or boat conformation (**Figure 5**).[44] The coordination of the carbohydrate hydroxyl groups by the proteins sidechains remains similar in all four examples (**Figure** 9, **Table** 2). These residues are either part of the TIM-barrel β-sheets or the (βα)-loops, with the exception of His88 and Gln164 which are part of domain B.[44]

The GH13 family possesses a conserved catalytic triad the numbering of which is referenced to the corresponding Taka amylase A (TAA) positions.[55] In BaSP this residues are the nucleophile Asp192[56] (TAA: 206) the acid/base catalyst Glu232[57] (TAA: 230) and a proposed transition state stabilizer Asp290[58] (TAA: 297). It was previously suggested that the stabilization by Asp290 is achieved via a strong hydrogen bond to 2-OH of the glucosyl oxocarbenium ion-like transition state.[58, 59] The interaction between Asp192 and the 6-OH moiety of the donor places the second, nucleophilic oxygen of Asp192 ca. 3.0 Å above C1 of glucose in a suitable angle and distance for the following attack. Asp192 belongs exclusively to the -1 site whereas Glu232 and Asp290 interact with OH-moieties of the donor substrate and the leaving group alike (**Figure** 9, **Table** 2).[44] Two further highly conserved residues in

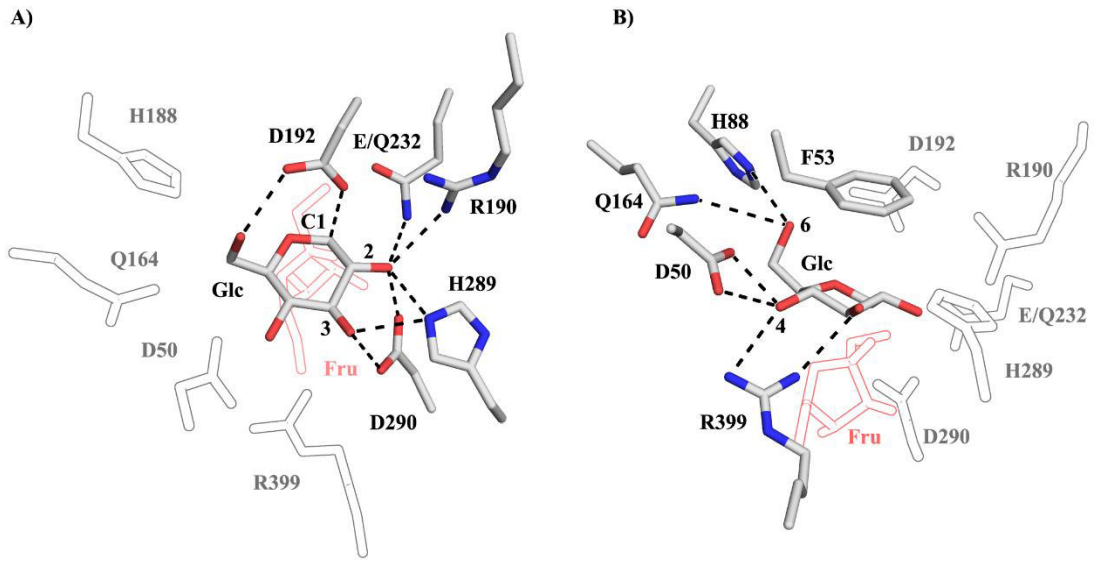


Figure 9 The -1-subsite of BaSP and the coordination of glucose (PDB ID 2gdu). The glucosyl moiety is taken from sucrose which was co-crystallized in the inactive E232Q variant. Asp192 is the catalytic nucleophile, Glu232 the acide-base catalyst. The outline describes the position of the fructosyl moiety of sucrose, the dashes the coordination of the carbohydrates OH-moieties by the residues of the -1 subsite. A) Residues coordinating the 2-OH and 3-OH of glucose, B) Residues coordinating 4-OH, 5-OH and 6-OH of glucose.

Table 2 Hydrogen bond distances between glucosyl moieties in BaSP and -1-subsite residues; the chair conformation values are the averages of the structures 2gdu chain A and B and 2gdv chain B. The values for the boat conformation reflect the crystal structure 2gdv chain A.

OH-moiety glucose	coordinating sidechain	distance chair conformation [Å]	distance boat conformation [Å]
2	*Arg190*	*3.0±0.05*	*3.0*
	Glu232	*3.1±0.2*	*2.9*
	His289	*2.9±0.2*	*3.1*
	Asp192	*2.6±0.2*	*2.7*
3	*His289*	*2.90±0.05*	*3.1*
	Asp290	*2.7*	*2.8*
	Arg399	*3.3±0.05*	*3.3*
4	*Asp50*	*2.6±0.1*	*2.7*
	Arg399	*2.8*	*3.0*
6	*His88*	*2.9±0.05*	*3.0*
	Asp192	*2.8±0.1*	*2.7*

the donor binding site are Arg190 (H-bond with 2-OH, TAA:204) and His 289 (H- bond with 2-OH and 3-OH, TAA: 298).[51]

The active site of GH13 enzymes also contains an aromatic residue, usually a tyrosine or phenylalanine above the donor glucosyl moiety, which contributes to the stabilization of the transition state via strong π-cation interaction with the oxocarbenium ion intermediate or transition state.[49] Additionally it helps to sterically control the positioning of the glucosyl moiety.[49] In BaSP this residue is Phe53.

The acceptor binding site/+1-site (fructose binding conformation)

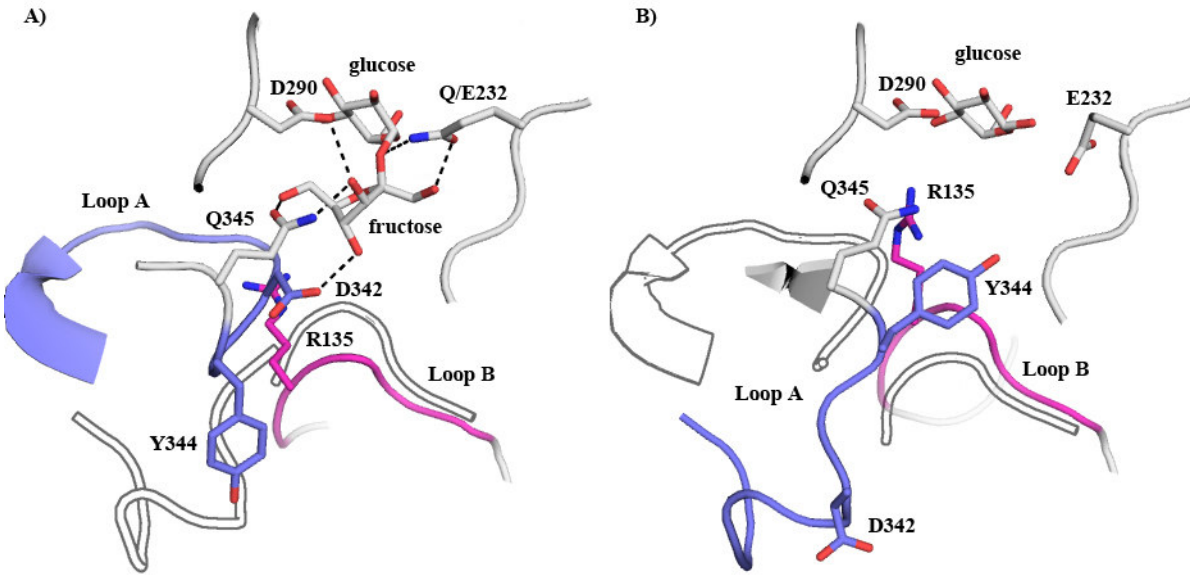


Figure 10 Orientation of loops A (violet) and B (pink) and the active site defining residues resulting in two -1-subsites. A) fructose coordination by BaSP (PDB ID 2gdu) the outlines indicate the orientation of loops A and B in the phosphate binding conformation of BaSP. (PDB ID: 2gdv chain B) B) phosphate binding conformation of BaSP (PDB ID: 2gdv chain B). The outlines indicate the orientation of loops A and B in the fructose binding conformation (PDB ID: 2gdu)

As described in **chapter 1.3.3**, sucrose phosphorylases undergo a rearrangement of loops A (^{336}AAASNLDLY^{344}) and B (^{132}YRPRP^{136}), resulting in two conformations and two distinct active sites (**Figure** 10).[44] Four sidechains are involved in the coordination of the fructosyl moiety: Glu232, Asp290, Asp 342 and Gln345 (**Figure** 10**, Figure 3**).[44] Glu232 and Asp290 are additionally involved in glucose coordination and contribute to the -1 and +1 subsite (**Figure** 9**, Table** 2) whereas Gln345 and Asp342 are exclusively involved in the coordination of fructose. The catalytic acid/base Glu232 forms bidentate hydrogen bonds with the fructosyl moiety via its O2 and 1-OH. The interaction with O2 is required for the protonation/deprotonation step in the catalytic cycle. The H-bond with the 1-OH moiety is considered a substrate assisted facilitation and serves to position Glu232 in a correct orientation for its catalytic purpose.[60] This interaction is considered responsible for the preference of BaSP towards 1,2-diols as acceptor substrates.[60] There is a weaker third hydrogen bond of Glu232 with the 3-OH of fructose.

The hydrogen bond of Gln345 with the 6-OH moiety of fructose plays an important part in the

Table 3 Hydrogen bond distances between the fructosyl moiety in BaSP E232Q (PDB ID 2gdu)

OH-moiety fructose	coordinating sidechain	distance [Å]
1	*Glu232*	*2.6*
3	*Glu232*	*3.4*
	Asp290	*3.0*
	Gln345	*3.0*
4	*Asp342*	*2.8*
6	*Gln345*	*2.8*

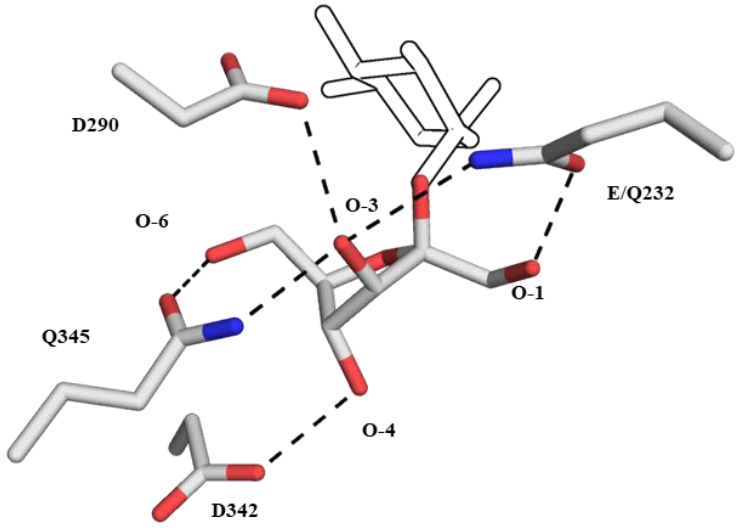


Figure 11 Accommodation of fructose (from sucrose, PDB ID 2gdu) in the -1 –site of BaSP in the *F*-conformation. The dashed lines indicate the hydrogen bonds; the outline shows the position of the glucose moiety.

acceptor substrates specificity of BaSP.[60] Gln345 maintains its position during the loop rearrangement while Asp342 is flexible and only found in the active site of the *F*-conformation (**Figure** 10). In contrast to the *P*-conformation Tyr344 and Arg135 are oriented away from the active site.[44]

Further contributions to fructose recognition come from Tyr196 via hydrophobic interactions with C1 of fructose and Tyr132 has been shown to be important for fructose recognition although its distance from the fructosyl moiety does not permit direct stabilizing interactions.[61]

The acceptor binding site or +1-site (phosphate binding conformation)

In the *P*-conformation loop B (^{132}YRPRP^{136}) moves closer to the -1-subsite and the sidechain of Arg135 and Tyr344 become part of the acceptor binding site while Asp342 moves away (**Figure** 10).[44] No crystal structure of a SP in complex with phosphate exists to date. Kinetic analysis and modelling studies concluded that Arg135 and Tyr344 are vital for the activity towards phosphate and glucose-α-D-1-phosphate and take part in binding interactions with the phosphate group.[61-63] An additional binding interaction comes from Gln345, which is directly adjacent to the flexible loop A, but maintains its orientation during the loop rearrangement.[61]

1.3.5. Exploitation of native reactions catalysed by Sucrose Phosphorylases

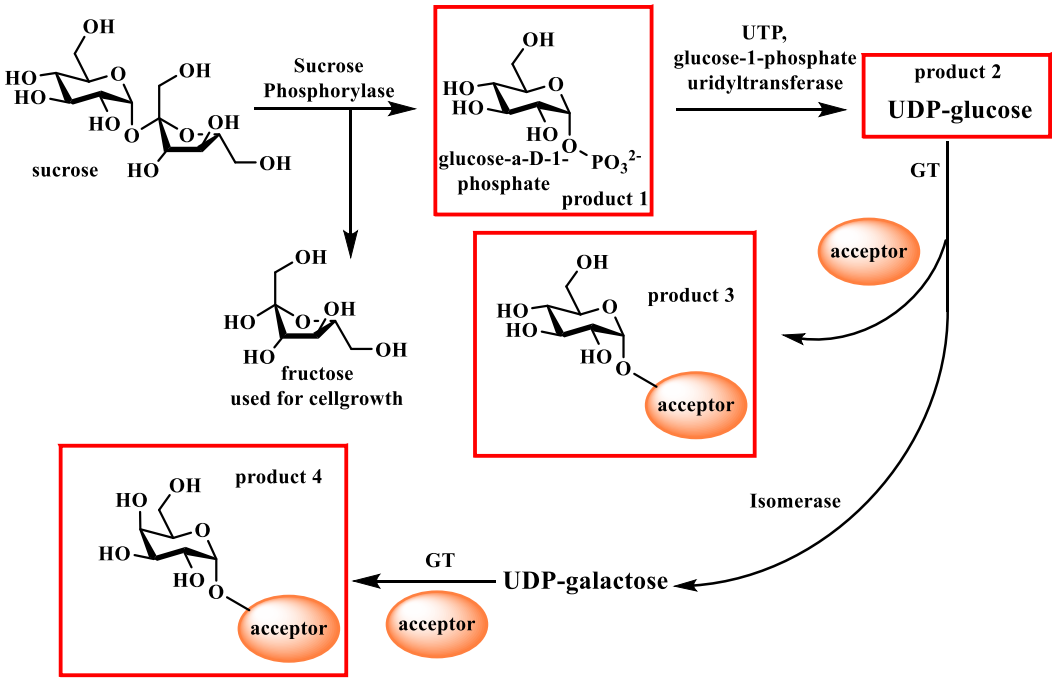


Figure 12 Examples of fermentation strategies exploiting the native SP'reaction resulting in 4 different classes of products (red boxes).

The most straightforward way to utilize SPs in the production of fine chemicals is the synthesis of glucose-α-D-1-phosphate from sucrose.[64] In addition several multi-enzyme

strategies and fermentation strategies for the conversion of sucrose into high value products based on SPs have been developed. In many cases whole cell bio-catalysis is employed as it allows for the efficient and cheap production and regeneration of otherwise expensive co-substrates like UTP. In several of these examples fructose released from sucrose by the SP is utilized by the microorganisms as a carbon or energy source.[65, 66]

The formation of glucose-α-D-1-phosphate is generally followed by its conversion into UPD-glucose which can present the final product.[67] Further conversion into target glucosides like hydroxybenzoate-β-1-D-glucosyl esters[65] or the glucosylated flavonol, astragalin[68] is achieved by co-expression of GT enzymes in the in the microorganisms. This strategy was further expanded by the use of UDP-glucose isomerases and the subsequent GT mediated synthesis of galacto-N-biose[69] and galactosylated or rhamnosylated quercetin.[66]

1.3.6. Non native reactions catalysed by wild type Sucrose Phosphorylases

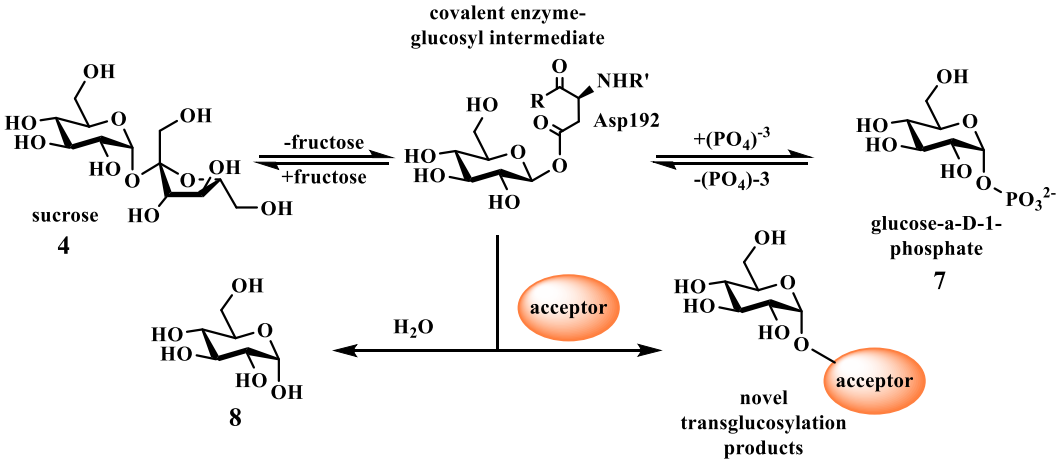


Figure 13 Possible reaction pathways of SPs

Sucrose Phosphorylases are frequently used for the synthesis of novel transglucosylation products. An acceptor nucleophile can intercept the covalent enzyme-glucosyl intermediate resulting in the formation of a novel transglucosylation product (**Figure 13**). Either sucrose or glucose-α-D-1-phosphate can be used as donor substrates. The use of sucrose is more common as it is the cheaper and better available compound. SPs display a degree of substrate promiscuity and over the last decades several classes of acceptor nucleophiles have been glucosylated by wild type SPs from several source organisms. (**Figure 14**).

Amongst those substrates are alternative carbohydrate acceptors like glucose (**8**)[70, 71], D- and L-arabinose (**9**)[70], or psicose (**10**)[72]. The glucosylation of these leads to the production of rare sugars. Since they are very close in polarity, size and shape to the natural acceptor substrate fructose, monosaccharides are the most readily accepted alternative substrates of SPs. It has been indicated that SPs also transfer to di- or trisaccharides,[40] but to the best of my

knowledge the resulting tri- or tetra-saccharides have not been isolated and this transfer reaction has not been studied in detail. It should be added that the active site, found in both crystallized SP conformations, does not contain a +2 subsite and lacks sufficient space for the accommodation of more than one acceptor carbohydrate moiety.

The glucosylation of sugar alcohols and their derivatives, like glycerol (**11**) or D- and L-arabitol (**12**) [40, 70] is the most intensely studied non-native reaction of SPs.[73-77] This is owed to the fact that the transglucosylation product of glycerol, Glycoin®, is used as a moisturizing agent in cosmetics and therefore of interest for industrial production.[73] The stereo-selectivity of SPs in respect to glycerol derivates depends on the substrate. Glucosylation of **13** results in a racemic mixture,[74] in contrast the (R)-enantiomer of **14** is preferred by LmSP and a d. e. of >99% is obtained.[75] Like the carbohydrates from which they are derived the sugar alcohols match the existing polarity of the active site and do not exceed the spatial limitations. Consequently they are accepted rather readily as substrates by SPs.

A range of usually mono-cyclic phenolic compounds like hydroquinone (**15**) and tri hydroxyl benzene (**16**) act as acceptors and SPs have been used to produce α-arbutin, a tyrosinase

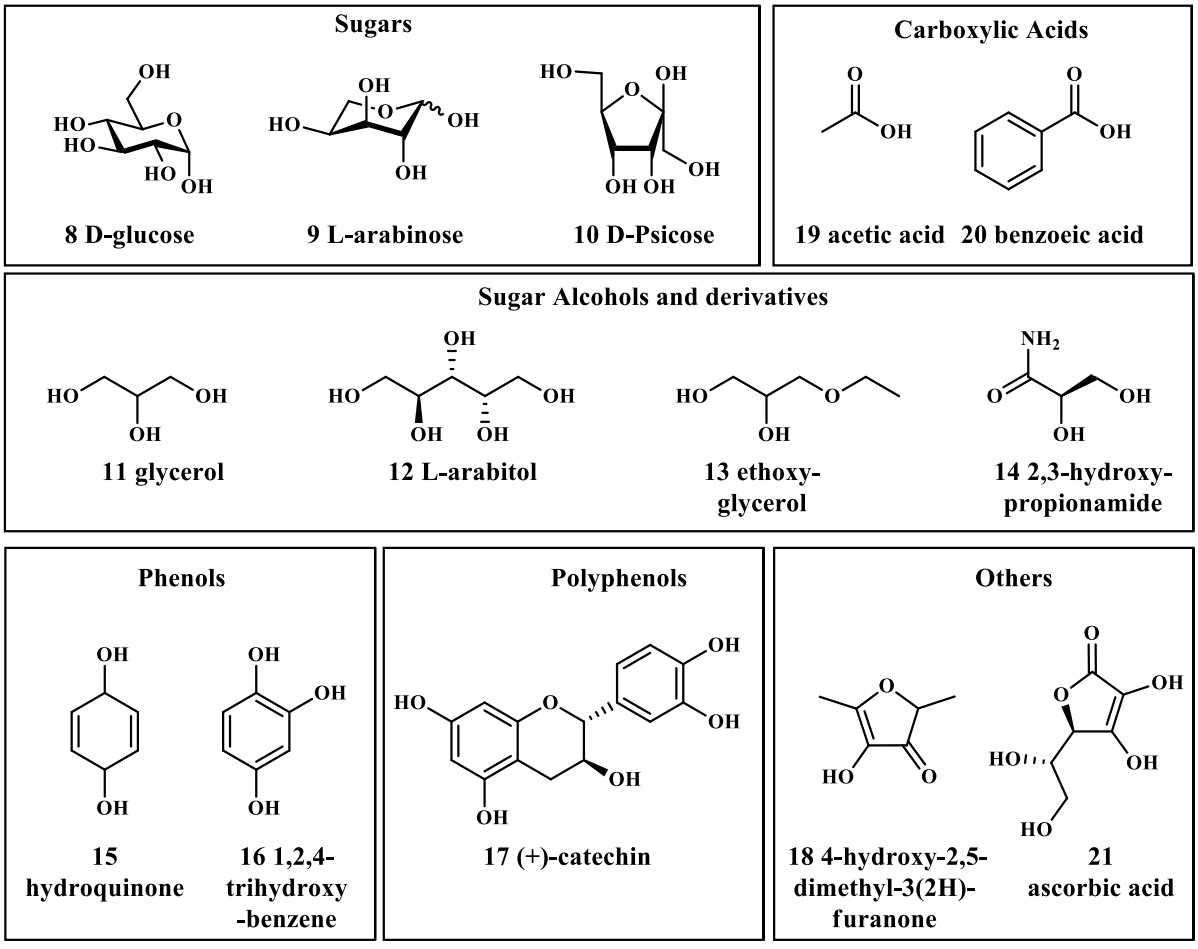


Figure 14 Selection of non-nativ acceptor substrates of SP used in transglucosylation reactions

inhibitor that is used in cosmetics and related compounds.[40, 78] Larger polyphenols like the flavanol catechin (**17**) are unusual substrates. However one example of the successful glucosylation of catechin by *Leuconostoc mesenteroides* sucrose phosphorylase (LmSP) has been reported.[42] Catechin, a tricyclic polyphenol is considerably larger than all of the other substrates discussed here. It does not fit in the active site of BaSP as it is found in the reported crystal structures.[43, 44] Additionally, hetero aromatic furan derivatives like **18** belong to the substrate spectrum of the phosphorylases.[79]

Sucrose Phosphorylases are further able to glucosylate carboxylic acids like acetic acid (**19**)[80, 81] or benzoic acid (**20**).[82] The resulting α-1-glycosyl-esters are however unstable and undergo acyl shifts to form more stable products.[81]

The final example is the glucosylation of ascorbic acid (**21**) which can be targeted either at the position 2 of the furan ring, or at the sidechain depending on the pH value.[83, 84]

1.3.7. Limits of the non-native transglucosylation reactions of SPs

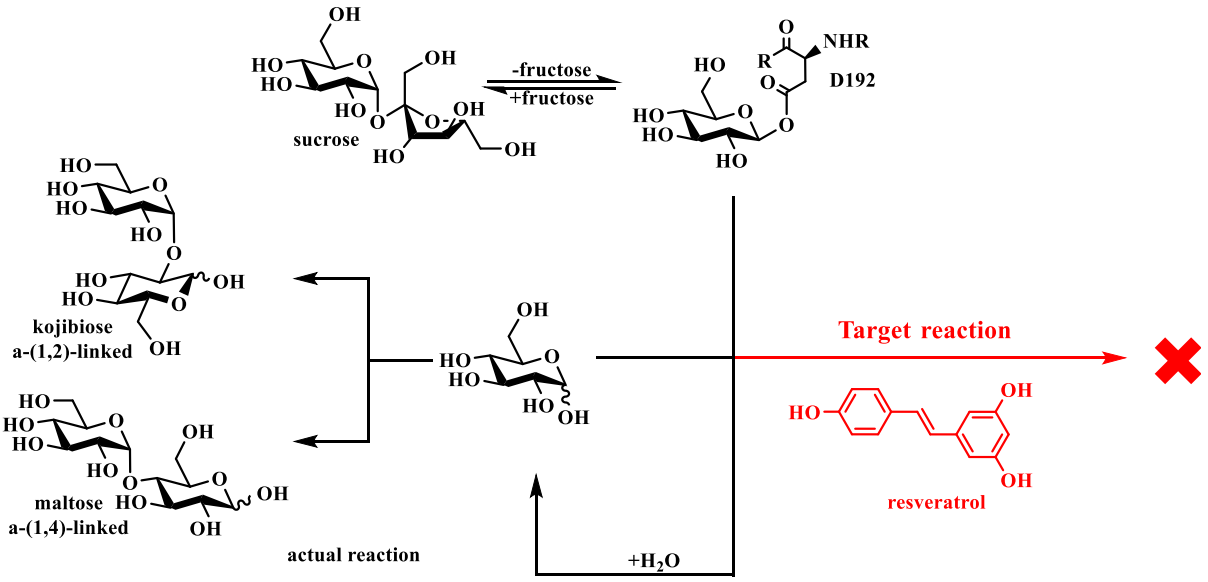


Figure 15 Reaction of BaSP with an inefficient target acceptor substrate. Instead of the glycosylation of resveratrol (red) sucrose (4) is hydrolysed and the resulting glucose (8) outcompetes resveratrol (2) as an acceptor. The main transglucosylation products are glucose-glucose disaccharides kojibiose (22) maltose (24).

While SPs are to a degree promiscuous and a variety of substrates is accepted, several limitations exist. The first challenge is the low affinity of SPs for non-natural substrates. A comprehensive study by Aerts et al. compared the reaction speed of six SPs with 83 acceptor substrates.[40] The efficiency of the substrates was scored by dividing the observed activity of the SP on the donor substrate in presence of an acceptor substrate (v_a) by the hydrolysis rate (v_h). It is necessary to keep in mind that the v_a/v_h ratio does not provide any information about the transfer to hydrolysis ratio observed with a given acceptor. The most

efficient alternative substrates are the sugars arabinose (v_a/v_h: 3.0-8.0) and sorbose (v_a/v_h: 2.6-8.9). The v_a/v_h value for most acceptors was 1-2, whereas BaSP has a v_a/v_h of 44 for phosphate and 38 for fructose.[40] Of the 27 non-sugar or sugar alcohol substrates tested only salicylic acid (v_a/v_h: 1.8) showed an increased reaction rate of BaSP. These findings underline the need to adapt the +1-sites of SPs in order to create efficient catalysts for most substrate classes, in particular for aromatic acceptor molecules.

One major complication is the release of glucose through sucrose hydrolysis. In general, trans-glycosidases have a less efficient transition state stabilization for the attack of the acceptor substrate on the covalent enzyme-donor intermediate ("deglycosylation") than glycosyl hydrolases. As a result TG's usually exhibit k_{cat}-values that are decreased by up to four orders of magnitude compared to structurally related GH enzymes.[39] A degree of transition state stabilization in TGs originates from interactions between the acceptor substrate and the +1-site. The result is a lower energy transition state for trans-glycosylation than hydrolysis and a preference for trans-glycosylation over hydrolysis. An acceptor that cannot exploit constructive interactions has to statistically compete with the ubiquitous potential acceptor water and hydrolysis becomes the preferred reaction pathway.[39] Most of the alternative acceptors discussed above are glycosylated inefficiently so that sucrose hydrolysis is the dominant reaction pathway. The straightforward approach to deal with this issue is the use of a significant excess of donor substrate and in this case yields are generally considered only as the conversion of acceptor substrate into product.

However, the release of glucose via hydrolysis presents a further complication. Glucose can be used as an acceptor by SPs (**Figure** 15). The linkage of the produced glucose-glucose disaccharides depends on the sucrose phosphorylase. Whereas LmSP synthesizes kojibiose (**22**) and nigerose (**23**), BaSP produces kojibiose (**22**) and maltose (**24**).[71, 85] While the synthesis of rare disaccharides from cheap and abundant starting materials like sucrose and glucose is a useful method, this is a serious issue for the synthetic application of SPs with low affinity substrates. If the desired acceptor substrate is inefficient and therefore a considerable portion of sucrose is hydrolysed the desired transfer reaction is usually outcompeted by the emerging glucose and glucose-glucose disaccharides that are formed instead (**Figure** 15). This complication is not limited to sterically demanding substrates, but has been observed with smaller compounds like ethoxy-glycerol (**13**).[74] Even if the desired product is obtained the excess of different sugars in the reaction mixture often complicates product purification.

1.4. Mutagenesis strategy to enable the glucosylation of polyphenols by BaSP

1.4.1. Previous mutagenesis studies

Previous mutagenesis studies focused primarily on mechanistic aspects of the native reactions of sucrose shosphorylases and elucidated which residue was involved in each step.[49, 57, 59, 62, 86] A complete alanine scanning of all acceptor binding site residues was published during the course of this work focussing on the impact on the natural reactions.[87] Two studies focused on the creation of SP variants with increased thermostability.[88, 89]

There is one example of a variant with a novel reactivity. Both the nucleophile Asp196 and the catalytic acid/base Glu237 were substituted and the resulting D196N-E237Q variant produced a β-linked glucosyl azide from glucose-α-D-1-phosphate and sodium azide with a very low turnover number of $3.6x10^{-5}s^{-1}$.[90]

1.4.2. Mutagenesis Strategy of this work

Basic considertations

The main limitations for the successful glucosylation of aromatic compounds like resveratrol by SPs are their relative size, their polarity and the resulting low or virtually non-existing affinity for the +1-subsite of SPs. The active sites of both BaSP conformations do not permit the accommodation of resveratrol. The slow but existing transfer to the isoflavon daidzein(**Figure 16**), which is slightly larger than resveratrol, suggests an existing natural mechanism to overcome this limitation.[40] Potentially this substrate is accommodated via an open conformation of BaSP that arises during the loop rearrangement. However, both aromatic substrates do not match the active site polarity and are inefficiently coordinated. This could be alleviated by the generation of a less polar active site via the exchange of polar amino acids against non-polar or aromatic residues.

It was decided to introduce the aromatic amino acids histidine, phenylalanine, tyrosin and tryptophan at an appropriate position into the active site of BaSP. These four amino acids were chosen due to their potential to introduce π-π-interactions with the aromatic substrates in addition to the predicted polarity change.

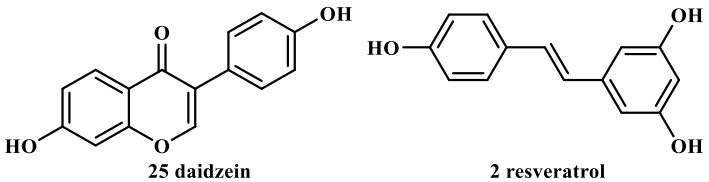


Figure 16 Poor acceptor substrates of SPs daidzein (25) and resveratrol (2).

Target amino acid

The above mentioned loop flexibility (**chapter 1.3.3** and **1.3.4**) poses an additional challenge for the mutagenesis strategy as it is not known which loop orientation will be relevant for aromatic compound glucosylation. It is further conceivable that neither the *F*- nor the *P*-conformation will be suitable for the accommodation of the sterically demanding substrates. If parts of the flexible loops are targeted the desired change may not be present in the active site of the conformation that offers the required space for the aromatic acceptors. Therefore, the target amino acid should be present in the active sites in both loop conformations.

A further criterion that needs to be met is an appropriate distance to the -1-site. A serious disturbance of the acceptor glucose coordination, i. e. through an exchange of Asp290, will lead to a collapse of the activity. Likewise introducing steric restrictions in the -1-site would render the enzyme inactive.

A partition of the exchanged residue in fructose coordination was considered to be simultaneously beneficial and problematic. Disturbing the fructose coordination will reduce the affinity to sucrose and impact the activity. On the other hand it is very likely that undesired, competing acceptor glucose is coordinated via the fructose binding conformation. A loss of hydrogen bonds to fructose may translate to the loss of hydrogen bonds with the unwanted acceptor glucose and further improve the selectivity for the target aromatic acceptors.

Gln345 satisfies these requirements best. It interacts with fructose via it's OH-3 and OH-6 but remains far enough from the -1-site. Therefore, Gln345 was chosen as the mutagenesis target and the four variants Q45H, Q345F, Q345Y and Q345W were constructed. It was envisioned that the aromatic ring would adopt an orientation comparable to Gln345 in the

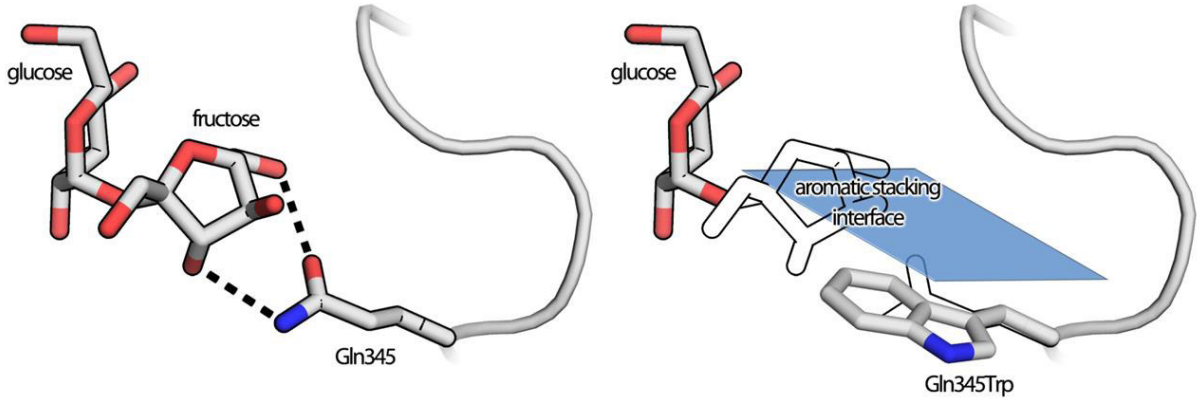


Figure 17 Visualisation of the mutagenesis strategy A) Interactions between Gln345 and the fructosyl moiety in BaSP (PDB ID 2gdu) B) Theoretical Q345W variant obtained via the mutagenesis wizard in pymol. The blue area indicates a potential binding site for aromatic residues. The outlines indicate the position of the fructosyl moiety and Gln345 in the wild type.

wildtype and that the acceptor aromatic compounds could occupy a position in parallel to the introduced aromatic sidechain and attack the covalent glucosyl-enzyme intermediate (**Figure** 17).

The initial tests revealed Q345Y and Q345W to be almost inactive, while Q345H behaved like the wildtype and did not present good levels of transfer to the aromatic substrates. The variant Q345F displayed a reduced activity towards sucrose and simultaneously a preference for the target substrates. The characterisation of this variant and the elucidation of the underlying structural changes are discussed in the following chapters.

1.4.3. Mutagenesis studies in parallel to this work

During the course of this work two examples of SPs with altered acceptor substrate specificity have been presented.

In the first study, published by Dirks-Hofmeister et al. in 2015[91], the transfer to larger polyphenolic substrates was achieved through an exchange of Arg134 (the equivalent to the phosphate coordinating Arg135 in BaSP) in *Thermoanaerobacterium thermosaccharolyticum* sucrose-6'-phosphate phoshphorylase (TtSPp). The transfer was enabled through the creation of space by exchanging the large arginine against alanine, valine and threonine. The fastest reaction was achieved with the R134A variant and the affinity for resveratrol is lower than for the natural substrates with K_M values between 56 and 185 mM.[91]

The second study, published in 2016 by Verhaeghe et al.[85], focused on the optimization of kojibiose production. The focus of this study was the flexible loop A. The double mutation L341I-Q345S provided a catalyst with a 95% selectivity for kojibiose over maltose.[85]

Chapter 2

REDESIGN OF THE ACTIVE SITE OF SUCROSE PHOSPHORYLASE BY A CLASH INDUCED CASCADE OF LOOP SHIFTS

Michael Kraus, Clemens Grimm and Jürgen Seibel

Published in ChemBioChem, volume 17, pages 33-36, 2009.

Summary

In order to obtain a SP variant with a preference for polyphenolic substrates, Gln345 of BaSP was chosen for mutagenesis and exchanged against all four aromatic amino acids, His, Phe, Tyr and Trp. The binding of aromatic acceptor substrates was envisioned via π-π-interactions introduced by the mutation. Of the four variants one, BaSP Q345F was able to perform the glucosylation of polyphenols in the initial screening reactions and maintained a suitable activity (8.6 % compared to the wild type) towards sucrose. This variant was chosen for further investigation. As a proof of concept for the new trans-glucosylation capability three polyphenols, (+)-catechin, (-)-epicatechin and resveratrol were chosen as acceptors. All yields are >80%, resveratrol is quantitatively converted into the glucoside.

To gain further insight into how the mutation affects the enzyme and how it enables the novel reaction a crystal structure of BaSP Q345F was solved. An unexpected domain shift was discovered that enlarges the active site and creates an access channel. The distortions in the glucose binding region provide an explanation for the lower activity towards sucrose and the increased space is needed for the accommodation of the polyphenols. As none of the acceptor substrates are present in the crystal structure the mechanism of the aromat accommodation needs to be addressed with further experiments. (See **Chapter 4**)

2.1. Abstract

Sucrose phosphorylases have been applied in the enzymatic production of glycosylated compounds for decades. Yet several desirable acceptors, such as flavonoids or stilbenoids, which exert diverse antimicrobial, anticarcinogenic or antioxidant properties, remain poor substrates. The Q345F exchange in sucrose phosphorylase from Bifidobacterium adolescentis allows efficient glucosylation of resveratrol, (+)-catechin and (-)-epicatechin in yields up to 97% whereas the wild type enzyme favors sucrose hydrolysis. Three previously undescribed products are made available. The crystal structure of the variant reveals a widened access channel with a hydrophobic aromatic surface that is likely to contribute to the improved activity towards aromatic acceptors. The generation of this channel can be explained by a cascade of structural changes arising from the Q345F exchange. The observed mechanisms are likely to be relevant for the design of other tailor-made enzymes.

2.2. Introduction

Plant polyphenols such as stilbenoids or flavonoids are in the focus of interest due to their antimicrobial[92-94] and antitumor[12, 95] activities and their role in lifespan/healthspan extension[96]. However, they often suffer from low bioavailability due to their poor water solubility.[97, 98] Glycosylation of natural products is a general strategy to improve their water solubility and biochemical or pharmaceutical properties.[99] In this context resveratrol - one of the most popular stilbenoids - was selected as a target of glycosylation. So far, besides a few exceptions, enzymatic glycosylation usually relies on glycosyltransferases (GT) which require expensive nucleotide diphosphate activated sugars.[99] In contrast, glycosidases and transglycosidases (GH) are a class of alternative enzymes that use cheap substrates, making them suitable for industrial use, but suffer from a limited acceptor range.[100] In recent studies redesign of enzymes either through random or rational approaches have been used to overcome these limitations.[91, 101-104] Here, we describe the structure-based redesign of sucrose phosphorylase (EC 2.4.1.7 GH13)[45] from *Bifidobacterium adolescentis* (BaSP) for efficient glycosylation of polyphenolic substrates.

2.3. General Strategy

BaSP was chosen for its compatibility with organic solvents and relative thermostability.[89] In addition, crystal structures for BaSP are available, which allow a rational approach for mutagenesis.[43, 44] Sucrose phosphorylases transfer glucose moieties to either fructose or phosphate.[44] The transfer to unnatural acceptors including stilbenoids and flavonoids has been observed, but is generally highly inefficient, most likely because the latter are larger

and less polar than the natural substrates[40, 42, 105]. In a nonconventional strategy, we aimed to introduce new non-polar interactions between the enzyme and the desired substrates. We envisaged that the introduction of an aromatic side-chain into the active site of the enzyme may enable π-π-stacking mediated coordination of the acceptor substrate. Introducing a larger sidechain should also cause rearrangement of adjacent residues, like the flexible Tyr344, resulting in a larger binding pocket[44]. Thus, Gln345 was chosen for the exchange with phenylalanine. It is located in short distance from the acceptor binding site (+1-site). Interactions of Gln345 with the donor substrate sucrose is limited to OH-3 and OH-6 of its fructose moiety.[44] Therefore the distance to the glucose binding pocket is sufficient and interference with sucrose binding capability should be tolerable. Minor interference with carbohydrate binding in the +1 site was intended in order to inhibit the unwanted transfer to glucose observed with wild type BaSP. The enzyme is known to undergo structural changes during its catalytic cycle.[44] Crystal structures of the sucrose binding conformation (PDB ID code 2GDU) as well as the phosphate binding conformation (PDB ID code 2GDV) were published.[44] The conformation of Gln345 is largely unaffected by the structural rearrangements observed between both conformations. We therefore reasoned that the introduced phenylalanine may likewise occupy similar positions in both conformations.[44] This is of interest as it is not known which conformation binds aromatic acceptors. Previously, a Gln345Ala variant was shown to poses a moderately decreased affinity for fructose and strongly decreased affinity for phosphate while no increase in the glucosylation of the aromatic compound pyridoxine was observed.[61]

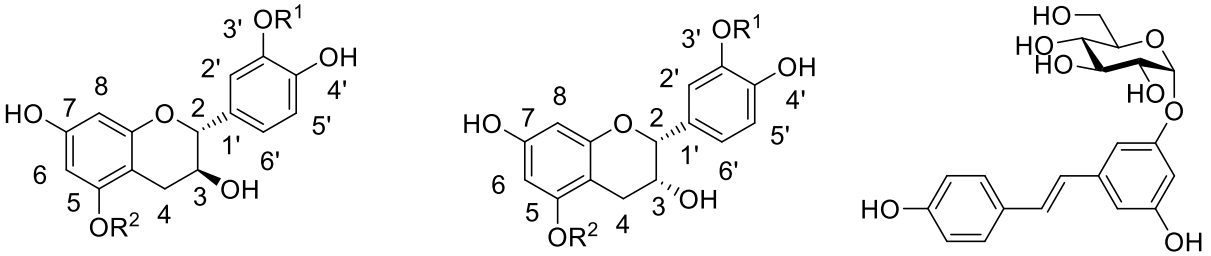
2.4. Results and Discussion

2.4.1. Enzyme expression and activity assay

BaSP wild type and the Q345F variant were expressed as *N*-terminally hexa-histidine tagged proteins allowing affinity purification via standard protocols. To avoid undesired phosphorolysis, reactions were carried out in MOPS buffer. Sucrose hydrolysis was used to determine activities, as no transfer reaction suitable for both wiltype enzyme and Q345F variant exists. Specific activities of 0.716 U/mg for the wild type and 0.062 U/mg for the variant were observed. Despite the fact that the targeted Gln345 is involved in substrate binding, the Q345F variant retains 8.6% of the specific activity of the wild type towards sucrose. For the variant under the same reaction conditions, a K_M-value for sucrose of 17.5 ± 1.04 mM was determined. The wild type enzyme reaches V_{max} at sucrose concentrations around 1 mM or lower; reactions with lower sucrose concentrations were not investigated due to assay limitations. BaSP can be efficiently produced by bacterial overexpression with a yield of 50 mg purified protein per liter culture medium. In addition

immobilization techniques for this enzyme have been reported, enabling biocatalyst reuse.[105, 106] Therefore the low specific activity of BaSP Q345F towards the donor substrate is not considered a severe drawback for synthetic application.

2.4.2. Glycosylation of polyphenols by BaSP Q345F



26: R^1 = Glc, R^2 = H; 82%
27: R^1 = R^2 = Glc 12%

28: R^1 = Glc, R^2 = H 7%
29: R^1 = H, R^2 = Glc 5%
30: R^1 = R^2 = Glc 86%

31: 97%

Figure 18 Product spectrum and yields obtained with BaSP Q345F using 600 mm sucrose as donor: 1,2: 100 mm (+)-catechin, 3-5: 150 mm (-)-epicatechin, 6: 75 mm resveratrol Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission.

To test for improved selectivity of polyphenols resveratrol, (+)-catechin and (-)-epicatechin were chosen as acceptor substrates. Reactions were carried out at 37 °C with 30% DMSO as a co-solvent to improve acceptor solubility. While optimal temperatures of 48 °C[70] and 60 °C[106] are reported for BaSP, prolonged incubation at these temperatures with high concentrations of organic solvents inactivates the protein.

BaSP Q345F produced (+)-catechin-3'-*O*-α-D-glucoside (**26**) from (+)-catechin and sucrose in 80% yield. An additional, previously not reported product, (+)-catechin-3',5-*O*-α-D-diglucoside (**27**) is produced with up to 24% yield. Reactions of the wild type enzyme under the same conditions afforded no significant amounts of either product **26** is an inefficient acceptor for further glycosylation by the Q345F variant. While sufficient amounts of diglucoside **27** can be produced from catechin, a reaction using purified **26** as sole acceptor displayed only minor amounts of **27** and significant amounts of sucrose hydrolysis and glucose disaccharide formation were observed. The NMR spectrum of **27** reveals a side product, which could not be fully characterized but might be the region-isomer with glucose attached to position 7 of the flavonoid.

A reaction of (-)-epicatechin and sucrose catalysed by the Q345F variant yielded three detectable products: (-)-epicatechin-3'-α-*O*-D-glucoside (**28**), **(**23%**)**, (-)-epicatechin-5-*O*-α-D-glucoside (**29**) (29%) and the diglucosylated product epicatechin-3',5-*O*-α-D-diglucoside (**29**). While **28** was previously reported, no entries in the SciFinder database exist for either **29** or **30**[107]. Again, the NMR spectra of **30** reveal roughly 8% of so far uncharacterized side

products, presumably additional regio-isomers. To a minor extent yet uncharacterized side products also appear accompanying the **29** monoglucosylation product.

The variant Q345F converts resveratrol into resveratrol-3-*O*-α-D-glucoside (**31**), a natural product found in *Eleutherococcus brachypus*.[108] Yields up to 97% of **31** were obtained with BaSP Q345F as compared to 4% for the wild type enzyme. Glucosylation of resveratrol merely constitutes a side reaction for wild type BaSP. HPAEC analysis revealed a preference of the wild type for sucrose hydrolysis and consequently the production of the glucose disaccharides kojibiose and maltose together with an uncharacterized saccharide. Under initial reaction conditions the Q345F variant uses 90% of the sucrose consumed for transfer to resveratrol and only about 10% hydrolysis occurs. Increased hydrolysis and disaccharide formation is observed only after most resveratrol is glucosylated. Thus the Q345F variant uses sucrose far more efficiently for transfer than the wild type. Product **31** is in contrast to resveratrol an inefficient substrate for BaSP Q345F. Yet if the reaction was allowed to continue after all resveratrol was monoglucosylated two further products are detected in low amounts of 1.5% and 2.8%. Those were not isolated, however the retention time suggests that at least one is a diglucosylation product.

2.4.3. Crystallographic investigation of BaSP Q345F

To determine the mechanism responsible for the altered catalytic properties of the Q345F variant, crystals of BaSP Q345F were grown in the presence of sucrose and a crystal

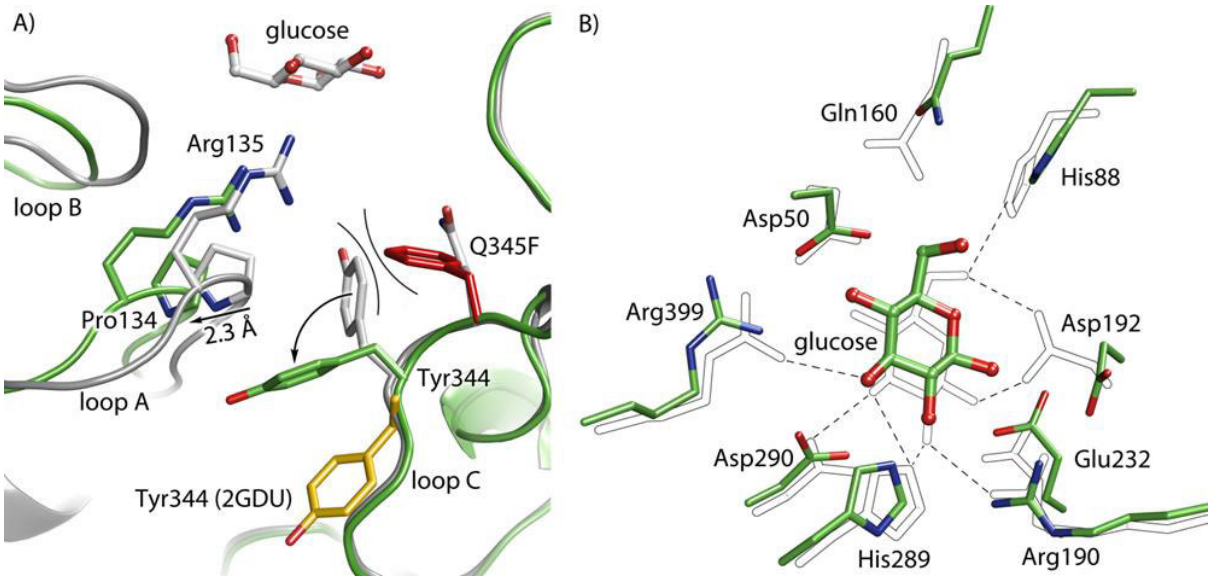


Figure 19 Structural changes in the active site of BaSP Q345F A) Steric hindrance induced shifts in BaSP Q345F. Alignment of BaSP Q345F (5C8B, green) and the wild type (PDB ID code2GDV, chain B, white) Steric hindrance induced by Phe345 (red) causes the rotation of Tyr344. Loss of the interaction between Pro134 and Tyr344 causes loop A to shift away from loop C, resulting in a wider access channel. The shift of loop B (residues 154-159, Table 4) provides additional space in the active site. (PDB ID code 2GDV) Additionally Tyr344 in sucrose binding conformation of BaSP wild type is shown (PDB ID code 2GDU chain A, yellow) b: Differences in glucose coordination; alignment of BaSP Q345F (PDB ID code: 5C8B, green) and the wild type (PDB ID code: 2GDV, chain B, white) In the Q345F variant the hydrogen bonds between glucose and His88, Gln160, Asp192 and E232 are disrupted. Dashed lines: disrupted H-bonds from the wild type. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission.

Table 4 Key regions in BaSP Q345F

Structural motif	Loop A	Loop B	Loop C	*β-sheet A*
Residues	*133-137*	*154-159*	*336-344*	*88-91* *160-162*
Shift of C_α	*2.5-3.7 Å*	*2.4-3.3 Å*	-	*1.8-2.1 Å*

structure in complex with glucose (a hydrolysis product) with a resolution of 2.7 Å was solved (PDB ID code: 5C8B). Only conformations corresponding to the presumed phosphate binding conformation (PDB ID code: 2GDV chain B, in complex with β-D-glucose) were observed with β-D-glucose bound to the glucose binding site.

The most pronounced differences between the crystal structure of the Q345F variant and the wild type enzyme are found in the region spanning residues 86-166 (domain B) that is shifted relatively to the rest of the protein.[44] The rearrangement of loops A and B as well as the β-sheet A (**Table** 4, **Figure** 19**A**, **Figure** 20) are responsible for the altered catalytic properties of BaSP Q345F. While many of the hydrogen bonds responsible for glucose coordination in the active site are disrupted, the substrate orientation remains conserved (**Figure** 19**B**). However the bound glucose is shifted towards β-sheet A by 0.9 Å. While hydrogen bonds to Asp50, Asp290 and Glu232 still remain all other hydrogen bonds observed in the wild type structure (2GDV) are disrupted (**Figure** 19**B**).[44] In comparison to the wild type structure, in BaSP Q345F the carboxyl group of Asp192, constituting the catalytic nucleophile, is rotated away by 115° from the C1 atom of glucose as the χ1 angle of the residue changes from *trans* to *gauche*$^+$ conformation. In addition, His88 and Gln160, both involved in 6-OH coordination

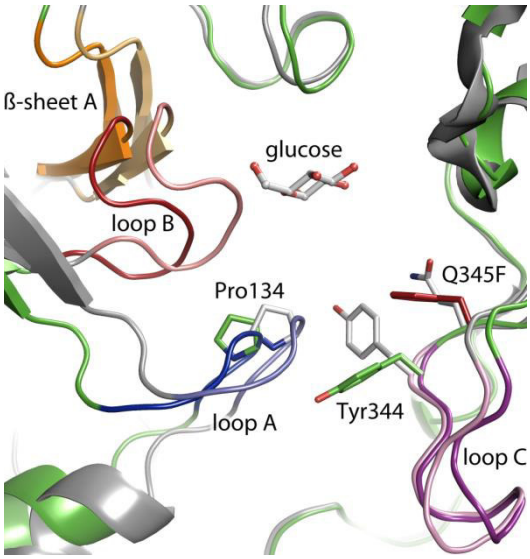


Figure 20 Alignment of BaSP Q345 green (5C8B) and wildtype enzyme grey (2GDV): The shift of loop A (blue) loop B (red) and β-sheet A (orange) is shown. The shifted part of C is a flexible region located at the proteins surface. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission.

are likewise shifted away from the bound glucose. Finally the catalytically essential Glu232 is shifted by 0.9 Å (**Figure** 19**B**). These changes explain the decreased activity towards and affinity for sucrose. As no crystal structure of the Q345F variant in complex with sucrose was obtained, it remains elusive how the loss of hydrogen bonds between the fructose moiety of sucrose and Gln345 affects binding of the substrate. However it is likely that a weakened interaction with this substrate due to the loss of OH-bonds between Gln345 and the fructose moiety contributes to the poor sucrose binding.

Tyr344 plays a key role in the wild type catalysis and undergoes major structural rearrangements during the catalytic cycle along with the entire loop C.[44] In the Q345F variant, Tyr344 is not rotated as far out of the active site as in the sucrose binding conformation of the wild type enzyme (2GDU, **Figure** 19**A**). In addition, loop C (**Figure** 19**A**, **Figure** 20) of BaSP Q345F is found in an orientation that is characteristic for the phosphate binding conformation. Steric hindrance due to the exchange of Gln345 against Phe forces the neighbouring Tyr344 to rotate by 85° out of the active site (**Figure** 19**A**). In the phosphate binding conformation of the wild type (2GDV chain B) Van der Waals interactions of Tyr344 with Pro134 are present.[44] Rotation of Tyr344 abolishes the interaction with Pro134 and favours the shift of loop A. As a consequence, also the more distant loop B and β-sheet A are repositioned as well as the rest of the domain B. A wider access channel, capable of accommodating the large polyphenolic acceptors is the result of the combined steric hindrance induced shifts. (**Figure** 19**A**, **Figure** 21) Our crystal structure of BaSP Q345F proves that the introduction of a sterically demanding residue can result in an enlarged active site by indirect effects. Furthermore, additional functional changes may be introduced by this exchange, as Tyr344 and Phe345 create an aromatic surface at the opening of the active site which could coordinate the desired aromatic acceptors via π-π-stacking. (**Figure** 21) Our study adds to the recent, interesting findings of Desmet *et al*.[91]

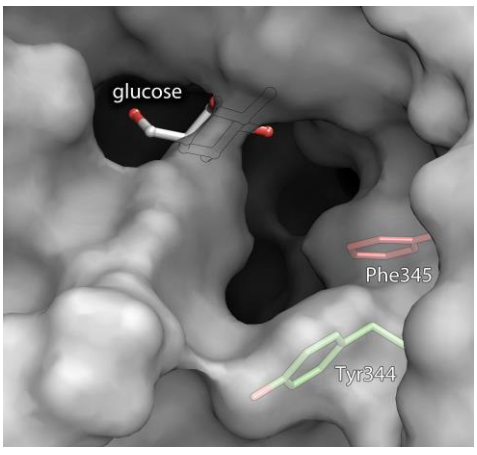


Figure 21 Access channel of BaSP Q345F displaying the non-polar aromatic surface created by Phe345 and Tyr344 Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission.

2.5. Conclusion

In order to create a new tool for the glucosylation of polyphenolic phytochemicals, we introduced an aromatic amino acid as a potential partner for π-π stacking into the active site of sucrose phosphorylase. In the Q345F variant the access channel is not enlarged by the conventional strategy of replacing a larger side chain with a smaller one. In contrast, a cascade of conformational changes is induced by the replacement of Gln345 with a spatially more demanding Phe residue. By this means, glucosylation of aromatic acceptors with yields up to 97% is enabled. In addition the three previously undescribed compounds, **27**, **30** and **28** are now available via enzymatic glucosylation. The structural data elucidates further how a single amino acid exchange affects enzymatic function. We are confident that these insights aid in further development of tailor made biocatalysts.

2.6. Experimental section

2.6.1. Materials and chemicals

(+)-Catechin hydrate, (-)-epicatechin were purchased from Sigma-Aldrich, resveratrol from Carbosynth, all other chemicals were purchased from Sigma-Aldrich or VWR. Solvents for chromatographie were distilled prior to use, all other chemicals were used without further purification. Ni-NTA resin was a product of Sigma Aldrich, PCR primers were ordered from Sigma-Aldrich.

2.6.2. Cloning of BaSP wild tape and BaSP Q345F

Freezedried cultures of *B. adolescentis* (DSMZ 20083) were obtained from DSMZ (Deutsche Sammlung von Mikroorgansimen und Zellkulturen GmbH), and grown under anaerobic conditions in DSMZ medium Nr.58 without resazurin. Cells were harvested and the genomic DNA isolated, using a GenJet Genomic DNA purification Kit (Thermo Fisher). The BaSP gene was amplified from genomic DNA using the primers

5’-ATAACCATGGCTATGAAAAACAAGGTGCAGCTCATCAC-3’ and 5’-CAATCCGCCTGTCGTCGCCCTCGAGTAAT-3’. The amplicon was inserted into pET-28b(+) using the NcoI and XhoI restriction sites yielding plasmid pET-28b(+)-BaSP-wt.

The Q345F mutation was constructed applying the Megaprimer method. The mutagenic primers 5’-CCAATCTCGACCTCTACTTCGTCAACAGCACCTAC-3’, and 5’-CAATCCGCCTGTCGTCGCCCTCGAGTAAT-3’ were used for the creation of the megaprimer and 5’-ATAACCATGGCTATGAAAAACAAGGTGCAGCTCATCAC-3’ was used

for the second PCR. Cloning and purification of the variant followed the procedure described for the wildtype.

2.6.3. Expression and purification of BaSP wild tape and BaSP Q345F

E. coli Bl21 star™ cells were heat shock transformed with plasmid pET-28b(+)-BaSP-wt. Overnight cultures of the transformed host in LB-medium containing 50 mg/l kanamycin sulfate were grown and 1.8 ml were used to inoculate 250 ml of LB-Medium (50 mg/l kanamycin sulfate). The cultures were incubated at 37 °C and 180 rpm until they reached an OD_{600} of 0.6, at which point the temperature was adjusted to 19 °C and IPTG was added to a final concentration of 0.5 mM. The cells were grown for additional 18 hours after which they were harvested by centrifugation (4000 g for 10 min). The sediment was resuspended in lysis buffer (60 mM phosphate, 250 mM NaCl, 11 mM imidazol, 5 mM β-mercaptoethanol pH=8). Cells were lysed using a sonifier and centrifugated at 17000 g for 10 min at 4 °C. The lysate was loaded onto 0.5 ml Ni-NTA columns equilibrated with lysis buffer and incubated at 4 °C and slow rotation for a minimum of 2 hours. The column was washed with 2.5 ml of lysis buffer and the protein was eluted with 1.5 ml of elution buffer (60 mM phosphate, 250 mM NaCl, 230 mM imidazol, 5 mM β-mercaptoethanol pH=8). The buffer was exchanged to 20 mM MOPS-buffer (pH=7) using 5 ml Hi-Trap columns from GE Healthcare.

2.6.4. Enzyme activity assays

General Procedure of the BCA-Assay

Solution A: 1.52 g 2,2′-Biquinoline-4,4′-dicarboxylic acid dipotassium salt trihydrate and 62.3 g Na_2CO_3 were dissolved in 1 l H_2O. **Solution B:** 3.50 g aspartic acide, 5.00 g Na_2CO_3 and 1.70 g $CuSO_4$*$5H_2O$ were dissolved in 150 ml H_2O.

Solutions A and B were stored at 4 °C and excluded from light. Minimum 1 h prior to use 31.1 ml Solution A, 0.92 ml Solution B and 8.00 ml of absolute EtOH were combined and kept at ambient temperature under the exclusion of light.50 µl of inactivated diluted reaction mixtures were added to 150 µl of BCA-solution in 96-well plates. The plates were tightly covered with adhesive aluminium foil and heated in an oven at 70 °C for a minimum of 75 min.

Activity assays:

Activity assays were performed at 37 °C in 100 mM MOPS-buffer at pH=7 in at total volume of 200 µl. 200 mM Sucrose was used as sole substrate. A final concentration of 0.36 g/l for Q345F and 0.066 g/l for the wildtype was used. After 0 min, 6 min, 12 min and 18 min 20 µl samples were diluted with 180 µl H_2O and inactivated at 95 °C for 6 min and centrifuged to remove precipitated proteins. Standards containing 25 µM to 500 µM glucose and fructose

were treated in the same manner as the samples. 50 µl of samples or standards were added to 150 µl BCA solution in 96-well plates. The plates were covered tightly with adhesive aluminum foil and incubated at 70 °C for 75 min. Concentrations of glucose and fructose release were determined at 540 nm using a Tecan Sunrise well plate reader. The assay was performed in quadruplets. Origin Pro 9.1G was used for data procession. To determine the K_M-value of BaSP Q345F reactions with 2 mM, 4 mM, 8 mM, 15 mM, 25 mM and 75 mM of sucrose were performed with conditions otherwise identically to those above.

Table 5 Kinetik Parameters of BaSP Q345F

v_{max} [U/mg]	v_{max} standard deviation	K_M [mM]	K_M standard deviation
$66.0*10^{-3}$	*$1.43*10^{-3}$*	*17.5*	*0.949*

2.6.5. Glucosylation of (-)-epicatechin, resveratrol and (+)-catechin.

40 ml of reaction mixtures contained 100 mM resveratrol 150 mM, (+)-catechin or (-)-epicatechin, 30% DMSO, 50 mM MOPS Puffer pH=7 and 800 mM sucrose. Cell lysates containing BaSP Q345F were used without further purification. The reactions were incubated at 37 °C under slight agitation and monitored via thin layer chromatographie. The reaction was stopped through heating at 95 °C for 15 min and precipitated protein was removed via centrifugation at 6000 g for 15 min. The solvent was removed and the residue purified via chromatographie on silica. Using water:isopropanol:ethylacetate 1:3:6. A second chromatographic purification using methanol:ethylacetate 1:12 was used to separate the monoglucosylated products from the unglucosylated acceptors. All products were characterized via NMR and mass spectrometry.

2.6.6. Determination of transglucosylation yields

(+)-Catechin

200 µl reaction mixtures containing 50 mM, 100 mM or 150 mM (+)-catechin, 30 % DMSO, 600 mM sucrose, 100 mm MOPS buffer pH=7 and 0.85 g/l BaSP Q345F were incubated at 37 °C. Samples of 10 µl were taken, diluted with 990 µl of water and the reaction was stopped at 95 °C for 6 min. The boiled samples were further diluted 1:10 and analyzed via HPAEC-PAD. Yields were determined using standards of known concentration.

(-)-Epicatechin

400 µl reaction mixtures containing 50 mM or 150 mM (-)-epicatechin, 30 % DMSO, 600 mM sucrose, 100 mM MOPS buffer pH=7 and 1.14 g/l BaSP Q345F were incubated at 37 °C. Samples of 40 µl were taken, diluted with 360 µl of water and the reaction was stopped at

95 °C for 6 min. The boiled samples were further diluted 1:50 with methanol and analyzed via HPLC. Yields were calculated from the relative peak areas. Standards of known concentrations were used to determine the relative extinction coefficiants.

Resveratrol

400 µl reaction mixtures containing 75 mM (-)-resveratrol, 30 % DMSO, 600 mM sucrose, 100 mM MOPS buffer pH=7 and 1.14 g/l BaSP Q345F were incubated at 37 °C. Samples of 40 µl were taken, diluted with 360 µl of water and the reaction was stopped at 95 °C for 6 min. The boiled samples were mixed with 400 µl of 60% DMSO to dissolve precipitated resveratrol and consequently diluted 1:25 with Methanol and analyzed via HPLC. Yields were calculated from the relative peak areas.

2.6.7. Chromatographic analysis

HPLC

Conversions of epicatechin and resveratrol were determined on analytical scale using a JASCO HPLC system (pump PU-1580, gradient unit LG-980-02S, degasser DG-2080-53 and UV detector MD-2010-plus) at ambient temperature. The system was controlled by the Galaxie Chromatography Data System (Agilent). A Symmetry-C_{18} column (Waters; 5 µm, 4.6×250 mm) was used for the chromatographic separation as stationary phase. Mobile phases were (A) H_2O with 0.05% trifluoroacetic acid and (B) MeOH with 0.05% trifluoroacetic acid.

Epicatechin and its glycosylated derivatives were resolved using a linear binary gradient programmed as follows: 0 min 10% B, 20 min 50% B, 21 min 100% B, 23 min 100% B, 24 min 10% B, 28 min 10% B. The flow rate was set to 0.8 mL/min and the detection wavelength to 220 nm.

Resveratrol and its glycosylated derivatives were resolved using a linear binary gradient programmed as follows: 0 min 10% B, 20 min 100% B, 24 min 100% B, 25 min 10% B, 28 min 10% B. The flow rate was set to 0.8 mL/min and the detection wavelength to 320 nm.

HPAEC-PAD

HPAEC-PAD analysis was performed with a Dionex ICS-5000+ SP system utilizing a Carbopac PA10 column. Eluents were 100 mM NaOH (A), 100mM NaOH, 1 M NaOAc (B), and 250 mM NaOH (C). Catechin and its glycosylated derivatives were resolved using a multistep gradient programmed as follows: 0 min to 9 min 100% A, 10 min 100% B, 10 min to 31 min 100% B, 31 min to 45 min 100% C, 45 min to 60 min 100% A.

2.6.8. Crystallization, data collection, structure determination and –refinement

Crystals were grown using the hanging drop method. 4 g/l Protein solution was mixed with precipitant solution containing PEG 4000 (20-30%), mM NaCl (150 mM) and MES-buffer (pH= 6.35-7, 100 mM) and sucrose (2%). Crystals were grew for 10 weeks at 14 °C up to assize of 0.05x0.04x0.08 mm. Crystals were then transferred to mother liquor supplemented with 25% glycerol, mounted in cryo loops and plunged into liquid nitrogen. At beamline ID29 of the ESRF Grenoble the mounted crystals were placed within a 100K nitrogen gas stream and datasets were collected over 180° oscillation range. The datasets were autoindexed, integrated and scaled with XDS [109]. The structure was solved by molecular replacement using chain B of PDB entry 2GDV as a search model within PHASER[110]. After initial refinement within Phenix, regions with distinct conformational changes were manually rebuilt within COOT[111]. After three more rounds of automated refinement and manual rebuilding including water and ligand placement, the R and R_{free} factors converged.

2.6.9. Crystal structure data collection and refinement statistics

See Appendix **chapter 7.5**

2.6.10. NMR and Mass Spektra

See Appendix **chapter 7.2**

2.7. Acknowledgements:

We thank the team of beamline ID29 of ESRF Grenoble, France for their excellent support during data collection. We thank Jan Wendrich for supplying the HPLC analysis and Julian Görl for generating the images.

Chapter 3

SYNTHESIS OF THE RARE DISACCHARIDE NIGEROSE BY STRUCTURE-BASED DESIGN OF A PHOSPHORYLASE MUTANT WITH ALTERED REGIOSELECTIVITY

Michael Kraus, Julian Görl, Malte Timm and Jürgen Seibel

Published in Chemical Communications, volume 52, pages 4625-4627, 2016.

Summary

Where **chapter 2** focused on the glucosylation of aromatic compounds by BaSP Q345F **chapter 3** describes the behaviour of the variant towards carbohydrate acceptors. In the absence of suitable acceptor substrates BaSP and its variants hydrolyse sucrose and use the emerging glucose as an acceptor, resulting in the production of glucose-glucose disaccharides. BaSP Q345F was found to produce nigerose and maltose whereas the wild type yields maltose and kojibiose. The nigerose/maltose ratio can be optimized through the addition of DMSO. The isolation of nigerose was facilitated by removal of all other carbohydrates via the use of baker's yeast. Docking studies led to the hypothesis, that the domain shift enables the production of nigerose by removing a steric clash between C-6 of the acceptor glucose and Tyr196. It was assumed that glucose is bound in the enzyme conformation responsible for sucrose binding, and that the domain shift extends to this conformation. New structural evidence presented in **chapter 4** and **chapter 5** falsifies this theory.

Recently Verhaeghe et al. presented the L341I_Q345S variant which selectively produces kojibiose.[85] As these mutations target the flexible Loop A (See **chapter 1.3.3**) we chose to additionally study the disaccharide production by the variants L341, D316C_L341C and D316C_N340C, which were created by Julian Görl.

3.1. Abstract

In the absence of the natural acceptor inorganic phosphate wild-type sucrose phosphorylase from *Bifidobacterium adolescentis* (BaSP) produces maltose (4-*O*-α-D-glucopyranosyl-d-glucose) and kojibiose (2-*O*-α-D-glucopyranosyl-d-glucose) as sole transfer products. A Q345F exchange switches the enzyme's regioselectivity from 2 to 3 exclusively, yielding the rare sugar nigerose (3-*O*-α-D-glucopyranosyl-d-glucose, sakebiose).

3.2. Introduction

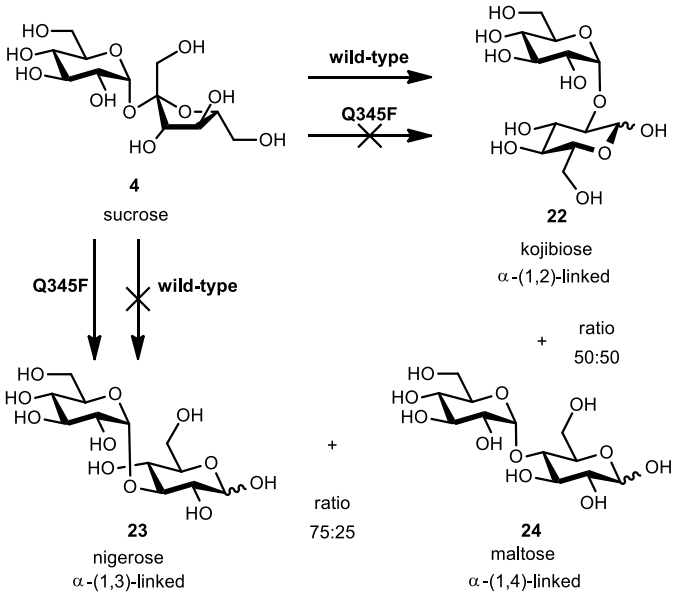


Figure 22 The Q345F variant causes a switch in product formation from kojibiose (22) to nigerose (23). Reproduced from Ref. 137 with permission from the Royal Society of Chemistry

Nigerose (**23**) (**Figure 22**) is the rare α-*O*-(1,3)-linked glucose disaccharide. In a study with mice, nigerose has shown immunopotentiating activity[112]. In nature, nigerose occurs as a dimeric building block in polysaccharides such as nigeran (glucose with alternating α-(1,3)- and α-(1,4)-glycosidic bonds) which is present in the cell walls of a few members of filamentous fungi such as *penicillium*[113] or *aspergillus*.[114] It is also found in Japanese sake (sakebiose) in small amounts.[112] While some methods for the production of nigerose have been reported[71, 72, 115-118] pure nigerose remains rarely available and expensive. They usually require costly starting materials or result in difficult to separate product mixtures.[71, 72, 115, 116, 118] Nihira *et al.* described a synthesis of nigerose from sucrose, but a multiple enzyme system was required.[117]

BaSP has already been subject to many detailed studies including elucidation of reaction mechanism, crystallization (of the wild-type and various variants)[44, 119, 120], thermostability

optimization[89], and immobilization experiments for potential industrial application[77]. The enzyme is known to transfer glucose to small molecules like glycerol (Glycoin®)[74, 121] and various acceptors utilizing α-D-glucose-1-phosphate or sucrose as donor via a double displacement mechnism.[40, 70, 74] It is also known that BaSP forms a mixture of the disaccharides kojibiose (**22**) and maltose (**24**) from sucrose as donor. In this regard Verhaeghe *et al.* optimized the product yield of the wild-type product kojibiose (**22**) from 50% to 95% yield by enzyme engineering of BaSP[122]. However, while α-(1,2)- and α-(1,4)-glucosylation of BaSP is known, so far no α-(1,3) nigerose formation of a BaSP variant is published.

Here, we report several variants of sucrose phosphorylase from *Bifidobacterium adolescentis* (DSM 20083) which switch the regioselectivity of the transfer reaction from α-(1,2) to α-(1,3), thus enabling the efficient synthesis and isolation of nigerose.

3.3. Results and discussion

3.3.1. Comparison of BaSP Variants

Table 6 Product distribution of several BaSP variants under optimized conditions after more than 90% consumption of sucrose (100 mM sucrose, 30% DMSO, 50 mM MOPS pH 7, 37 °C). a 100 mM glucose, b 400 mM glucose, c 100 mM glucose, 400 mM sucrose 55 °C, only 78% of sucrose is consumed due to enzyme degradation

BaSP variant	kojibiose [%]	nigerose [%]	maltose [%]	total transfer [%]
wild-type[a]	*28*	*n.d.*	*40*	*68*
L341C[b]	*39*	*2*	*41*	*91*
D316C L341C[b]	*37*	*2*	*28*	*80*
D316C N340C[b]	*23*	*10*	*28*	*61*
Q345F 37 °C[a]	*2*	*30*	*12*	*44*
Q345F 55 °C[c]	*2*	*31*	*13*	*46*

When wild-type BaSP is supplemented with 100 mM sucrose in the absence of inorganic phosphate, hydrolysis of sucrose and subsequent transfer to the released glucose is observed. The disaccharides kojibiose (**22**) and maltose (**24**) are the sole glucose disaccharides detected (HPAEC using Pulsed Amperometric Detection PAD) in the reaction mixture (50 mM MOPS buffer pH 7, 37 °C). We generated several mutations which were assumed to have an impact on the loop ^{341}LDLYQ^{345} and thus on the selectivity of acceptor binding.[44, 119, 120] Screening of these BaSP variants revealed four promising candidates for nigerose synthesis (**Table 6**).The best result was obtained with the Q345F variant, which was recently reported by Kraus *et al.* to transfer glucose to aromatic acceptors such as catechin and resveratrol.[119]

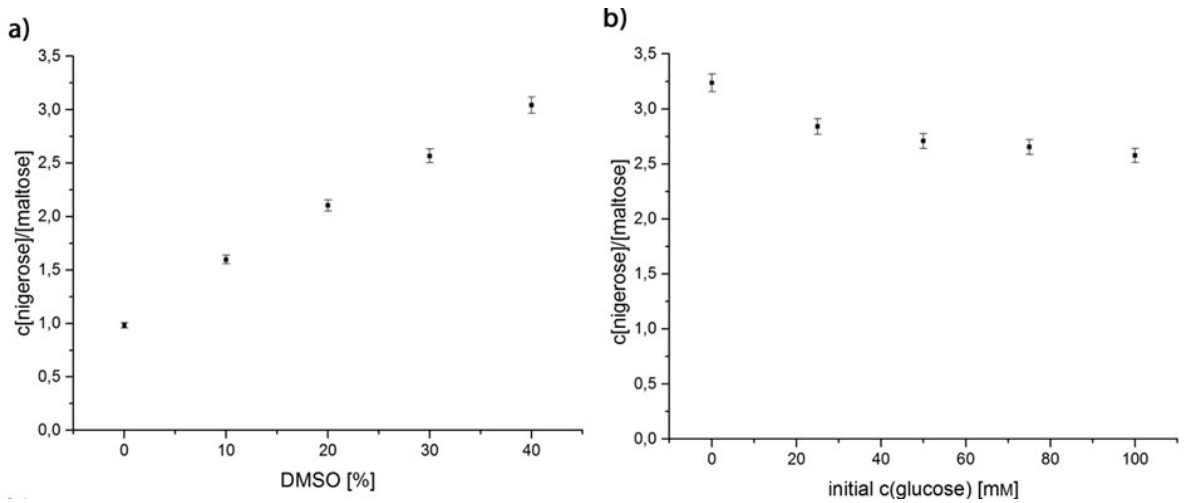


Figure 23 a) shift of nigerose/maltose ratio with increaseing DMSO concentration (100 mM sucrose, 50 mM glucose, 50 mM MOPS, 37 °C); b) shift of nigerose/maltose ratio with increasing glucose concentration (100 mM sucrose, 30% DMSO, 50 mM MOPS, 37 °C). Reproduced from Ref. 137 with permission from the Royal Society of Chemistry

3.3.2. Optimization of nigerose production

Including DMSO as a co-solvent in the reaction mixture improves BaSP wild-type activity[40] and suppresses hydrolysis resulting in higher yields of disaccharides. As shown in **Figure** 23**A**, raising the DMSO concentration from 0% to 40% at 100 mM sucrose and 50 mM glucose shifts of the nigerose/maltose ratio from 0.98 to 3.04 in favor of nigerose, equivalent with an increase of nigerose yield from 3.9 mM to 25.3 mM. Remarkably, a linear correlation between the DMSO concentration and the product ratio is observed. Other organic solvents (EtOH, *i*-PrOH, acetone, *t*-BuOH and *n*-BuOH) were tested at various concentration, but only minor variations of the nigerose/maltose ratio were observed. A unique synergistic effect between the Q345F mutation and DMSO appears to be present, since no other variant profits from DMSO in this way.

Varying the initial concentration of glucose from 0 mM to 100 mM at 100 mM sucrose and 30% DMSO increases the yield of nigerose from 13.5 mM to 30.0 mM (**Figure** 23**B**). Maltose production profits slightly more, and the nigerose to maltose ratio is reduced from 3.24 to 2.58. Higher glucose concentrations (>300 mM) do not increase the nigerose yield and nigerose concentration stabilizes around 30 mM. Production of nigerose was additionally tested at 55 °C with 30% DMSO. Since elevated levels of hydrolysis were observed, and higher concentrations of sucrose seem to stabilize the enzyme[123], reactions were performed with 400 mM sucrose and 100 mM glucose yielding 77.9 mM (26.7 g/l, 31%) nigerose (**Table 6**).

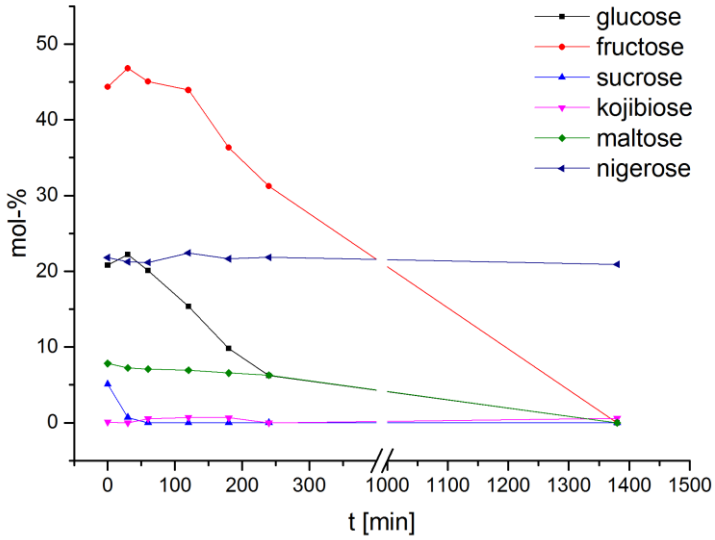


Figure 24 Consumption of sugars of the reaction mixture by baker's yeast (mol-% is based on the total sugar concentration at time 0). Reproduced from Ref. 137 with permission from the Royal Society of Chemistry

Nigerose purification from the reaction mixture can be simplified by the addition of baker's yeast immobilized on calcium alginate beads. The yeast was selected as it metabolizes the remaining sucrose, glucose, fructose and finally maltose in the given order (**Figure** 24) forming ethanol and glycerol, but does not degrade nigerose. Nigerose was obtained in good purity (>97.5% by HPAEC). The α-(1,3)-linkage of the product was determined by 2D-NMR and the obtained NMR-data is in agreement with the one previously published.[124]

3.3.3. Docking studies

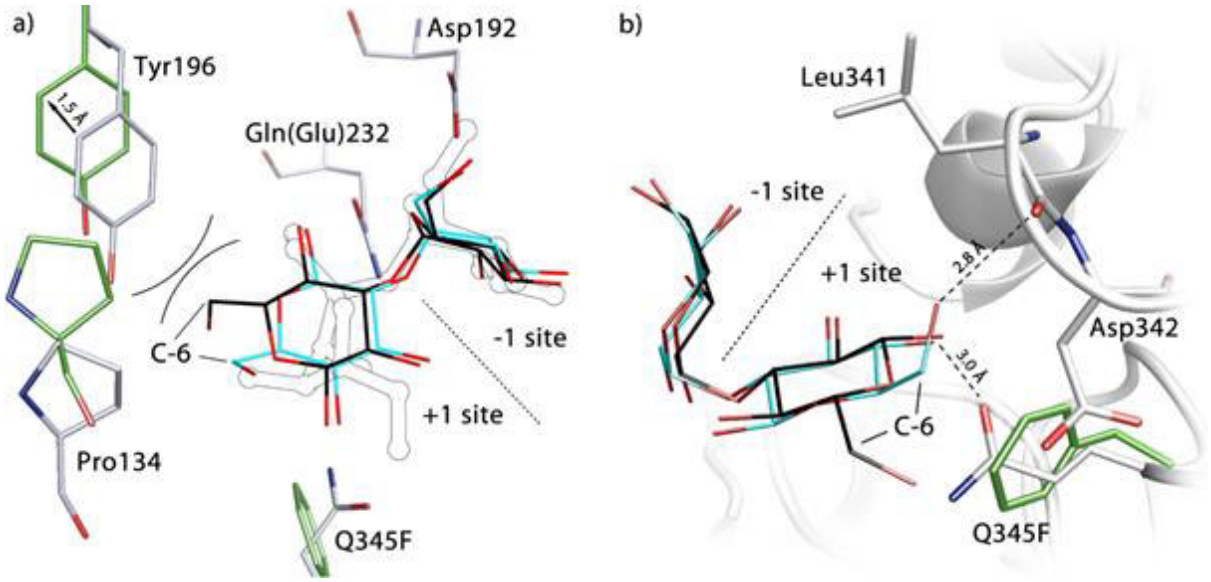


Figure 25 Comparison of kojibiose (cyan) docked into BaSP (2gdu, white) and nigerose (black) superimposed on kojibiose; green: X-ray structure of Q345F (5c8b); a) C-6 of nigerose is in proximity to Tyr196. The shift of Tyr196 in the Q345F (5c8b, green) variant creates additional space allowing the accommodation of nigerose in the active site; outlined: co-crystallized sucrose (2gdu). b) Close-up view of loop ^{341}LDLYF(Q)^{345} responsible for product specificity. Highlighted are potential H-bonds to 4-OH and 6-OH of docked kojibiose. Reproduced from Ref. 137 with permission from the Royal Society of Chemistry

In order to determine the reason of the selectivity switch, molecular docking studies (AutoDock Vina)[125] were performed with kojibiose or nigerose as ligands (supporting information). BaSP is known to convert sucrose into α-D-glucose-1-phosphate via a double displacement mechanism. The enzyme adopts two different conformations, one for sucrose binding (2gdu, 2gdv chain A) and a second for phosphate binding (2gdv chain B) with loops ^{341}LDLYQ^{345} and ^{134}PRP^{136} undergoing structural rearrangement.[44] For phosphate binding, Tyr344 and Arg135 are directed towards the +1 site, while for fructose binding Leu341 and Asp342 face the +1 site.

Very recently our group published the crystal structure (5c8b) of the Q345F variant.[119] Unfortunately, only structures with the proposed phosphate binding conformation of BaSP were obtained. Consequently the structure of the E232Q variant 2gdu[44] was chosen as a template and its co-crystallized sucrose was deleted for docking.

As known from literature BaSP prefers 1,2-diols as acceptors, and previous work suggests that productive binding of carbohydrate acceptors is achieved via Gln(Glu)232 and Gln345 in the 2gdu structure.[60] Conformations obtained from docking experiments were considered productive if they fulfilled two criteria: 1. the glucoside moiety in the +1 site needs an H-bond to Gln(Glu)232; 2. the glucose in the -1 site has to be in agreement with the co-crystallised sucrose from the 2gdu.

Wild-type BaSP synthesizes only kojibiose but not nigerose. We were able to confirm this by our docking studies as only kojibiose but not nigerose yielded productive docking modes (see **Figure** 26). Nigerose was superimposed with the best binding mode of kojibiose (**Figure** 25). The positioning of C-6 is the most significant difference between kojibiose and nigerose binding. With the wild-type enzyme C-6 of nigerose comes in steric conflict with

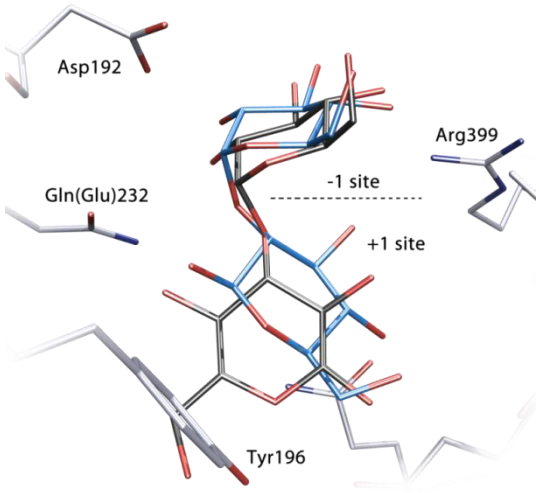


Figure 26 Best docking modes for nigerose (black) and kojibiose (blue): while the non-reducing glucose-moiety of kojibiose (-1 site) is in agreement with the sucrose (not shown for reasons of clarity) co-crystallized in the 2gdu structure, the best docking mode of nigerose is significantly shifted. Reproduced from Ref. 137 with permission from the Royal Society of Chemistry

Tyr196, explaining why the wild-type is not able to synthesize nigerose. So why can the Q345F variant produce nigerose? Recent work showed that the Q345F exchange causes several loop shifts, and as a result, Tyr196 is shifted by 1.5 Å. By that the steric clash is removed and nigerose synthesis is enabled (**Figure** 25**A**). C-6 of kojibiose points to Pro134, part of loop ^{134}PRP^{136}. In addition to the observed shifts, sequence and structure alignments (data not shown) of BaSP and LmSP (sucrose phosphorylase from *Leuconostoc mesenteroides*, which is known to produce nigerose)[71] reveal an RKD motif in this loop as the main difference in the active site.

According to the docking results, 6-OH of kojibiose can form H-bonds with the backbone of Leu341 (**Figure** 25**B**). Leu341 is part of a loop ^{341}LDLYQ^{345} which is made responsible for the product specificity.[122] The 5c8b structure reveals a rearrangement of Tyr344 due to steric hindrance. The findings above suggest long-range effects on loop ^{341}LDLYF^{345} in the fructose binding conformation of BaSP Q345F structure, favoring nigerose production.

3.4. Conclusion

In conclusion, through the Q345F exchange we achieved the switch in the regioselectivity from α-(1,2)- to α-(1,3)-glucosylation of glucose, introducing a new reaction into the repertoire of the well-established sucrose phosphorylase from *Bifidobacterium adolescentis*. We also described the synthesis of nigerose in a 31% yield. However a multi enzyme process with a yield of 61% was previously reported.[117] In addition the use of DMSO induces a preference for the formation of nigerose shifting from α-(1,4)- to α-(1,3)-glucosylation. Docking studies highlight the importance of the loops ^{134}PRP^{136} and ^{341}LDLYQ^{345} for product specificity in respect to kojibiose/nigerose formation. The shift of Tyr196 offers an explanation for the Q345F variant's ability to produce nigerose. Rearrangements in loop ^{341}LDLYQ^{345} induced by the Q345F exchange may also explain the suppression of kojibiose formation.

3.5. Experimental Section

3.5.1. Cloning, expression and purification of the enzymes

Cloning, expression and purification of wt-BaSP:

Freeze-dried cultures of *Bifidobacterium adolescentis* (DSM 20083) were obtained from DSMZ (Deutsche Sammlung von Mikroorgansimen und Zellkulturen GmbH), and grown under anaerobic conditions in DSMZ medium Nr.58 without resazurin. Cells were harvested and the genomic DNA isolated, using a GenJet Genomic DNA purification Kit (Thermo Fisher). The BaSP gene was amplified from genomic DNA using the primers 5'-ATAACCATGGCTATGAAAAACAAGGTGCAGCTCATCAC-3' and

5’-CAATCCGCCTGTCGTCGCCCTCGAGTAAT-3’. The amplicon was inserted into pET-28b(+) using the NcoI and XhoI restriction sites yielding plasmid pET-28b(+)-BaSP-wt.

E. coli Bl21 star™ cells were heat shock transformed with plasmid pET-28b(+)-BaSP-wt. Overnight cultures of the transformed host in LB-medium containing 50 mg/l kanamycin sulfate were grown and 1.8 ml were used to inoculate 250 ml of LB-Medium (50 mg/l kanamycin sulfate). The cultures were incubated at 37 °C and 180 rpm until they reached an OD_{600} of 0.6, at which point the temperature was adjusted to 19 °C and IPTG was added to a final concentration of 0.5 mM. The cells were grown for additional 18 h after which they were harvested by centrifugation (4000 g for 10 min). The sediment was resuspended in lysis buffer (60 mM phosphate, 250 mM NaCl, 11 mM imidazole, 5 mM β-mercaptoethanol, pH=8). Cells were lysed using a sonifier and centrifuged at 17000 g for 10 min at 4 °C. The lysate was loaded onto 0.5 ml Ni-NTA columns equilibrated with lysis buffer and incubated at 4 °C and slow rotation for a minimum of 2 hours. The column was washed with 2.5 ml of lysis buffer and the protein was eluted with 1.5 ml of elution buffer (60 mM phosphate, 250 mM NaCl, 230 mM imidazole, 5 mM β-mercaptoethanol, pH=8). The buffer was exchanged to 20 mM MOPS-buffer (pH=7) using 5 ml Hi-Trap columns from GE Healthcare.

Construction and purification of the variants

The mutations were constructed applying the Megaprimer method. 5’-CAATCCGCCTGTCGTCGCCCTCGAGTAAT-3’ was used as reverse primer for the creation of the megaprimer and 5’-ATAACCATGGCTATGAAAAACAAGGTGCAGCTCATCAC-3’ was used as the forward primer in the second PCR. For the construction of the D316C_N340C and D312C_L341C, the variants N340C and L341C, respectively, were created first. The mutations D316C and D312C, respectively, were introduced in a second megaprimer PCR.

Table 7 Mutagenesis Primers

variant	forward mutagenic primer in first PCR
Q345Y	*5’-CCAATCTCGACCTCTACTTCGTCAACAGCACCTAC-3’*
N340C	*5’-CCGCCGCATCCTGTCTCGACCTC-3’*
L341C	*5’-CCGCATCCAATTGCGACCTCTACC-3’*
D316C	*5’-GGATGAGGACGTGTGCAACCTCGTCAAC-3’*
D312C	*5‘-GGTCTCGTGCCGTGTGAGGACGTGG-3’*

Cloning and purification of the variants followed the procedure described for the wildtype. The desired mutations were confirmed by DNA sequencing (GATC biotech AG, Konstanz, Germany).

3.5.2. Enzyme assays (final concentrations are given).

If not stated otherwise enzyme assays were carried out in MOPS buffer ((3-(*N*-morpholino) propane sulfonic acid, 50 mM, pH 7) supplemented with sucrose (100 mM) in a total volume of 100 or 200 µl. Reactions were incubated at 37 °C. The variants D316C_N340C and D312C_L341C were pretreated with 15 mM TCEP (tris(2-carboxyethyl)phosphine) for 1 h at 30 °C to ensure reduction of disulfide bonds.

Unit definition

1 U was defined as the enzyme activity that hydrolases 1 µmol of sucrose in one minute (200 mM sucrose, 50 mM MOPS buffer pH 7, total reaction volume 200 µL, 37 °C).

3.5.3. Immobilization of baker's yeast

1.25 g sodium alginate (purchased from VWR) was dissolved at 60 °C in 80 ml water. 10.0 g baker's yeast (Dr. Oetker, "frische Backhefe") was suspended in the solution and the resulting slurry was loaded into a syringe and added dropwise (6.0 ml min^{-1}, via syringe pump) to 500 ml of a stirred $CaCl_2$-solution (150 mM).[126]

3.5.4. Production of nigerose

In a total volume of 10 ml, sucrose (400 mM) was supplemented with glucose (200 mM) in MOPS buffer (20 mM, pH 7) and 30% DMSO. 1.0 mL of BaSP Q345F (5 mg/mL, activity: 62 U/g) was added and the reaction was incubated at 37 °C under slow agitation. After 4 d (90% sucrose consumption) the reaction was stopped by the addition of 20 ml MeOH. The occurring precipitate was removed by centrifugation (10 min, 6000 g). After evaporation of the solvent and freeze-drying, the residual syrup was supplemented with 50 ml water and baker's yeast (immobilized on calcium alginate beads, 20 beads). The consumption of sugars at 20 °C was monitored by HPAEC (**Figure** 24). After reaction completion, the baker's yeast was removed by filtration and the solvent was removed by freeze-drying. Silica gel chromatography (0.063-0.200 mm, MeCN/MeOH = 4:1) yielded pure nigerose (430 mg, 24%).

3.5.5. HPAEC-Analysis

Sample preparation

Collected samples were diluted (1:50) with water and boiled at 95 °C for 5 min to stop the catalytic reaction. After centrifugation at 17.000 g for 5 min the samples were furthermore diluted (final total sugar concentration: 100-200 µM).

HPAEC-PAD analysis

High Performance Anion Exchange Chromatography with Pulsed Amperometric Detection (HPAEC-PAD) was performed with a Dionex ICS-5000+ SP system utilizing a Carbopac PA10 (2x250 mm) column. The saccharides were resolved using an isocratic program (100 mM NaOH and 30 mM NaOAc at 250 µL/min). Yields were determined by peak area using external standards (fructose, glucose, sucrose, kojibiose, nigerose and maltose) of known concentrations (10 µM, 25 µM, 50 µM, 75 µM and 100 µM).

3.5.6. Determination of Glucose and solvent influence on Nigerose Production

Glucose

In a total volume of 100 µl, sucrose (100 mM) was supplemented with glucose (0-100 mM, in 25 mM intervals), in MOPS-buffer (50 mM, pH 7) and 30% DMSO. 95 U/l BaSP Q345F was added and the reaction mixture was incubated at 37 °C. After 52 h, samples of 10 µl were taken and yields were determined using HPAEC-PAD.

DMSO

In a total volume of 100 µl, sucrose (100 mM) was supplemented with glucose (50 mM) in MOPS-buffer (50 mM, pH 7) and DMSO (0-40% in 10% intervals). 95 U/l BaSP Q345F was added and the reaction mixture was incubated at 37 °C. After 52 h, samples of 10 µl were taken and yields were determined using HPAEC-PAD.

Other solvents.

In a total volume of 100 µl, sucrose (100 mM) was supplemented with glucose (50 mM) in MOPS-buffer (50 mM, pH 7). Different solvents (EtOH, *i*-PrOH, acetone, *n*-BuOH, *t*-BuOH, 0-30% in 10% intervals) were added and the reactions were started by the addition of 95 U/l BaSP Q345F. Samples were taken after incubation at 37 °C for 17, 48 and 72 h and yields were determined using HPAEC-PAD.

3.5.7. Molecular Docking

Protein and ligand setup

The crystal structure of the BaSP E232Q mutant (pdb code 2gdu, chain A) was retrieved from PDB as “receptor” for the docking calculations. All water molecules and ligand entries were removed, non-polar hydrogens were added using AutoDockTools 1.5.6r.[127] Grid box center and grid dimensions (16x16x16 Å, grid spacing: 1.0 Å) were determined via AutoDockTools and transferred to the AutoDock Vina configuration file.

Ligands were built and geometries were optimized using ChemBioOffice 13 (MM2 energy minimization with default settings). Both α- and ß-anomers of kojibiose and nigerose were

prepared. Gasteiger charges were added and rotatable bonds were assigned using AutoDockTools.

AutoDock Vina

AutoDock Vina[125] was used for docking calculations. The docking parameters "exhaustiveness" and "energy_range" were set to "25" and "4", respectively. Sucrose was used a test ligand for the docking procedure resulting in an excellent agreement with the binding mode of the co-crystallized sucrose of the 2gdu structure. Conformations in which the non-reducing glucose-moiety was in agreement with the one of the 2gdu structure were considered productive binding modes and within these the best-scored modes were chosen for closer analysis. From the analyzed binding modes no significant difference was observed between the α- and ß-anomers of the ligands.

3.5.8. NMR Spektra

See Appendix **chapter 7.2**

3.5.9. Structure and sequence alignments.

Structure alignments of BaSP (2gdu and 2gdv, chain A and B) with LmSP were performed using I-TASSER[128] (data not shown). The section of the sequence alignment (ClustalW2) neighboring the ^{134}PRP^{136} motif is given below.

BaSP 123 ATEEDLAGIYRPRPGLPFTHYKFAG 147

LmSP 126 PTQADVDLIYKRKDKAPTQEITFDD 150

Chapter 4

SWITCHING ENZYME SPECIFICITY FROM PHOSPHATE TO RESVERATROL GLUCOSYLATION

Michael Kraus, Clemens Grimm and Jürgen Seibel

Published in Chemical Communications, volume 53, pages 4625-4627, 2016.

Summary

This chapter follows up on the unsolved question from **chapter 2** whether the domain shift is in fact responsible for the novel activity towards the aromatic substrates. A crystal structure of an inactive BaSP Q345F in complex with resveratrol-3-α-D-glucoside features the domain shift offering proof that the glucosylation of the aromatic substrates is a product of the domain shift. The selectivity of the variant for aromatic substrates is further tested via kinetic studies and BaSP Q345F is revealed to have a high affinity for these substrates. The initial strategy, the installation of π-π-stacking is also shown to be successful, albeit different from the envisioned geometry, as T-shaped π-π-interactions are found.

Chapter 3 is also revisited, as the crystal structure of BaSP Q345F in complex with nigerose points out that the same conformation that enables the glycosylation of aromatic compounds is responsible for nigerose production and not a domain shifted *F*-conformation. The issue of the *F*-conformation of BaSP Q345F will be addressed in **chapter 5**.

4.1. Abstract

Here we present a point mutation-triggered domain shift which switches the acceptor preference of a sucrose phosphorylase from phosphate to a variety of large polyphenolic compounds including resveratrol and quercetin, enabling their efficient glucosylation. The variant possesses a high affinity for aromatic substrates due to newly introduced π-π- and hydrophobic interactions in the altered active site. The domain shift brings about a substantially enlarged and multifunctional active site for polyphenol glucosylation and rare disaccharide production. The crystal structure of the variant with its product resveratrol-3-α-D-glucoside allows the prediction of the substrate scope and regioselectivity of the aromatic compounds' glucosylation sites

4.2. Introduction

Polyphenols, in particular the extensively studied resveratrol and quercetin, exhibit antitumor activities[12, 13] and play a key role in lifespan and health span extension[14-16]. Glycosylation is desired in order to increase bioavailability, fine-tune bioactivities and pharmaceutical properties and to improve delivery of polyphenol drugs to target cells.[99, 129] Thus engineering of carbohydrate processing enzymes towards accepting polyphenols as substrates is the subject of several recent investigations.[91, 99, 100, 119, 130-132] We chose sucrose phosphorylase (SP, EC 2.4.1.7, GH13) as the target for protein engineering because it is an industrially important enzyme that utilizes the abundant glucosyl donor substrate sucrose to transfer glucose moieties to various acceptors.[40, 45, 73] However, SPs do not glucosylate resveratrol or

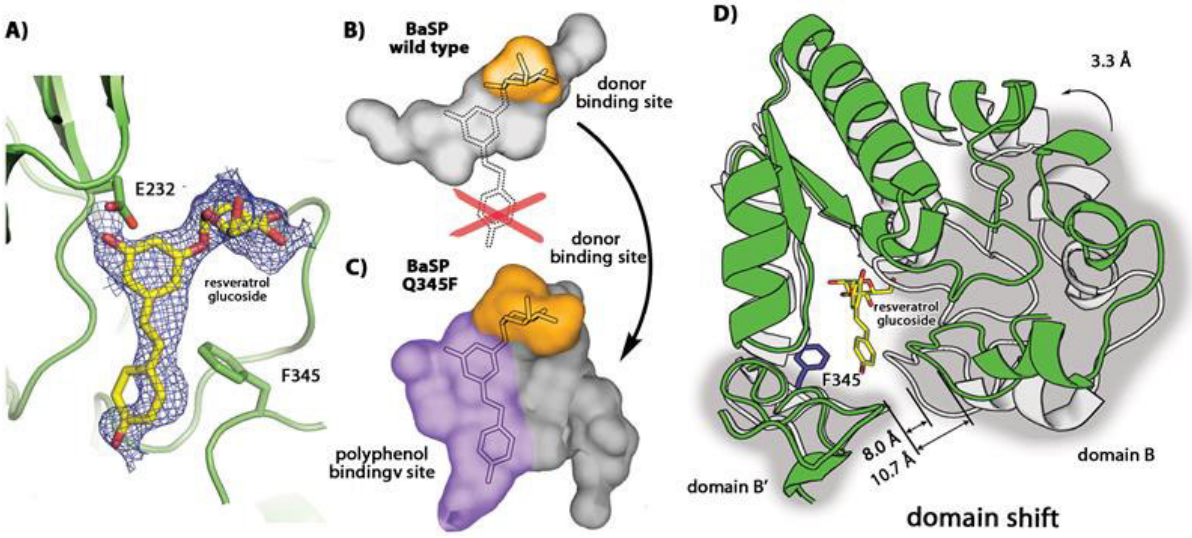


Figure 27 Enabling of aromatic compounds glucosylation via domain shift in BaSP Q345F; A) Electron density of bound resveratrol-3-α-D-glucoside 1.0 σ; B) Active site dimensions and substrate positioning in the wild-type enzyme, dotted line: required position of resveratrol; C) Domain shift responsible for the creation of a novel active site, green: BaSP Q345F, grey: BaSP wild-type (PDB ID 2gdv chain B); D) Engineered active site of BaSP Q345F with bound resveratrol-3-α-D-glucoside. Reproduced from Ref. 140 with permission from the Royal Society of Chemistry

other polyphenols efficiently due to the spatial limitations of the active site (**Figure 27**).[40, 91, 119] The overall poor solubility of this compounds in aqueous system presents a further challenge and call for a high affinity of the engineered enzymes towards their target substrates.[91] In general, the introduction of new activities into enzymes has been achieved by state of the art protein engineering including directed evolution[133] and rational *de novo* approaches where a theoretical active site is constructed *in silico* and accommodated in an existing protein scaffold.[134] To date, optimization of *de novo* enzymes via directed evolution is required to achieve activities comparable to naturally occurring proteins.[135] Therefore structure based, (semi-) rational exchanges remain a common tool and are performed mostly to enlarge the active site, or to fine-tune its polarity and ligand protein interactions.[136] We followed an unconventional strategy during the redesign of *Bifidobacterium adolescentis* sucrose phosphorylase (BaSP)[119] and exchanged a glutamine residue (Gln345), located at the acceptor binding site against phenylalanine, which actually introduces a larger sidechain into the active site. By introducing a non-polar aromatic residue into the acceptor binding site, we envisioned to enable π-π stacking with the desired aromatic substrates. Furthermore the exchanged glutamine is involved in phosphate binding and may play a role in utilizing glucose as an acceptor. Reducing this undesired side reactions by eliminating polar interactions to this substrate was an additional goal of the design strategy.

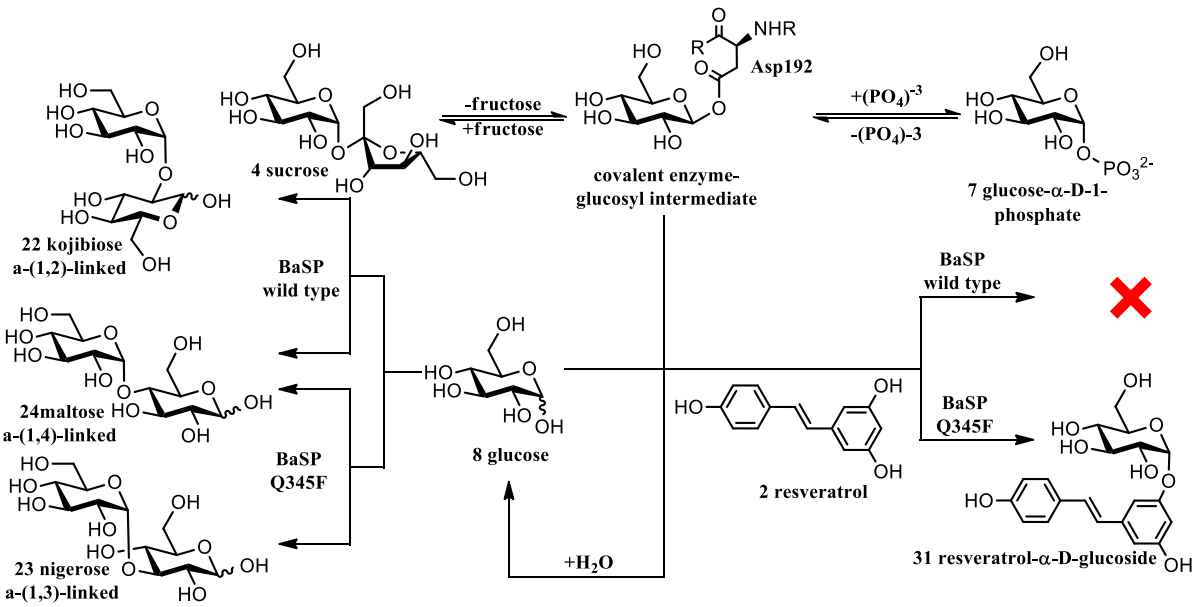


Figure 28 Overview of reactions catalysed by BaSP wild-type and BaSP Q345F. The interconversion of Sucrose and glucose-1-Phosphate via a covalent enzyme glucosyl intermediate comprises the native reaction of all sucrose phosphorylase. The covalent intermediate can be intercepted by a variety of acceptors, including water, which leads to hydrolysis and the subsequent formation of glucose-glucose disaccharides. The Q345F variant prefers the transfer of glucose to polyphenolic substrates like resveratrol. Reproduced from Ref. 140 with permission from the Royal Society of Chemistry

4.3. Results and discussion

4.3.1. Analysis of the crystal structure

The Q345F exchange in fact enables the efficient glucosylation of bulky polyphenol species (**Figure** 28**).** This effect is explained by the crystal structure of the engineered BaSP in complex with glucosylated resveratrol **(Figure** 27**A**). The active site volume increases from 272 to 557 Å^3 providing the space required for resveratrol and flavonoid accommodation and glucosylation (**Figure** 27**B and C**) due to an initially unexpected shift of one domain (**Figure** 27**D**). As the observed domain shift might be of further interest to manipulate the active site of TIM-barrel enzymes in particular the vast GH13 family, we moved on to a closer inspection of its underlying mechanics. The central domain of SP and the GH13 family with its 37,000 members is a TIM-barrel which harbours the enzyme's active site near the C-terminal ends of its eight parallel β-strands. The loops at the C-terminal end of the strands are often replaced by versatile domains and thus structurally and functionally diverse active sites are created. In sucrose phosphorylases the acceptor binding site is defined by two of these domains, domain B (residues 86-166 **Figure** 27**D**) and domain B' (residues 292-355).[44] In the BaSP variant Q345F the domain shift responsible for the active site remodelling consists of domain B moving 3.3 Å away from B' (**Figure** 27**D**). This rearrangement is independent of the presence of a ligand in the active site. Therefore an induced fit effect as the reason behind the domain shift can be excluded. Ultimately, the Q345F exchange causes an opening of BaSP and enables resveratrol and flavonoid accommodation via a *provoked fit* effect. A potentially adverse effect connected to the observed domain shift could be the displacement of the highly conserved residues His88 and Gln160 involved in substrate binding (**Figure** 29). These residues coordinate the 4-OHand 6-OH of the donor glucose moiety. An additional defined water molecule is found in the BaSP Q345F crystal structure coordinated by residues His88 and Gln160. The water molecule bridges the increased

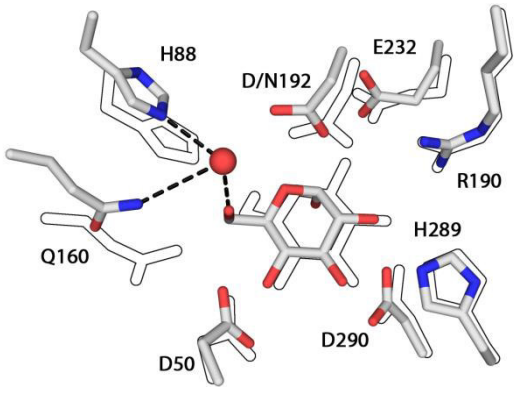


Figure 29 Coordination of the glucose moiety of resveratrol-3-α-D-glucoside in the -1 site of BaSP D192N/Q345F: Outlines indicate corresponding positions in the structure of the wild-type enzyme (PDB ID code 2GDV chain B). The red sphere inticates the position of the water molecule. Reproduced from Ref. 140 with permission from the Royal Society of Chemistry

distance and recovers the lost hydrogen bond to the 6-OH of the glucose moiety, thus healing the distortions in the active site (**Figure** 29).

4.3.2. Kinetic investigation of BaSP Q345F

The engineered variant BaSP Q345F glucosylates various polyphenolic acceptors efficiently, achieving a yield of 97% in the synthesis of resveratrol-3-α-D-glucoside from resveratrol **(Figure** 30**)** More importantly, our engineered variant displays a high affinity for the desired aromatic acceptors (K_M 0.08-1.55 mM), while the affinity for phosphate is simultaneously reduced (K_M from 4.8 to 26 mM, **Table** 8)

K_M values for the wild-type enzyme with polyphenolic substrates could not be determined due to its virtually non-existing affinity towards this compounds. The variants improved transfer to polyphenolic substrates is due the significant stabilizing interactions between the enzyme and the novel acceptor substrates and to the creation of space in the active site.

Substrate	**K_M [mM]**	**k_{cat} [s^{-1}]**	***K_{cat}/K_M [$M^{-1}s^{-1}$]***
Phosphate	*25.7 ±1.86*	*0.179 ±0.007*	*6.96*
Resveratrol	*0.92±0.09*	*0.131 ±0.002*	*142*
Quercetin	*0.52±0.06*	*0.094 ±0.003*	*181*
Fisetin	*0.32±0.12*	*0.067 ±0.004*	*209*
(-)-Epicatechin	*1.55±0.43*	*0.104 ±0.007*	*67.1*
(+)-Catechin	*0.95±0.41*	*0.074 ±0.010*	*77.9*
Naringenin	*0.08±0.01*	*0.002 ±0.0001*	*250*

Table 8 Kinetic data of BaSP Q345F acceptor substrates Values based on Michaelis-Menten fittings. A detailed analysis including Lineweaver-Burk plot, Hanes-Woolf plot and direct linear plot[139] included in Table 9.

4.3.3. Investigation of BaSP Q345F product profiles

Investigation of both the product profiles and the crystal structure of BaSP Q345F in complex with resveratrol grant further insight into the variants binding mechanism. BaSP Q345F prefers 1,2- and 1,3-aromatic diols and utilizes the 3-OH of resveratrol and the 3'-OH of flavonoids, while ignoring the 4'-OH of both acceptors (**Figure** 30)

This fact is explained by the chemical environment of the resveratrol moiety of resveratrol-3-α-D-glucoside in the engineered acceptor binding site of the enzyme. The second non-glucosylated OH group of resveratrol is coordinated by Glu232 and the peptide nitrogen of Ala193 (**Figure** 32), and a comparable arrangement is possible for the 1,2- and 1,3- diol motifs present in flavonoids (**Figure** 31). If a carbonyl functionality is present in 4-position of a flavonoid the variant glucosylates the acceptor at position 7, whereas the catechins which lack the carbonyl functionality are glucosylated at the 5-OH moiety (**Figure** 30)

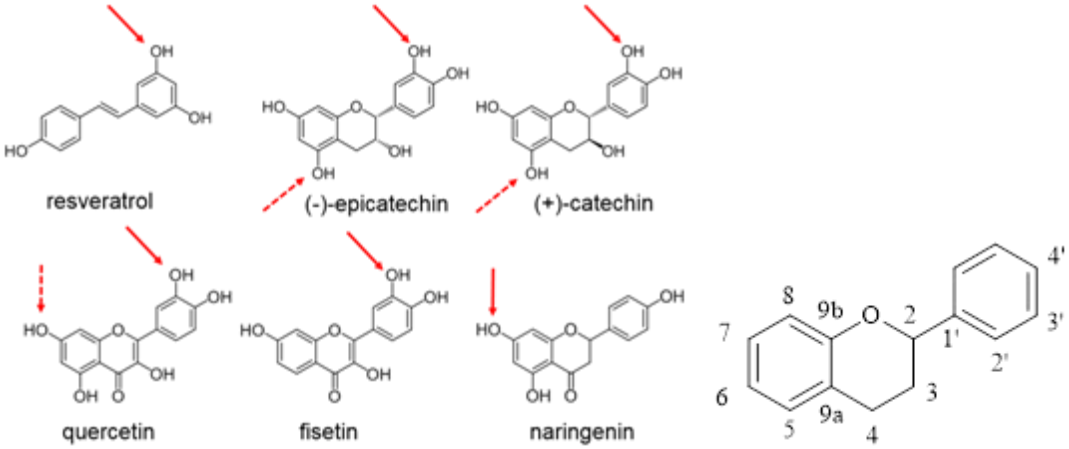


Figure 30 Glucosylation pattern of BaSP Q345F. Red arrow main glucosylation site, dashed arrow, secondary glucosylation site and Numbering of positions in flavonoids Conditions: 600 mM Sucrose, 30% (v/v) DMSO, 50 °C, acceptor concentrations: (-)-epicatechin 32, (+)-catechin 17: 100 mM, quercetin 3, fisetin 33: 50 mM, naringenin 34: 25 mM, resveratrol 2 75 mM. The numbering of positions in flavonoid systems is described at the right side. Reproduced from Ref. 140 with permission from the Royal Society of Chemistry

Glucosylation at the 4'-position of flavonoids would however require an almost linear orientation of the glucose moiety and all three flavonoid rings, which would cause a steric clash with the residues outlining the active site, in particular Tyr132, Tyr196 and Phe205. We therefore conclude that this region constitutes a newly identified hotspot for the modification of SP regioselectivity towards complex aromatic acceptors.

Detailed analysis of product distributions revealed that (-)-epicatechin is glucosylated in equal measures in positions 3' and 5 while (+)-catechin, is almost exclusively glucosylated in

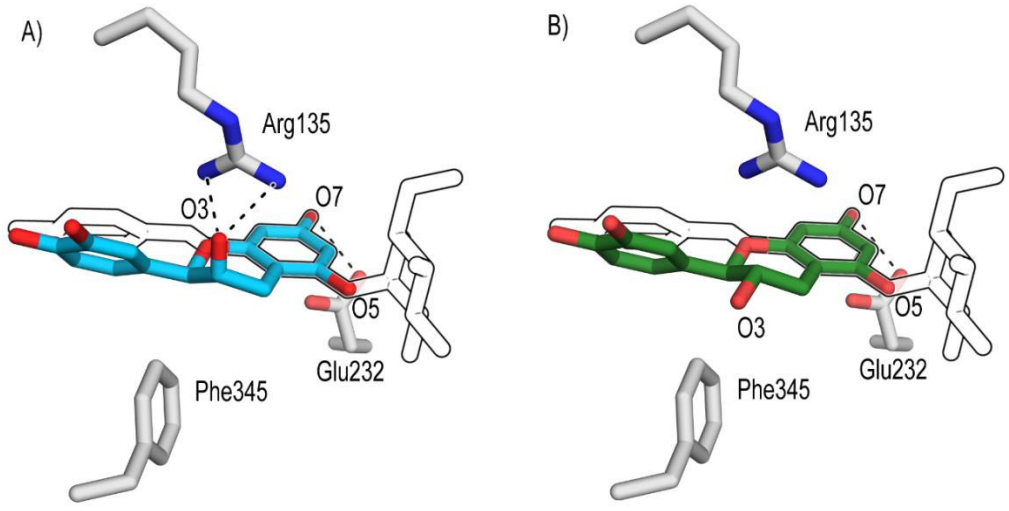


Figure 31 Investigation of the O5-glucosylatoion of (-)-epicatechin and (+)-catechin by BaSP Q345F (-)-Epicatechin (A, light blue) and (+)-catechin (B, green) were superpositioned with the resveratrol moiety complexed by the D192N/Q345F-variant, which comprises the most likely productive binding mode, leading to (-)-epicatechin-5-glucosid and (+)-catechin-5-glucosid. The A ring of the flavanol is able to perfectly mimic resveratrol. The key difference is found at the 3-OH moiety which is oriented towardsArg[135] in epicatechin in a distance and orientation suitable for a hydrogen bond. In the case of (+)-catechin the 3-OH group faces the nonpolar sidechain of F345. This observations are in accordance with the fact that glucosylation in 5-position is a slow side reaction for (+)-catechin, while glucosylation (-)-epicatechin is equally distributed between 3'-OH and 5'OH. Reproduced from Ref. 140 with permission from the Royal Society of Chemistry

3'-position. Superimposition of both flavanols with the resveratrol moiety bound to the active site (**Figure** 31) point to a hydrogen bond between the 3-OH moiety of (-)-epicatechin and Arg135 as the cause of this variance in regioselectivity.

4.3.4. Aromat binding by BaSP Q345F

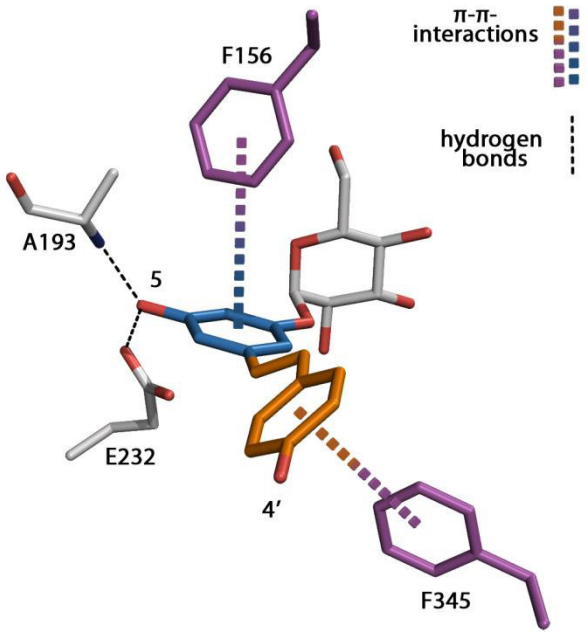


Figure 32 π-π-interactions: The A-ring of the resveratrol moiety (blue) is stabilized via hydrogen bonds to Ala193 and Glu232 and displays T-shape π-π-interactions to Phe156 (74° angle between the aromatic rings), the B-ring and the conjugated double bond (orange) undergo π-π-interactions with Phe345 (88°). Reproduced from Ref. 140 with permission from the Royal Society of Chemistry

Our initial strategy to introduce π-π-interactions between the polyphenols and the enzyme to force the acceptor substrate in a productive position for glucosylation is reflected in **Figure** 32. The binding of nonpolar substrates is further facilitated through the increased hydrophobic and aromatic character of the engineered active site. The interaction interface between the resveratrol moiety and the enzyme has an area of 308 Å^2. In addition, T-type π-π stacking interactions between Phe156 and the A-ring of resveratrol stabilize the substrate (**Figure** 32). A second T-shape π-π interaction exists between the π-system formed by the central double bond and the B-ring of resveratrol and the engineered residue Phe345 (**Figure** 32).

4.3.5. Disaccharide production of BaSP Q345F

While the domain shift-created novel active site of BaSP Q345F is particularly well suited for large polyphenols, it displays a certain degree of substrate promiscuity. In the absence of suitable acceptors BaSP slowly hydrolyses sucrose and utilizes the resulting glucose as an acceptor to form maltose (4-O-α-D-glucopyranosyl-D-glucose) and nigerose (3-O-α-D-

glucopyranosyl-D-glucose) in case of the Q345F variant.[137] In contrast the wild-type enzyme synthesizes maltose and kojibiose (2-O-α-D-glucopyranosyl-D-glucose).[85] A crystal structure (**Figure** 33) of the variant in complex with the non-natural product nigerose illustrated that nigerose production is enabled by the same, engineered acceptor binding site that allows polyphenol glucosylation.

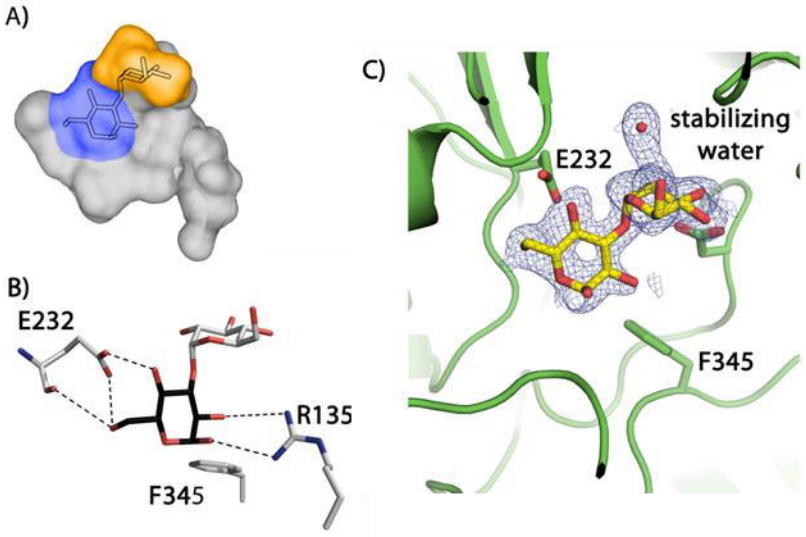


Figure 33 Nigerose coordination in BaSP Q345F A) active site dimensions and acceptor glucose positioning; B) Hydrogen bonds between the acceptor glucose of nigerose and BaSP D192N/Q345F; C) Electron density of bound nigerose 1.0 Reproduced from Ref. 140 with permission from the Royal Society of Chemistry

4.4. Conclusions

In summary, we present the creation of a new multifunctional acceptor binding site *via* a domain shift and the introduction of favourable π-π-interactions. To the best of our knowledge this is the first example of an active site remodelling by a domain shift which is visualized by protein structures with and without the substrates. The domain shift is triggered by a single amino acid exchange and is responsible for remodelling the acceptor-binding site of BaSP into a polyphenol binding site. The engineered variant is capable of the glucosylation of a wide variety of bulky flavonoids, including quercetin and resveratrol as well as of the synthesis of rare disaccharides. The crystal structures of our engineered sucrose phosphorylase, in complex with its respective products resveratrol-3-α-D-glucoside and nigerose explain the mode of substrate binding and may reveal hotspots for future modification of this and possibly other SP variants.

4.5. Acknowledgements

We thank the team of beamline ID30B of ESRF Grenoble, France for their excellent support during data collection. The corresponding coordinates and structure factors are available from the PDB under accession code 5M9X (BaSP D192N/Q345F in complex with

resveratrol-3-α-d-glucoside), 5MAN (BaSP D192N/Q345F in complex with nigerose) and 5MB2 (BaSP Q345F apo form).

4.6. Experimental Section

4.6.1. Materials and Methods

Materials

Resveratrol-3-α-D-glucoside was produced and purified following the previously reported method.[119] Nigerose was produced and purified following the previously reported method.[137] All other chemicals were purchased from Sigma-Aldrich or VWR. Ni-NTA resin was a product of Sigma Aldrich, PCR primers were ordered from Sigma-Aldrich. Mutagenesis was performed applying the QuikChange II Site-Directed Mutagenesis Kit from Agilent.

4.6.2. Production of nigerose (as previously described)[137]

In a total volume of 10 ml, sucrose (400 mM) was supplemented with glucose (200 mM) in MOPS-NaOH buffer (20 mM, pH 7) and 30% (v/v) DMSO. 1.0 mL of BaSP Q345F (5 mg/mL, activity: 62 U/g) was added and the reaction was incubated at 37 °C under slow agitation. After 4 d (90% sucrose consumption) the reaction was stopped by the addition of 20 mL MeOH. The occurring precipitate was removed by centrifugation (10 min, 6000 g). After evaporation of the solvent and freeze-drying, the residual syrup was supplemented with 50 mL water and baker's yeast (immobilized on calcium alginate beads, 20 beads). The consumption of sugars at 20 °C was monitored by HPAEC After reaction completion, the baker's yeast was removed by filtration and the solvent was removed by freeze-drying. Silica gel chromatography (0.063-0.200 mm, MeCN/MeOH = 4:1) yielded pure nigerose (430 mg, 24%).

4.6.3. Production, isolation and characterization of glucosylated polyphenols

Ca. 1.66 – 4.38 mmol polyphenol were dissolved in 20 mL 30% (v/v) DMSO containing 100 mM MOPS-NaOH-buffer pH = 7 and 1.1 M sucrose and 40 mg BaSP Q345F. After 64 h at 50 °C the reaction was stopped by incubating the reaction mixture at 95 °C for 15 min. The solvent was removed under reduced pressure and the crude product was purified by column chromatography (silica, ethyl-acetate : methanol 12:1 -> water : isopropanol : ethyl-acetate 1:3:6). These reactions were performed to obtain sufficient amounts of glucosylated product for characterization, and not optimized for maximum conversion.

Yields:

233 mg naringenin-7-α-D-glucoside (27%)

323 mg fisetin-3'-α-D-glucoside (42%)

400 mg glucosylated quercetin (mixture of quercetin-3'-α-D-glucoside, quercetin-7'-α-D-glucoside and quercetin-3',7-α-D-diglucoside, 48%)

936 mg catechin-3'-α-D-glucoside (53%) 236 mg catechin-3',5-α-D-diglucoside (10%)

103 mg epicatechin-3'-α-D-glucoside (12%) 274 mg mixture of epicatechin-3'-α-D-glucoside and epicatechin-5'-α-D-glucoside (31%), epicatechin-3',5-α-D-diglucoside (20%)

1.19 g resveratrol-3-α-D-glucoside (70%)

4.6.4. Acetylation of Quercetin-3',7-α-D-diglucoside

12.5 mg (20 µmol) of Quercetin-3',7-α-D-diglucoside and 50.1 mg (0.490 mmol) acetic anhydride were dissolved in 1.0 mL pyridine and stirred overnight at room temperature. Then 10 mL of water were added, and after 30 minutes of stirring the reaction mixture was extracted 3x with 10 mL ethylacetate. The combined organic layers were washed (3x 10 mL 1M HCl, 2x 10 mL saturated $CuSO_4$-solution, 1x 10 mL brine) and the solvent removed under reduced pressure to yield 22.1 mg (17.4 µmol, 87%) Quercetin-3',7-α-D-diglucoside-undecaacetate.

4.6.5. Cloning expression and purification of BaSP variants

Cloning of BaSP wild type

Freeze-dried cultures of *B. adolescentis* (DSMZ 20083) were obtained from DSMZ (Deutsche Sammlung von Mikroorgansimen und Zellkulturen GmbH), and grown under anaerobic conditions in DSMZ medium Nr.58 without resazurin. Cells were harvested and the genomic DNA isolated, using a GenJet Genomic DNA purification Kit (Thermo Fisher). The BaSP gene was amplified from genomic DNA using the primers

5'-ATAACCATGGCTATGAAAAACAAGGTGCAGCTCATCAC-3' and 5'-CAATCCGCCTGTCGTCGCCCTCGAGTAAT-3'. The amplicon was inserted into pET-28b(+) using the NcoI and XhoI restriction sites yielding plasmid pET-28b(+)-BaSP-wt.

Construction of BaSP Q345F

The Q345F mutation was constructed applying the Megaprimer method. The mutagenic primers 5'-CCAATCTCGACCTCTACTTCGTCAACAGCACCTAC-3', and 5'-CAATCCGCCTGTCGTCGCCCTCGAGTAAT-3' were used for the creation of the megaprimer and 5'-ATAACCATGGCTATGAAAAACAAGGTGCAGCTCATCAC-3' was used for the second PCR. Cloning and purification of the variant followed the procedure described for the wildtype.

Construction of BaSP D192N/Q345F

The D192N exchange was achieved with the QuikChange II Site-Directed Mutagenesis Kit from Agilent following the manufacturer's protocol. The mutagenic primers were 5'-CTACATCCGCCTCAACGCCGTCGGC-3' and 5'-CCGACGGCGTTGAGGCGGATGTAG-3'.

Expression and Purification BaSP variants

E. coli Bl21 star™ cells were heat shock transformed with plasmid pET-28b(+)-BaSP-wt. Overnight cultures of the transformed host in LB-medium containing 50 mg/L kanamycin sulfate were grown and 1.8 mL were used to inoculate 250 mL of LB-Medium (50 mg/L kanamycin sulfate). The cultures were incubated at 37 °C and 180 rpm until they reached an OD600 of 0.6, at which point the temperature was adjusted to 19 °C and IPTG was added to a final concentration of 0.5 mM. The cells were grown for additional 18 hours after which they were harvested by centrifugation (4000 g for 10 min). The sediment was resuspended in lysis buffer (60 mM phosphate, 250 mM NaCl, 11 mM imidazol, 5 mM β-mercaptoethanol pH=8 pH was adjusted before the addition of β-mercaptoethanol and imidazole via NaOH). Cells were lysed using a sonicator and centrifuged at 17000 g for 10 min at 4 °C. The lysate was loaded onto 0.5 mL Ni-NTA columns equilibrated with lysis buffer and incubated at 4 °C and slow rotation for a minimum of 2 hours. The column was washed with 2.5 mL of lysis buffer and the protein was eluted with 1.5 mL of elution buffer (60 mM phosphate, 250 mM NaCl, 230 mM imidazol, 5 mM β-mercaptoethanol pH=8, pH was adjusted before the addition of β-mercaptoethanol and imidazole via NaOH). The buffer was exchanged to 20 mM MOPS-NaOH-buffer (pH=7) using 5 mL Hi-Trap columns from GE Healthcare.

4.6.6. Crystallization, soaking data collection

Crystals were grown using the hanging drop method. 4-20 g/L Protein solution were mixed with precipitant solution containing PEG 8000 (20-30% (w/v)), NaCl (200 mM) and Tris-HCl-buffer (pH= 7-8, 100 mM). Crystals were grown for 10 weeks at 16 °C up to a size of 0.05x0.04x0.08 mm. Crystals were then transferred to cryo solution containing PEG 1500 (30%(w/v)), glycerol (20%(w/v)) NaCl (200mM) Tris-HCl-buffer (100 mM pH=8) and resveratrol-3-α-D-glucoside or nigerose (100 g/L) and soaked over night, mounted in cryo loops and plunged into liquid nitrogen. At beamline ID30B of the ESRF Grenoble the mounted crystals were placed within a 100K nitrogen gas stream and datasets were collected over 180° oscillation range. The datasets were auto indexed, integrated and scaled with XDS.

4.6.7. Structure determination and -refinement

The structures of apo SP and the two complexes were solved by molecular replacement using chain B of PDB entry 2GDV as a search model within PHASER.[110] After initial

refinement within Phenix,[138] regions with distinct conformational changes were manually rebuilt within COOT[111] and the appropriate ligands were modelled into the active site. After three more rounds of automated refinement and manual rebuilding including water and ligand placement, the R and R_{free} factors converged. Two conformations were modeled for Glu232 in PDB 5M9X as it displays a degree of flexibility. This is likely due to a subpopulation of ligand free protein molecules. Of note, several residues residing in β-turn regions appeared close to the border of or within the disallowed region of the Ramachandran plot. After inspection of the surrounding electron density we conclude that these are true outliers, most likely in a stressed conformation.

4.6.8. Data collection and refinement statistics

See **chapter 7.5**

4.6.9. Determination of the active site volume

The active site volume was determined with CAVER-Analyst 1.0 using the settings:

Outer Probe Radius: 2.29 Å

Inner Probe Radius: 1.80 Å

4.6.10. Determination of Kinetic Parameters

General Considerations:

The experiments required for determination of the activities at the lower acceptor substrate concentrations were challenging due to the detection limit of HPAEC-PAD and the presence of an excess of sucrose. This leads to relative high experimental errors, especially at the lower substrate concentrations.

(-)-epicatechin and (+)-catechin interfere with the pulsed amperometric detection resulting large experimental inaccuracies. These experiments where nonetheless included as the K_M-values are comparable with those of the other substrates. In addition previous experiments have shown that for catechin concentrations above 10 mM saturation was already reached.

The error-bars indicate the average of three repeats ± one standard deviation.

The direct linear plot follows the procedure described by Eisenthal and Cornish-Bowden.[139] The intersections were calculated and ranked using the algorythm included in **chapter 7.3** of this supplementary information.

Table 9 Kinetic parameters of BaSP Q345F a) 50 mM sucrose; b) 20 mM resveratrol

kcat [s^{-1}] K_M [mM]	Michaelis-Menten fit	Direct linear regression	Hanes Plot	Lineweaver-Burke Plot
k_{cat} resveratrol[a]	*0.131 ± 0.00179*	*0.131*	*0.131*	*0.131*
K_M resveratrol[a]	*0.924 ±0.0854*	*0.989*	*0.888*	*0.888*
k_{cat} quercetin[a]	*0.0941 ± 0.00325*	*0.096*	*0.094*	*0.094*
K_M quercetin[a]	*0.526 ± 0.0603*	*0.564*	*0.0524*	*0.0524*
k_{cat} fisetin[a]	*0.0673 ± 0.00433*	*0.071*	*0.066*	*0.066*
K_M fisetin[a]	*0.325 ± 0.120*	*0.612*	*0.262*	*0.262*
k_{cat} (-)-epicatechin[a]	*0.104 ± 0.00739*	*0.092*	*0.097*	*0.097*
K_M (-)-epicatechin[a]	*1.55 ± 0.453*	*0.787*	*1.11*	*1.11*
k_{cat} (+)-catechin[a]	*0.0743 ± 0.00953*	*0.074*	*0.069*	*0.069*
K_M (+)-catechin[a]	*0.953 ± 0.406*	*0.770*	*0.626*	*0.626*
k_{cat} naringenin[a]	*0.0221 ± 0.00015*	*0.022*	*0.022*	*0.022*
K_M naringenin[a]	*0.0782 ± 0.0099*	*0.085*	*0.078*	*0.078*
k_{cat} glucose[a]	*0.0258 ± 0.00143*	*0.024*	*0.026*	*0.026*
K_M glucose[a]	*41.5 ± 3.78*	*37.3*	*42.3*	*42.3*
k_{cat} phosphate[a]	*0.179 ± 0.0066*	*0.176*	*0.178*	*0.178*
K_M phosphate[a]	*25.7 ± 1.86*	*23.5*	*25.1*	*25.1*
k_{cat} sucrose[b]	*0.118 ± 0.00293*	*0.107*	*0.119*	*0.119*
K_M sucrose[b]	*5.41 ± 0.512*	*4.35*	*5.07*	*5.07*

Activity assays were performed at 37 °C in 50 mM MOPS-NaOH-buffer at pH=7 in at total volume of 100 µL containing 30% (v/v) DMSO, 50 mM of sucrose and 0.219 g/L BaSP Q345F or 0.307 mg/L (sucrose parameters) and 1.22 mg/L (phosphate parameters) BaSP wild type respectively. The acceptor molecules were added to obtain appropriate concentrations and samples of either 10 or 20 µL were taken at the appropriate timepoints, diluted with 230 to 490 µL water and the reaction was stopped by heating the samples at 95 °C for 6 minutes. Product formation was determined via HPAEC-PAD (High performance anion exchange chromatography with pulsed amperometric detection), by comparing the amount of glucose and fructose released, in the case of aromatic acceptors and phosphate and direct determination in the case of glucose used as an acceptor. This was necessary, since the glucosylated aromats cannot be directly detected at the concentrations occurring in the

Table 10 Kinetic parameters of BaSP wild type a) 50 mM sucrose; b) 20 mM phosphate

kcat [s^{-1}] K_M [mM]	Michaelis-Menten fit	Direct linear regression	Hanes Plot	Lineweaver-Burke Plot
k_{cat} phosphate[a]	*94.0 ± 3.82*	*97.6*	*95.3*	*95.3*
K_M phosphate[a]	*5.63 ± 0.611*	*6.12*	*5.74*	*5.74*
k_{cat} sucrose[b]	*75.4 ± 0.95*	*74.7*	*75.4*	*75.4*
K_M sucrose[b]	*0.908 ± 0.019*	*0.898*	*0.908*	*0.908*

applied assay. All experiments were performed in triplicates.

HPAEC-PAD Methods

HPAEC-PAD analysis was performed with a Dionex ICS-5000+ SP system utilizing a Carbopac PA10 column. Eluents were 100 mM NaOH (A), 100mM NaOH, 1 M NaOAc (B), 10 mM NaOH (C) and 250 mM NaOH (D).

Samples containing (+)-catechin and (-)-epicatechin and their glucosylated derivatives were resolved using a multistep gradient programmed as follows: 0 min to 9 min 100% A, flow = 0.25 mL/min, 10 min 100% B flow = 0.25, 11 min 100% B flow = 0.40. 11 min to 25 min 100% B flow = 0.4 mL/min, 25 min to 35 min 100% D flow = 0.4 mL/min, 35 min to 41 min 100% A flow = 0.4 mL/min, 41 min to 42 min 100% A flow = 0.25 mL/min, 42 min to 45 min 100% A flow = 0.25 mL/min.

Samples containing phosphate and glucose-1-α-D-phosphate were resolved using a multistep gradient programmed as follows: 0 to 5 min 100 A flow = 0.25 mL/min, 20 to 20.5 min 70% A , 30 % B flow = 0.25 mL/min , 21 min to 23 min 100% D flow = 0.5 mL/min, 23 min to 27.5 min 100% A flow = 0.5 mL/min, 28 min to 30 min 100% A flow = 0.25 mL/min.

Glucose concentration of samples containing fisetin or quercetin were resolved using a multistep gradient programmed as follows: The Amount of solution A was raised nonlinear, using curve 9 from 5% A and 95% C to 100% A in 20 minutes. 20 to 26 min 100% A, 26 to 32 min 5% A and 95% C. The flow was set to 0.25 mL/min.

Fructose concentrations of samples containing fisetin and quercetin and fructose and glucose concentrations of samples containing naringenin and resveratrol were determined under isocratic conditions with 100% A and a flow of 0.25 mL/min for 16 min.

Maltose and nigerose concentrations were determined under isocratic conditions with 97% A 3% B and a flow of 0.25 mL/min for 16 min.

4.6.11. Determination of yields

Reactions were performed in triplicates at 50 °C and contained 600 mM sucrose 100 mM MOPS-NaOH-buffer at pH=7, 30% (v/v) DMSO and 4.0 g/L BaSP Q345F in a final volume of 200 µL. Acceptor concentrations were 100 mM (-)-epicatechin and (+)-catechin, 75 mM resveratrol. 50 mM fisetin and quercetin and 25 mM (rac)-naringenin. 10 µL samples were diluted with 190 µL of water and inactivated at 95 °C for 5 minutes and diluted in 1 mL MeOH:water (final concentration 50% (v/v) MeOH), and subjected to HPLC-Analysis. Yields were calculated using the total area under all peaks to include trace products that cannot be isolated.

HPLC-Methods

Conversions of polyphenols were determined on analytical scale using a YMC-ODS-AQ column (C_{18} column; 5 µm, 4.6×250 mm) and a Sykam S3345 detector for resveratrol separation.

(-)-Epicatechin, (+)-catechin and its glucosylated derivatives were resolved using a linear ternary gradient programmed as follows Solvent A: 0.1% (v/v) TFA, B: Acetonitril, C: 20 mM $NH_4H_2PO_4$:
0 to 10 min 4% B 50% C, 25 min 28% B, 50% C, 27to 35 min 4% B 50% C. The flow rate was set to 1.0 mL/min and the detection wavelength to 220 nm.
(+)-Catechin-5-glucosid and catechin-3'-5-diglucosid were resolved using a linear ternary gradient programmed as follows Solvent A: 0.1 % (v/v) TFA, B: Acetonitril, C: 20 mM $NH_4H_2PO_4$:
0 to 1 min 4% B 50% C, 12 min 31% B, 50% C, 12 to 13 min 31% B, 50% C 14 to 21 min 4% B 50% C. The flow rate was set to 1.0 mL/min and the detection wavelength to 220 nm.

(rac)-naringenin and its glucosylated compound were resolved using a linear binary gradient programmed as follows: Solvent A: 0.1% (v/v) TFA, B: MeOH with 0.1% (v/v) TFA 0 to 3 min 40% B, 18 min 65% B, 19 to 25 min 40% B The flow rate was set to 1.0 mL/min and the detection wavelength to 220 nm.

Quercetin, Fisetin and their glucosylated compounds using a linear binary gradient programmed as follows: Solvent A: 20 mM $NH_4H_2PO_4$, Solvent B Acetonitrile. 0 to 5 min 10% B, 20 min 25% B, 21 to 25 min 50% B, 26 to 35 min 10%B. The flow rate was set to 1.0 mL/min and the detection wavelength to 320 nm.

Resveratrol and its glucosylated derivatives were resolved during 20 minutes using 65% of 0.1% TFA and 65% Acetonitrile for B, 20 min. The flow rate was set to 0.8 mL/min and the detection wavelength to 250 nm.

4.6.12. Determination of the resveratrol-enzyme interface

The area of the interface between the resveratrol moiety and BaSP Q345F was determined using the PISA-server. For this purpose all atoms of the carbohydrate moiety were manually deleted before submitting the pdb-file.

4.6.13. NMR- and MS-Data

See Appendix c**hapter 7.2**

4.6.14. A little tool to perform the direct linear plot analysis

See Appendix **chapter 7.3**

Chapter 5

REVERSIBILITY OF A POINT MUTATION INDUCED DOMAIN SHIFT: EXPANDING THE CONFORMATIONAL SPACE OF A SUCROSE PHOSPHORYLASE

Michael Kraus, Clemens Grimm and Jürgen Seibel

Published in Scientific Reports, volume 8, Article number 10490, 2018.

Summary

This chapter presents the missing crystallographic snapshot of the reaction coordinate of the BaSP Q345F variant. The crystal structure of BaSP Q345F in the *F*-conformation finally presents information about the binding of the donor substrate sucrose. Unlike assumed in **chapter 3** the *F*-conformation does not feature the domain shift. With this piece of information a complete image of the catalytic cycle of the useful new BaSP variant can be drawn and the mechanistic insight gained this way will aid in the design of the next generation of bespoke biocatalysts.

5.1. Abstract

Despite their popularity as enzyme engineering targets structural information about Sucrose Phosphorylases remains scarce. We recently clarified that the Q345F variant of *Bifidobacterium adolescentis* Sucrose Phosphorylase is enabled to accept large polyphenolic substrates like resveratrol via a domain shift. Here we present a crystal structure of this variant in a conformation suitable for the accommodation of the donor substrate sucrose in excellent agreement with the wild type structure. Remarkably, this conformation does not feature the previously observed domain shift which is therefore reversible and part of a dynamic process rather than a static phenomenon. This crystallographic snapshot completes our understanding of the catalytic cycle of this useful variant and will allow for a more rational design of further generations of Sucrose Phosphorylase variants.

5.2. Introduction

Sucrose Phosphorylases (CAZy Family GH 13,[45] SPs) are popular targets for enzyme engineering and employed in various transglucosylation reactions[40, 42] because they utilize the cheap and abundant donor substrate sucrose[70], exhibit thermostability[40] and organic solvent compatibility[40, 105]. Several SP variants have been created in recent years in order to establish a variety of novel transglucosylation reactions.[85, 91, 119, 137] While the native reaction of SPs is the interconversion of sucrose and α-D-glucose-1-phosphate via a covalent enzyme-glycosyl intermediate, [44] the two main target reactions for enzyme design are the synthesis of rare disaccharides[85, 137] and the glucosylation of polyphenols[119, 140, 141] Several crystal structures of *Bifidobacterium adolescentis* Sucrose Phosphorylase[43, 44] have elucidated the catalytic mechanism and substrate binding of wild type SPs[44] and constitute a common starting point of engineering strategies.

The most crucial insight into the mechanism of SPs was gained in 2006 when Mirza *et. al.* revealed the existence of two distinct conformations of BaSP, one responsible for the accommodation of sucrose, the other for α-D-glucose-1-phosphate.[44] BaSp switches between those two conformations via the rearrangement of two flexible loops: Loop A (^{336}AAASNLDLY^{344}, part of domain B') and loop B (^{132}YRPRP^{136}, part of domain B). Of note, the invariant residue Gln345 targeted in this study borders loop A. but maintains the identical position in both loop conformations. In the sucrose binding conformation (2gdu, 1r7a, 2gdv Chain A) loop A points into the active site and Asp342 becomes part of the acceptor binding site forming a H-bond to OH-3 of the fructosyl moiety. The sidechain of Tyr344 faces away from the +1 site and is not involved in substrate binding.

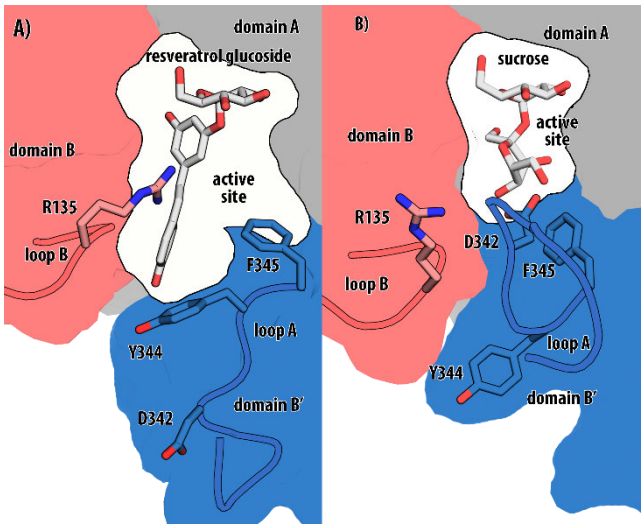


Figure 34 Schematic representation of BaSP Q345F loop conformations. Blue: Domain B', Red: Domain B, Grey: Domain A, White: active site cavity. A: Aromat binding conformation of BaSP Q345F (in complex with resveratrol-α-D-glucoside (PDB ID code 5man), B: Sucrose binding conformation of BaSP Q345F, sucrose superimposed from 2gdu). It should be noted that the increase in acceptor site space is due to the domain shift and not a result of the loop rearrangement.

The proposed phosphate binding conformation[44] (2gdv, Chain B, features loop A facing away from the active site, while the sidechain of Tyr344 now points into it and contributes to solvent shielding. The change in loop B mostly consists in the rearrangement of Arg135 which is now oriented towards the catalytic centre and facilitates phosphate binding through its positive charge.[44] Neither conformation features a defined access channel, consequently, access via substrate diffusion must occur by one or more unknown open conformations.

The first structural insight into the mechanism of sucrose phosphorylase variant Q345F was presented recently.[119, 140] The BaSP Q345F variant features a loop orientation that resembles the wildtype phosphate binding conformation (**Figure** 34). The two key differences between the wildtype and variant crystal structures are (1): the orientation of Tyr344, which is not part of the active site, and (2): a movement of the entire domain B by 3.3 Å[119, 140] referred to as domain shift. We recently demonstrated that this domain shift is in fact responsible for the altered acceptor specificity spectrum that enabled the synthesis reaction of resveratrol-3-α-D-glucoside and nigerose[119, 137, 140] While we were able to demonstrate that the domain shift is ligand independent,[140] the question how the variant binds sucrose and whether the domain shift is static and permanently present in the Q345F variant remained unsolved. We now present the missing link, a crystal structure of BaSP Q345F in the sucrose binding conformation. Based on this structural evidence the full catalytic cycle of this BaSP variant is now understood and can now be targeted by further design studies.

5.3. Results

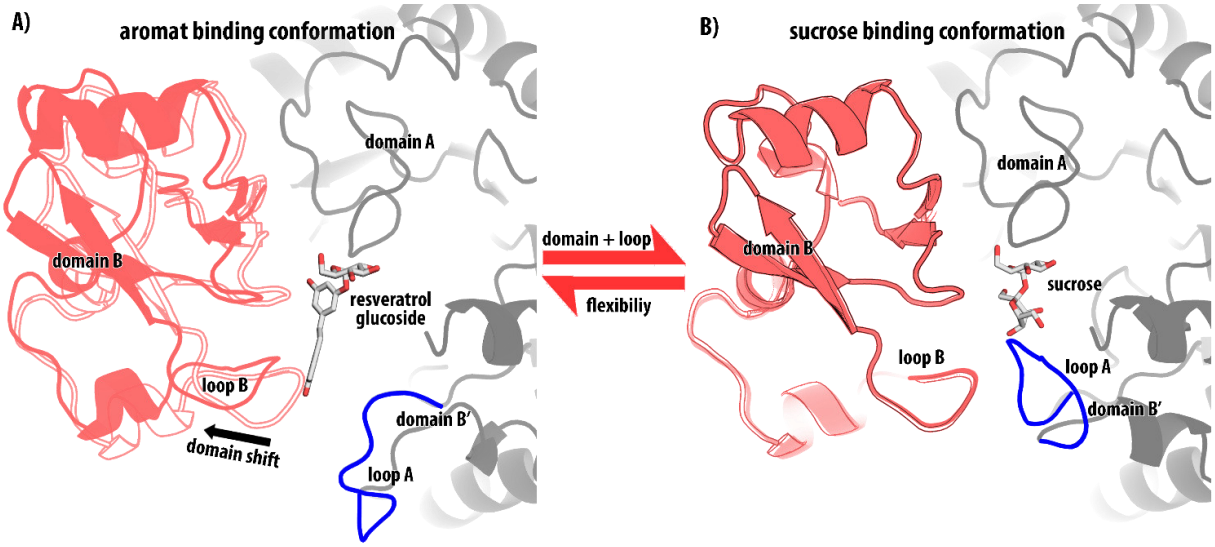


Figure 35 Reversible domain shift induced by the Q345F exchange. The domain shift occurs together with the rearrangement of loop A (blue) A: Aromat binding conformation of BaSP Q345 F in complex with the resveratrol-3-α-D-glucoside. Domain B (red) shifts by 3 Å (red outlines). B: Sucrose binding conformation of BaSP Q345F doamin B (red) occupies the same condition as found in the wildtype (red outline). The two crystal structures likely represent the two extremes of a dynamic equilibrium.

In contrast to all previous structures of the Q345F variant, no domain shift is observed in the new crystal structure presented here. This shows that the domain shift is reversible and part of a dynamic process. (**Figure 35**) The orientation of residue Asp342 presents the sole significant difference to wildtype BaSP (**Figure** 36). This residue usually interacts with the 4-OH group of fructose and is rotated by 81° towards the -1 site relative to its orientation found in the wild type enzyme. The phenyl ring of Phe345 is rotated by 31° around the Cβ-Cγ axis relative to the amide of Gln345. During the loop rearrangement and domain shift the benzene

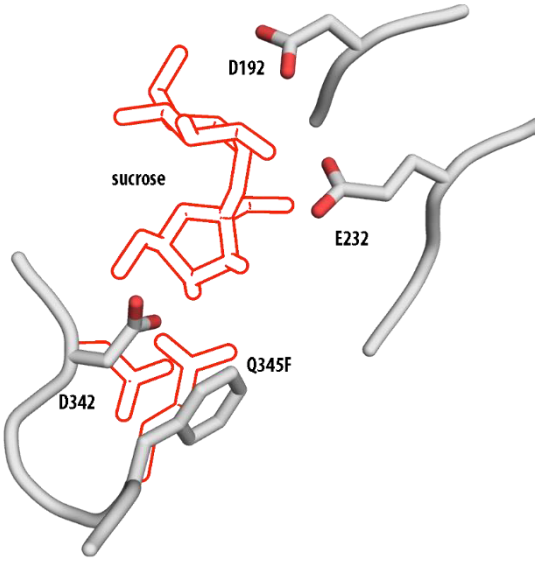


Figure 36 Key active site residues of BaSP Q345F in the sucrose binding conformations. The outlined side chaines represent the key differences between the wild type and the variant: The rotation of D342 and the Q345F exchange. The position of sucrose from the aligned wild type structure is indicated as outlines as well.

ring of Phe345 rotates by 82°. This rotation is observed in all structures that display the domain shift, regardless which, if any ligands are present and causes the displacement of the neighbouring Tyr344 which is also linked to the domain shift.

The BaSP Q345F variant displays a lowered affinity for (K_M = 5.41 mM vs 0.91 mM wildtype) and activity towards (8.6% of the wildtype) sucrose.[119, 140] This was to some degree expected as Gln345 exhibits hydrogen bonds with OH-3 and OH-6 of the fructosyl moiety of sucrose.[44] To evaluate the impact of the Q345F exchange and the rotation of Asp342 on sucrose binding docking studies with Autodock Vina were conducted. Docking of sucrose into BaSP Q345F yields a relative binding energy that is 10.1 kcal lower than the one of the wild type. This is due to the orientation of Asp342, which clashes with OH-4 and OH-6 of fructose and prevents sucrose from assuming the correct position in the enzyme (**Figure** 37**A**). When Asp342 is defined as flexible the docking results show it facing away from its natural position by 88° (**Figure** 37**B**). This does not recover its native H-bond but removes the steric clash and the difference in affinity to the wildtype is reduced to 2.0 kcal/mol. If Phe345 is set as a further flexible residue it rotates slightly and Asp342 can now adopt its native conformation and the total loss of affinity is reduced to 1.5 kcal/mol (**Figure** 37**C**). The loss of the H-bonds between Gln345 and the fructosyl moiety were expected to reduce the affinity for sucrose. To examine this influence Gln345 was exchanged *in silico* against alanine to remove any

Table 11 Docking of sucrose into BaSP wild type and BaSP Q345F as well as in-silico generated variants.

	flexible residues	Relative Affinity in kcal/mol
wild type (2gdu)	-	*-12.8*
Q345A (from 2gdu)	-	*-11.9*
D342A (from 2gdu)	-	*-12.4*
Q345F	-	*-2.7*
Q345F	*D342*	*-10.8*
Q345F	*D342, Q345F*	*-11.3*

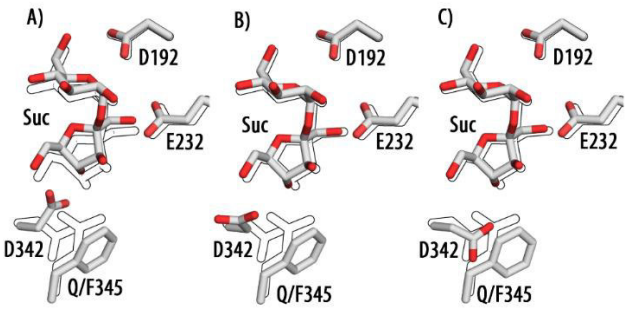


Figure 37 Orientation of sucrose and key side chains in BaSP Q345F as calculated by AutodockVina The outlines indicate the orientation of sucrose and the amino acid in the wild type (PDB ID 2gdu) A: all side chains rigid B: Asp342 defined as flexible C: Asp342 and Phe345 defined as flexible

interactions and docking with sucrose was performed. The result indicates that Gln345 contributes by ca. 0.9 kcal/mol to the donor binding. The remaining 0.6 kcal/mol difference is likely due to a slight steric hindrance induced by Pher345 (**Table 11**). It can be concluded that the Q345F exchange did not affect sucrose accommodation beyond the loss of the polar interactions of Gln345.

5.4. Discussion

The previously observed domain shift is responsible for the ability of BaSP Q345F to glucosylate resveratrol and further polyphenols as well as synthesize nigerose.[10] The fact that the domain shift is absent in the sucrose binding conformation of BaSP Q345F indicates a reversible, dynamic process induced by the mutation. We conclude from the crystal structure presented here that BaSP Q345F exists as an equilibrium and the crystal structures represent the two endpoints of the process: The open conformation, which is required for the glycosylation of polyphenolic acceptors and the sucrose binding conformation which is needed for donor substrate conversion.

From these structures the catalytic mechanism of BaSP Q345F can be proposed: First sucrose enters BaSP Q345F via an open conformation which features the domain shift. The enzyme binds sucrose and the loop rearrangement occurs and simultaneously domain B shifts back into the wildtype-like position. The loss of H-bonds between Gln345 and the fructosyl moiety and a minor steric clash from Phe345 reduce the binding energy of sucrose, resulting in a lowered activity of the variant compared to the wild type otherwise the mutation does not affect sucrose recognition. Then fructose cleavage takes place in the sucrose binding conformation, which is identical for both wildtype and variant. Subsequently the loop rearrangement and the domain shift occurs again and BaSP Q345F transforms into a glucosyl-linked open conformation. A polyphenolic acceptor is recruited and after glucosyl transfer and product release the cycle begins anew. The reversibility of the domain shift is necessary for the efficient conversion of sucrose the domain shift itself for the ability to glucosylate the aromats. Further investigations into the mechanics of the domain movement, while beyond the scope of this publication, could lead to the ability to predict this phenomenon. A reversible domain shift may then constitute a powerful novel tool for enzyme design as it allows to drastically alter the active site of a flexible enzyme, while maintaining the original structural features of the protein simulatiously.

5.5. Conclusions

In summary the domain shift of BaSP Q345F introduced by the mutation is not a static effect but part of a dynamic process. Sucrose binding by BaSP Q345F takes place in a manner closely related to the wild type and the Q345F mutation has no effect in sucrose coordination aside from the predicted loss of the H-bonds between Gln345 and OH-3 and OH-6 of fructose. The existence of a BaSP Q345F conformation without the domain shift allows further interpretation of previous results. Now the complete set of conformations required for the catalytic cycle of BaSP Q345F are known and this information can be used for further engineering of the versatile Sucrose Phosphorylases and potentially be transferred to other memebers of the vast glycosdase family GH13.

5.6. Experimental

5.6.1. Cloning expression and purification of BaSP Q345F

As previously described[119]: Freeze-dried cultures of *B. adolescentis* (DSMZ 20083) were obtained from DSMZ (Deutsche Sammlung von Mikroorgansimen und Zellkulturen GmbH), and grown under anaerobic conditions in DSMZ medium Nr.58 without resazurin. Cells were harvested and the genomic DNA isolated, using a GenJet Genomic DNA purification Kit (Thermo Fisher). The BaSP gene was amplified from genomic DNA using the primers

5'-ATAACCATGGCTATGAAAAACAAGGTGCAGCTCATCAC-3' and 5'-CAATCCGCCTGTCGTCGCCCTCGAGTAAT-3'. The amplicon was inserted into pET-28b(+) using the NcoI and XhoI restriction sites yielding plasmid pET-28b(+)-BaSP-wt.

5.6.2. Construction of BaSP Q345F

As previously described[119]: The Q345F mutation was constructed applying the Megaprimer method. The mutagenic primers 5'-CCAATCTCGACCTCTACTTCGTCAACAGCACCTAC-3', and 5'-CAATCCGCCTGTCGTCGCCCTCGAGTAAT-3' were used for the creation of the megaprimer and 5'-ATAACCATGGCTATGAAAAACAAGGTGCAGCTCATCAC-3' was used for the second PCR. Cloning and purification of the variant followed the procedure described for the wildtype.

5.6.3. Cloning Expression and Purification BaSP Q345F

As previously described[119]: *E. coli* Bl21 star™ cells were heat shock transformed with plasmid pET-28b(+)-BaSP-wt. Overnight cultures of the transformed host in LB-medium containing 50 mg/L kanamycin sulfate were grown and 1.8 mL was used to inoculate 250 mL of LB-medium (50 mg/L kanamycin sulfate). The cultures were incubated at 37 °C and

180 rpm until they reached an OD600 of 0.6, at which point the temperature was adjusted to 19 °C and IPTG was added to a final concentration of 0.5 mM. The cells were grown for additional 18 hours after which they were harvested by centrifugation (4000 g for 10 min). The sediment was resuspended in lysis buffer (60 mM phosphate, 250 mM NaCl, 11 mM imidazol). Cells were lysed using a sonicator and centrifuged at 17000 g for 10 min at 4 °C. The lysate was loaded onto 0.5 mL Ni-NTA columns equilibrated with lysis buffer and incubated at 4 °C and slow rotation for a minimum of 2 hours. The column was washed with 2.5 mL of lysis buffer and the protein was eluted with 1.5 mL of elution buffer (60 mM phosphate, 250 mM NaCl, 230 mM imidazol). The buffer was exchanged to 20 mM MOPS-NaOH-buffer (pH=7) using 5 mL Hi-Trap columns from GE Healthcare.

5.6.4. Crystallization, soaking data collection

Crystals were grown using the hanging drop method. 4-20 g/L Protein solution were mixed with precipitant solution containing PEG 8000 (20-30% (w/v)), NaCl (200 mM) and Tris-HCl-buffer (pH= 7-8, 100 mM). Crystals were grown for 10 weeks at 16 °C up to a size of 0.05x0.04x0.08 mm. Crystals were then transferred to cryo solution containing PEG 1500 (30%(w/v)), glycerol (20%(w/v)) NaCl (200mM) Tris-HCl-buffer (100 mM pH=8) and plunged into liquid nitrogen. At beamline ID30B of the ESRF Grenoble the mounted crystals were placed within a 100K nitrogen gas stream and datasets were collected over 180° oscillation range. The datasets were auto indexed, integrated and scaled with XDS.

5.6.5. Structure determination and -refinement

The structures of BaSP Q345F was solved by molecular replacement using chain B of PDB entry 2GDV as a search model within PHASER.[110] After initial refinement within Phenix,[138] regions with distinct conformational changes were manually rebuilt within COOT[111] and the appropriate ligands were modelled into the active site. After three more rounds of automated refinement and manual rebuilding including water and ligand placement, the R and R_{free} factors converged.

5.6.6. Crystal structure data collection and refinement statistics

See Appendix **chapter 7.5**

5.6.7. Docking

The crystal structures of the BaSP E232Q (PDB ID 2gdu, chain A) and BaSP Q345F F-conformation (PDB 6FME) were used as “receptor” for the docking calculations. All water molecules and ligand entries were removed, non-polar hydrogens were added using AutoDockTools 1.5.6r.[127] For dockings with flexible residues the respective amino acids were defined as flexible and Gasteiger charges were added and rotatable bonds were assigned

using AutoDockTools. The *in silico* mutations were introduced with pymol 1.8.0.3 using the mutagenesis wizard. Grid box center and grid dimensions (20x20x20 Å, grid spacing: 1.0 Å) were determined via AutoDockTools and transferred to the AutoDockVina configuration file.

As ligand sucrose as present in the crystal structure of BaSP E232Q was used. Gasteiger charges were added and rotatable bonds were assigned using AutoDockTools.

AutoDockVina[125] was used for docking calculations. The docking parameters “exhaustiveness” and “energy_range” were set to “20” and “8”, respectively.

Chapter 6

SUMMARY

Strategy

The initial goal was the conversion of *Bifidobacterium adolescentis* Sucrose Phosphorylase (BaSP) into a polyphenol glucosidase by structure based enzyme engineering. BaSP was chosen because of its ability to utilize sucrose, an economically viable and sustainable donor substrate, and transfer the glucosyl moiety to various acceptor substrates. Additionally it offers thermostability and organic solvent compatibility. Finally crystal structures of BaSP are available and allow a rational choosing of the mutagenesis targets. Previously BaSP was crystallized in two conformations the fructose binding conformation (*F*-conformation) and the phosphate binding conformation (*P*-conformation). The different conformations are achieved via the rearrangement of two loops.

The active site of wild type BaSP offers neither sufficient space nor a fitting polarity for the target polyphenolic acceptor substrates. These are significantly larger and less polar than the natural substrates phosphate and sucrose. This is in accordance with the finding that wild type BaSP only converts trace amounts of the target acceptor substrates resveratrol and quercetin. Instead an unwanted side reaction predominates: Hydrolysis of sucrose and a subsequent formation of glucose-glucose disaccharides from sucrose and hydrolytically produced sucrose.

In order to overcome this problem a mutagenesis strategy was developed: The introduction of aromatic residues into the active site was considered a viable way to render it more suitable for aromatic acceptor compounds by reducing its polarity and potentially introducing π-π-interactions with the polyphenols.

An investigation of the active site revealed Gln345 as a suitable mutagenesis target. This residue is at sufficient distance to the -1-site as not to disturb the donor binding and it occupies the same position in both loop conformations. It was exchanged against all four aromatic amino acids histidine, phenylalanine, tyrosine and tryptophan and of the four variants one, BaSP Q345F, maintained sufficient activity and displayed transfer to the target substrates. The detailed examination of this variant is presented in 3 published studies and one unpublished study.

Kraus M., Grimm C., Seibel J., *ChemBioChem*, 2016, 17, 33-36:

As a proof of concept BaSP Q345F was employed in the glycosylation of (+)-catechin, (-)-epicatechin and resveratrol. The variant was selective for the aromatic acceptor substrates and the glucose disaccharide side reaction was only observed after almost quantitative conversion of the aromatic substrates. The glucosylation of resveratrol gave a single product resveratrol-3-α-D-glucoside. Two of the five potential glucosylation sites of catechin and epicatechin were addressed by BaSP Q345F and a mixture of mono- and di-glucosides was obtained. Simultaneously the variant displays a lowered activity for sucrose and binds it less efficiently.

A crystal structure of BaSP Q345F in complex with glucose was obtained and it displayed an unexpected shift of an entire domain by 3.3 Å. Orientation of the flexible loops resembled the *P*-conformation. The domain shift created a widened active site and a visible access channel leading to the hypothesis that this creation of space allows the accommodation of the large target substrates. No corresponding channel is observed in wild type BaSP crystal structures. The crystal structure also revealed the unexpected distortion of the donor binding site and an inefficient binding of glucose. In particular, the H-bonds between OH-6 of glucose and His88, Gln164 and Asp192 were disturbed.

The desired activity was successfully installed and a possible structural cause for the change in selectivity appeared. In the absence of an aromatic ligand in the crystal structure no definitive conclusion about the role of the domain shift could be made.

Kraus M., Görl J., Timm M., Seibel J., *Chemical Communications*, 2016, 52, 4625-4627:

An investigation into the background reaction, the formation of glucose-glucose disaccharides of BaSP Q345F and three further variants that addressed the same region (L341C, D316C-L341C and D316C-N340C) revealed the formation of nigerose by BaSP Q345F. Nigerose, the rare α-O-(1,3)-linked glucose disaccharide is not synthesized by BaSP wild type that produces maltose (α-O-(1,4)-linked) and kojibiose (α-O-(1,2)-linked). BaSP Q345F also synthesizes maltose but not kojibiose and remarkably the nigerose to maltose ratio can be shifted towards nigerose by the addition of DMSO.

Docking studies were performed in order to explain the shift in regio-selectivity. For this purpose two assumptions were made: 1. Glucose is used as an acceptor via the *F*-conformation. 2. The domain shift is present in all conformations of BaSP Q345F. The dockings suggested that the domain shift induced movement of Tyr196 removes a steric clash with glucose in the orientation needed for nigerose synthesis and thus permits it.

Kraus M., Grimm C., Seibel J., *Chemical Communications*, 2017, 53, 12181-12184

A crystal structure of BaSP D192N-Q345F, an inactive variant in complex with resveratrol-3-α-D-glucosid, the glucosylation product of resveratrol, synthesized by BaSP Q345F was solved. It proved that the domain shift is in fact responsible for the ability of the variant to glycosylate aromatic compounds. Simultaneously a ligand free crystal structure of BaSP Q345F disproved an induced fit effect as the cause of the domain shift. Investigation into ligand binding revealed a critical hydrogen bond with the catalytic acid/base Glu232 and explains the preference of BaSP Q345F for aromatic 1,2- or 1,3 diols. The π-system of resveratrol is also oriented in a way to permit T-shaped π-π-interactions with Phe156 and the engineered residue Phe345. While the distortion of the -1-site persists, glucose adopts its natural position and the missing H-bonds to its OH-6 are bridged via a water molecule coordinated by His88, Gln164 and Asp192.

A detailed kinetic study of BaSP Q345F with eight acceptor substrates reveals a high affinity of the variant for the aromatic substrates (K_M 0.08 to 1.55 mM), which is even higher than the affinity of the wild type for its natural substrate phosphate (5.61 mM).

Unexpectedly a crystal structure of BaSP Q345F in complex with nigerose displays the domain shift and falsifies the assumption that all glucose-glucose disaccharides are formed via the *F*-conformation. Instead the domain shift creates a novel multifunctional active site.

Kraus M., Grimm C., Seibel J., *Scientific Reports*, accepted for publication

The missing link, a crystal structure of BaSP Q345F in the *F*-conformation is obtained. This does not feature the domain shift, but is in outstanding agreement with the wildtype structure. The domain shift is therefore not static but rather a step in a dynamic process. It is further conceivable that the domain shifted conformation of BaSP Q345F resembles the open conformation of the wild type and that an adjustment of a conformational equilibrium as a result of the Q345F point mutation is observed.

Docking studies indicate that the loss of activity for sucrose is mainly a result of the loss of the H-bonds between fructose and Gln345 coupled with a slight steric clash with the engineered Phe345.

With this data the complete catalytic cycle of the SP' variant BaSP Q345F has been observed.

Conclusion:

An useful catalyst for the synthesis of glucosylated polyphenols has been established via structure based enzyme engineering. The BaSP variant Q345F prefers transfer to larger aromatic substrates and possesses a high affinity for them (K_M~ 1 mM)

The point mutation in BaSP Q345F triggers a domain shift that is part of a dynamic process and leads to a novel multifunctional active site. First sucrose is recruited into the active site comparable to the wild type and the donor glucose is covalently attached to the enzyme. Then the domain shift takes place and sufficient space for the binding of the target aromatic substrates is created in the active site. The accommodation of resveratrol and related compounds is further aided by π-π-interactions with Phe156 and Phe345. The same domain shift enables the production of the rare disaccharide nigerose.

The enzymes, products and structures presented in this work are one additional step on the path to tailor made biocatalysts.

ZUSAMMENFASSUNG

Strategie:

Saccharose Phosphorylase aus *Bifidobacterium adolescentis* (BaSP) sollte durch strukturbasiertes Enzym-Engineering in die Lage versetzt werden Polyphenole zu glukosylieren. BaSP katalysiert die Bildung von Glukose-α-D-1-Phosphat aus Saccharose sowie die Rückreaktion. Das Enzym ist dadurch in der Lage den günstigen nachwachsenden Rohstoff Saccharose als Glukosedonor zu verwenden. BaSP vermag weiterhin unterschiedliche Klassen an Akzeptorsubstraten umzusetzen. Zusätzlich handelt es sich bei BaSP um ein relativ thermostabiles Enzym, dass seine Aktivität auch in der Gegenwart organischer Lösemittel nicht verliert. Schließlich konnte auf der Basis publizierter Kristallstrukturdaten eine rationale Auswahl der Mutagenesestellen erfolgen. Frühere Arbeiten zeigten, dass BaSP in zwei unterschiedlichen Konformationen vorliegt. Saccharose und Fruktose werden von der *F*-Konformation erkannt. Die Koordination von Glukose-α-D-1-phosphat und Phosphat erfolgt über die *P*- Konformation. Der Wechsel zwischen beiden Konformationen geht mit der Umlagerung zweier Loops einher.

Polyphenole wie Resveratrol und Quercetin sind signifikant größer und unpolarer als die als die natürlichen Substrate Saccharose und Phosphat. Die Aktive Tasche von Wildtyp BaSP bietet weder ausreichen Raum noch eine Oberfläche geeigneter Polarität um Polyphenole effizient zu binden. Dies ist in Einklang mit der Beobachtung, dass Wildtyp BaSP Resveratrol und Quercetin nur in Spuren umsetzt und stattdessen die Hydrolyse von Saccharose und in Folge die Bildung von Glukose-Glukose Disacchariden bevorzugt.

Um diese Einschränkungen zu umgehen wurde folgende Strategie entwickelt: In die katalytische Tasche sollten aromatische Seitenketten eingeführt werden um die Polarität an jene der gewünschten Akzeptorsubstrate anzupassen und eine weitere Stabilisierung durch π-π-Wechselwirkungen zwischen Enzym und Substrat zu erlauben.

Eine genaue Analyse der katalytischen Tasche führte zur Konzentration auf Gln345 als Mutageneseziel. Diese Aminosäure befindet sich in ausreichendem Abstand zur -1-subsite und nimmt in beiden Enzymkonformationen dieselbe Stelle und Orientierung ein. Gln345 wurde durch alle vier natürlichen aromatischen Aminosäuren (Histidin, Phenylalanin, Tyrosin und Tryptophan) ersetzt. Eine der vier Varianten BaSP Q345F war in der Lage die Zielsubstrate zu glukosylieren und behielt gleichzeitig ausreichen Aktivität bei. Die weitere Untersuchung dieses Enzyms ist in vier Studien beschrieben.

Kraus M., Grimm C., Seibel J., *ChemBioChem*, 2016, 17, 33-36:

Die Variante BaSP Q345F ist in der Lage Resveratrol (+)-Catechin und (-)-Epicatechin zu glukosylieren. Die Variante setzt die aromatischen Substrate selektiv um, und die Bildung von Glukose-Glukose Disacchariden erfolgt erst nachdem die Polyphenole verbraucht sind. Die Glukosylierung von Resveratrol verläuft regioselektiv und führt zu einem einzigen Produkt, Resveratrol-3-α-D-glukosid. Im Fall von (+)-Catechin und (-)-Epicatechin werden zwei der fünf möglichen Glukosylierungstellen adressiert und eine Mischung aus Mono- und Diglukosiden wird gebildet. Die Aktivität der Variante gegenüber Saccharose ist auf 8.6% der Aktivität des Wildtyps reduziert und die Affinität von BaSP Q345F zu Saccharose sinkt ebenfalls.

Eine Kristallstruktur von BaSP Q345F im Komplex mit Glukose wurde gelöst. Dabei wurde eine unerwartete Verschiebung einer kompletten Domäne um 3.3 Å beobachtet. Die Orientierung der flexiblen Loops entspricht dabei derjenigen der *P*-Konformation. Durch die Verschiebung der Domäne vergrößert sich die aktive Tasche und die Bindung der sterisch anspruchsvollen Polyphenolsubstrate wird ermöglicht. Außerdem ist die aktive Tasche der Variante im Gegensatz zu den bereits bekannten Wildtyp Strukturen nach außen offen. Desweitern wirkt sich die Verschiebung der Domäne auch auf die Donorbindungsstelle aus und führt zu einer ineffizienteren Bindung von Glukose. Verantwortlich dafür ist der Verlust der Wasserstoff-Brücken-Bindungen zwischen His88, Gln164 und Asp192 zur OH-6 der Glukose.

Die gewünschte Aktivität, der Transfer von Glukose auf aromatische Substrate wurde erhalten und die wahrscheinlich verantwortliche strukturelle Veränderung, die Verschiebung der Domäne wurde deutlich. Ohne die Anwesenheit eines aromatischen Liganden kann allerdings keine definitive Aussage über die Bedeutung der Domänenverscheibung für die neue Aktivität getroffen werden.

Kraus M., Görl J., Timm M., Seibel J., *Chemical Communications*, 2016, 52, 4625-4627:

Eine genauere Untersuchung der Nebenreaktion, der Bildung von Glukose-Glukose Disacchariden von BaSP Q345F und drei weiteren Varianten mit Mutationen im selben Berreich (L341C, D316C-L341C und D316C-N340C) zeigte, dass BaSP Q345F in der Lage ist Nigerose zu synthetisieren. Wildtyp BaSP produziert Kojibiose (α-O-(1,2)-verküpft) und Maltose (α-O-(1,4)-verknüpft) aber nicht den seltenen Zucker Nigerose (α-O-(1,3)-verknüpft). BaSP Q345F bildet ebenfalls Maltose aber keine Kojibiose. Interessanterweise kann der Anteil von Nigerose in der Produktmischung durch die Zugabe von DMSO gesteigert werden.

Um die Änderung der Regioselektivität von BaSP Q345F in Bezug auf Glukose zu erklären wurden Docking-Studien durchgeführt. Diese wurden auf der Basis von zwei Annahmen angefertigt. 1. Akzeptorglukose wird über die *F*-Konformation gebunden. 2. Die Domänenverschiebung ist in beiden Konformationen präsent. Im Wildtyp wird ein sterischer Block zwischen Tyr 196 und der für die Bildung von Nigerose notwendigen Glukose Konformation beobachte. Die Dockings legen nahe, dass dieser durch Verschiebung von Tyr196 in der Q345F Variante verschwindet und somit die Bildung von Nigerose ermöglicht wird.

Kraus M., Grimm C., Seibel J., *Chemical Communications*, 2017, 53, 12181-12184

Die Kristallstruktur einer inaktiven Variante, BaSP D192Q-Q345F in komplex mit dem Glukosylierungsprodukt Resveratrol-3-α-D-Glukosid wurde gelöst. Dadurch konnte gezeigt werden, dass die Verschiebung der Domäne in der Tat für die Fähigkeit der Variante Glukose auf aromatische Substrate zu übertragen, verantwortlich ist. Eine weitere, ligandenfreie Kristallstruktur von BaSP Q345F konnte ausschließen, dass die Verschiebung der Domäne durch die Substrate initiiert ist. Eine genaue Analyse der Ligandenbindung ergab, dass die E232, die katalytische Säure/Base eine Wasserstoff-Brücken-Bindung mit der 5-OH Gruppe von Resveratrol eingeht. Dies erklärt die Präferenz von BaSP Q345F für Aromaten mit 1,2- und 1,3-Diolmotiven. Die Orientierung des π-Systems von Resveratrol erlaubt weiterhin T-förmige π-π-Wechselwirkungen mit Phe156 und Phe345. Die Verzerrung der -1-site besteht fort, Glukose nimmt jedoch denselben Platz wie im Wildtyp ein. Die Wasserstoff-Brücken-Bindungen zwischen der 6-OH von Glukose und His88, Gln164 und Asp192 werden durch ein Wassermolekül überbrückt.

Die detaillierte kinetische Untersuchung von BaSP Q345F mit acht Akzeptorsubstraten ergab eine starke Affinität der Variante zu den aromatischen Substraten (K_M 0.08 bis 1.55 mM). Dies übersteigt sogar die Affinität des Wildtyps für das natürliche Substrat Phosphat(K_M 5.61 mM).

BaSP Q345F in Komplex mit Nigerose zeigt überraschenderweise ebenfalls die Domänenverschiebung und wiederlegt somit die Annahme, dass die Übertragung auf Glukose zwingend über die *F*-Konformation erfolgt. Die Verschiebung der Domäne erschafft eine neue multifunktionale aktive Tasche.

Kraus M., Grimm C., Seibel J., *Scientific Reports*, accepted for publication

Das fehlende Puzzelteil, eine Kristallstruktur von BaSP Q345F in der *F*-Konformation wurde gelöst. In dieser Konformation findet sich keine Verschiebung der Domäne, stattdessen stimmen die Strukturen von Variante und Wildtyp nahezu vollständig überein. Daraus folgt,

dass die Verschiebung der Domäne Teil eines dynamischen Prozesses ist. Es ist weiterhin denkbar, dass die Domänen-verschobene Form von BaSP Q345F einer offenen, bisher unbeobachteten Form des Wildtyps entspricht und die Punktmutation Q345F eine Verschiebung des Verhältnisses der unterschiedlichen Konformationen zur Folge hat.

Docking-Studien deuten darauf hin, dass die Aktivitätsverringerung in Bezug auf Saccharose auf den Verlust von Wasserstoff-Brücken-Bindungen zwischen Gln345 und Fruktose sowie auf eine schwache sterische Hinderung durch Phe345 zurück zu führen sind.

Mit diesen Informationen ist nun der komplette katalytische Zyklus von BaSP Q345F beschrieben.

Schlussfolgerung

Durch strukturbasiertes Enzym-Engineering wurde ein nützlicher Biokatalysator für die Synthese von Glukosylierten Polyphenolen geschaffen. Die Variante BaSP Q345F bevorzugt große polyphenolische Substrate und besitzt eine hohe Affinität zu diesen (K_M~ 1 mM).

Die Punktmutation Q345F löst eine Verschiebung einer Domäne aus. Dies ist Teil eines dynamischen Prozesses und führt zu einer neuen, multifunktionalen aktiven Tasche. Zunächst wird Saccharose in der aktiven Tasche gebunden und Glukose wird kovalent mit dem Enzym verknüpft. Dies geschieht auf eine mit dem Wildtyp vergleichbare Art und Weise. Dann findet die Umlagerung der Domäne statt und die aktive Tasche bietet ausreichen Platz für die größeren Polyphenole. Deren Bindung wird durch π-π-Wechselwirkungen mit Phe156 und Phe345 unterstützt. Die Verschiebung der Domäne ermöglicht darüber hinaus auch die Bildung von Nigerose.

Die Enzyme, Produktmoleküle und Kristallstrukturen in dieser Arbeit sind ein weiterer Schritt auf dem Weg zu maßgefertigten Biokatalysatoren.

Chapter 7

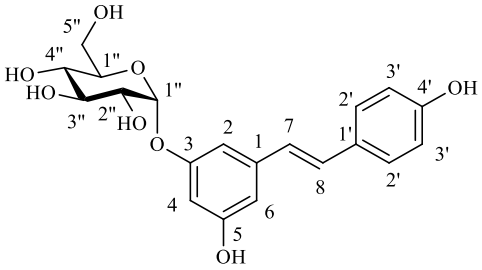
APPENDIX

7.1. Abbreviations

BASP: *Bifidobacterium adlosecens* Sucrose Phosphorylase
CAZy: Carbohydrate active Enzyme Database
DMSO: Dimethyl Sulfoxide
GH: Glycosyl Hydrolase
GT: Glycosyl Hydrolase
HPAEC-PAD: High Performance Anion Exchange Chromatography with Pulsed Amperometric Detection
HPLC: High Performance Liquide Chromatoraphy
LMSP: *Leuconostoc mesenteroides* Sucrose Phosphorylase
MS: Mass Spectrometry
NMR: Nuclear Magnetic Resonance (Spectroscopy)
Ni-NTA: Nickel Nitrolo Triacetic Acid
PCR: Polymerase Chain Reaction
SP: Sucrose Phosphorylase
TAA: Taka alpha-Amylase

7.2. NMR- and MS-Data

7.2.1. Resveratrol-3-O-α-D-glucosid (31)



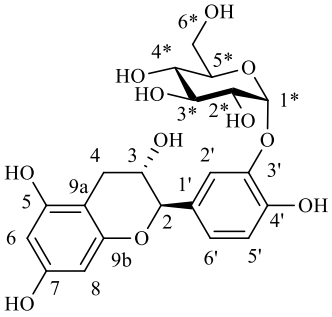
^{1}H-NMR (400 MHz; CD_3OD): δ= 7.38-7.36 (m, 2H, *J*= *8.2* Hz, *H*-2'), 7.03-6.99 (d, 1H, *J*= 16.4 Hz, *H*-8), 6.87-6.83 (m, 2H, *H*-2, *H*-7), 6.78 -6.76 (m, 2H, *J*= 8.7 Hz, *H*-3'), 6.63-6.62 (dd, 1H, *H*-6), 6.52-6.51 (dd, 1H, *H*-4), 5.48-5.47 (d, 1H, *J*= 3.7 Hz, *H*-1''), 3.88 (dd, 1H, *J*=9.7Hz 9.0 Hz, *H*-3''), 3.80-3.67 (m, 3H, *H*-5'', *H*-6''), 3.59-3.55 (dd, 1H, *J*= 9.8, *J*=3.7 Hz, *H*-2''), 3.47-3.42 (dd, 1H, *J*= 9.7 Hz, *J*= 8.2 Hz, *H*-4) ppm.

^{13}C-NMR (100 MHz; CD_3OD): δ= 159.9 (*C*-3), 159.6 (*C*-5), 158.5 (*C*-4'), 141.4 (*C*-1), 130.3 (*C*-8), 129.9 (*C*-1'), 128.9 (*C*-2'), 126.7 (*C*-7), 116.5 (*C*-3'), 108.3 (*C*-6), 107.4 (*C*-4), 104.5 (*C*-2), 99.3 (*C*-1''), 75.0 (C-5''), 74.3 (C-3''), 73.3 (*C*-2'') 71.5 (*C*-4''), 62.3 (*C*-6'') ppm.

MS (ESI positive):

Ion Formula:	$C_{20}H_{22}O_8Na^+$ $[M+Na]^+$
m/z calculated:	413.12069
m/z experimental:	413.12034
error [ppm]:	0.84

7.2.2. (+)-catechin-3'-O-α-D-glucopyranosid (26)



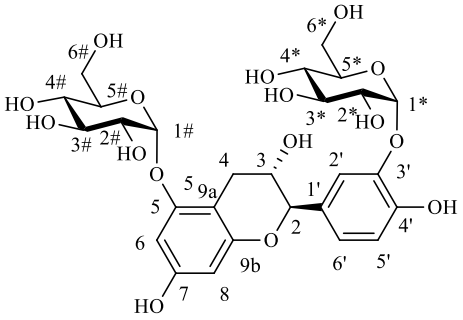
^{1}H-NMR (400 MHz; CD_3OD): δ= 7.32-7.31 (d, 1H, *J*= 2.0 Hz, *H*-2'),7.00-6.97 (d, 1H, *J*= 8.6 Hz, 2.0 Hz, *H*-6'), 6.86-6.84 (d, 1H, *J*= 8.3 Hz, *H*-5'), 5.93-5.84 (d, 1H, *J*= 2.4 Hz, *H*-6), 5.85-5.84 (d, 1H, *J*= 2.2 Hz, *H*-8),5.34-5.33 (d, 1H, *J*= 3.7 Hz, *H*-1#), 4.58-4.56 (d, 1H, *J*= 8.0 Hz, *H*-2), 4.02-3.96 (m, 1H, *H*-3), 3.90-3.85 (dd, 1H, *J*= 9.3 Hz, *H*-3*), 3.82 – 3.72 (m, 3H, *H*-5*, *H*-6*), 3.60-3.56 (dd, 1H, *J*= 9.7 Hz 3.7 Hz, *H*-2*), 3.47-3.43 (d, 1H, *J*= 9.1 Hz, *H*-4*), 2.92-2.87 (dd, 1H, *J*= 16.2 Hz 5.6 Hz, *H*-4a), 2.53-2.47 (dd, 1H, *J*= 16.2 Hz, *H*-4b) ppm.

^{13}C-NMR (100 MHz; CD_3OD): δ= 157.9 (C7), 157.6 (C5), 156.9 (C9a), 148.7 (C4') 146.3 (C3'), 132.3 (C1') 124.0 (C6'), 118.6 (C2'), 116.9 (C5') 101.6 (C1*), 100.9 (C9b), 96.3 (C6), 95.5 (C8), 82.8 (C2), 74.9 (C3*), 74.5 (C5*), 73.5 (C2*), 71.2 (C4*), 68.8 (C3), 62.2 (C6*), 29.0 (C4) ppm.

MS (ESI positive):

Ion Formula:	$C_{21}H_{24}O_{11}Na^+$ [M+Na]$^+$
m/z calculated:	475.12108
m/z experimental:	475.12102
error [ppm]:	0.14

7.2.3. (+)-catechin-3‘,5-O-α-D-diglucosid (27)



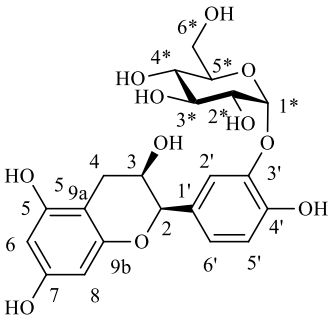
^{1}H-NMR (600 MHz; CD_3OD): δ= 7.30 (d, 1H, *J*= 2.0 Hz, *H*-2‘), 7.00-6.96 (dd, 1H, *J*= 8.3 Hz 2.2 Hz, *H*-6‘),6.86-6.84 (d, 1H, *J*= 8.1 Hz, *H*-5‘), 6.34-6.33 (d, 1H, *J*= 2.3 Hz, *H*-6), 6.01 (d, 1H, *J*= 2.3 Hz, *H*-8), 5.50 (d, 1H, *J*= 3.4 Hz, *H*-1#), 5.32 (d, 1H, *J*= 3.6 Hz, *H*-1*), 4.65-4-64 (d, 1H, *J*= 7.7 Hz, *H*-2), 4.05-4.01 (m, 1H, *H*-3), 3.89-3.87 (m, 2H, *H*-3*, *H*-3#), 3.80-3.70 (m, 5H, *H*-6*, *H*-6#, *H*-2*), 3.60-3.57 (m, 3H, *H*-4*, *H*-4#, *H*-2#), 3.47-3.43 (m, 2H, *H*-5*, *H*-5#), 2.92-2.89 (dd, 1H, *J*= 16.2 Hz 5.4 Hz, *H*-4a), 2.72-2.68 (dd, 1H, *J*= 16.4 Hz, 7.9 Hz, *H*-4b) ppm.

^{13}C-NMR (150 MHz; CD_3OD): δ= 158.1 (C7), 157.2 (C5), 156.6 (C9a), 148.7 (C4‘), 146.3 (C3‘), 132.2 (C1’) 123.8 (C6’) 118.4 (C2’), 116.9 (C5’), 103.2 (C9b), 101.6 (C1*), 98.7 (C1#), 97.9 (C8) 96.8 (C6) 82.6 (C2), 74.9 74.86 (C3*, C3#), 74.5, 74.5 (C2*, C2#), 73.4 73.3 (C4*, C4#), 71.3 71.2 (C5*, C5#), 68.5 (C3), 62.2 62.18 (C6*, C6#) 28.5 (C4) ppm.

MS (ESI positive):

Ion Formula:	$C_{27}H_{34}O_{16}Na^+$ [M+Na]$^+$
m/z calculated:	637.17391
m/z experimental:	637.17462
error [ppm]:	-1.12

7.2.4. (-)-epicatechin-3‘-O-α-D-glucopyranosid (28)



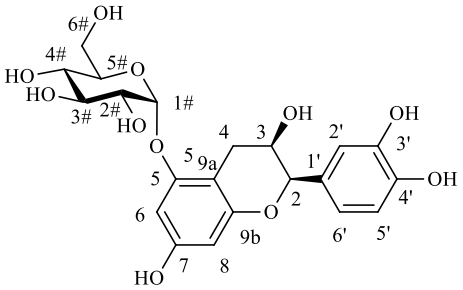
^{1}H-NMR (400 MHz; CD_3OD): δ= 7.43 (d, 1H, *J*= 2.2 Hz, *H*-2‘), 7.08-7.05 (dd, 1H, *J*= 8.2 Hz, 2.0 Hz, *H*-6‘), 6.85-6,83 (d, 1H, *J*= 8.2 Hz, *H*-5‘), 5.94-5.92 (m, 2H, *H*-6, *H*-8), 5.36-5.35 (d, 1H, *J*= 3.8 Hz, *H*-1*), 4.86 (s, 1H, *H*-2), 4.20-4.18 (m, 1H, *H*-3), 3.90-3.85 (dd, 1H, *J*= 9.2 Hz, *H*-3*), 3.84-3.80 (m, 1H, *H*-6*a), 3.80-3.76 (m, 1H, *H*-5*), 3.76-3.70 (m, 1H, *H*-6*b), 3.59-3.56 (dd, 1H, *J*= 9.7 Hz, 3.8 Hz, *H*-2*), 3.44-3.40 (dd, 1H, *J*= 9.6 Hz, 9.1 Hz, *H*-4*), 2.90-2.89 (dd, 1H, *J*= 16.8 Hz, 4.6 Hz, *H*-4a), 2.77-2.72 (dd, 1H, *J*= 16.9 Hz, 2.7 Hz, *H*-4b) ppm.

^{13}C-NMR (100 MHz; CD_3OD): δ= 158.0 157.7 157.3 (C9a, C7, C5), 148.1 (C4‘), 146.1 (C3‘), 132.5 (C1’), 123.2 (C6’), 118.1 (C2’), 116.6 (C5’), 101.4 C1*), 100.0 (C9b), 96.4 95.9 (C6, C8), 79.7 (C2), 74.9 (C3*), 74.5 (C5*), 73.5 (C2*), 71.4 (C4*), 67.4 (C3), 62.5 (C6*), 29.4 (C4) ppm.

MS (ESI positive):

Ion Formula:	$C_{27}H_{34}O_{16}Na^+$ $[M+Na]^+$
m/z calculated:	637.17391
m/z experimental:	637.17462
error [ppm]:	-1.12

7.2.5.(-)-epicatechin-5-O-α-D-glucopyranosid (29)



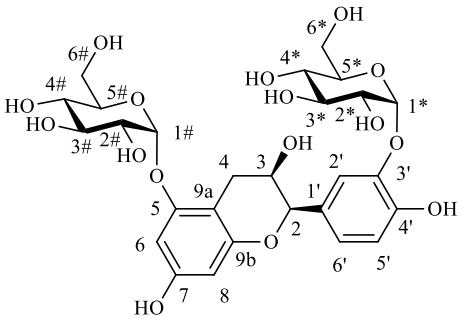
^{1}H-NMR (400 MHz; CD_3OD): δ= 6.97 (d, 1H, *J*= 1.92 Hz, *H*-2‘), 6.81-6.75 (m, 2H, *H*-5‘, *H*-6‘),6.34-6.33 (d, 1H, *J*= 2.32 Hz, *H*-6), 6.07 (d, 1H, *J*= 2.3 Hz, *H*-8), 5.51 (d, 1H, *J*= 3.6 Hz, *H*-1#), 4.83 (s, 1H, *H*-2), 4.19-4.17 (m, 1H, *H*-3), 3.91-3.86 (dd, 1H, *J*= 9.7 Hz 9.1Hz, *H*-3#), 3.72-3.71 (d, 2H, *J*= 3.5 Hz, *H*-6#), 3.63-3.58 (m, 2H, *H*-2#, *H*-5#), 3.50-3.45 (dd, 1H, *J*= 9.9 Hz, 8.8 Hz, *H*-4#), 3.06-3.00 (dd, 1H, *J*= 17.0 Hz 4.6 Hz, *H*-4a), 2.82-2.77 (,m, 1H, *J*= 17.1 Hz, 2.9 Hz, *H*-4b) ppm.

^{13}C-NMR (100 MHz; CD_3OD): δ= 157.8 (C7), 157.5 (C5), 157.1 (C9a), 146.0, 145.8 (C3’, C4‘), 132.1 (C1’), 119.4 (C6’), 115.9 (C5’), 115.3 (C2’), 102.5 (C9b), 98.4 (C1#), 98.2 (C8), 96.7 (C6), 80.0 (C2), 75.0 (C3#) 74.4 C2#) 73.4 (C5#), 71.3 (C4#), 67.3 (C3), 62.1 (C6#), 29.4 (C4) ppm.

MS (ESI positive):

Ion Formula:	$C_{21}H_{24}O_{11}Na^+$ $[M+Na]^+$
m/z calculated:	475.12108
m/z experimental:	475.12005
error [ppm]:	2.17

7.2.6. (-)-epicatechin-3‘,5-O-α-D-diglucosid (30)

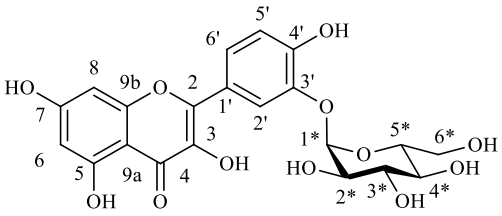


^{1}H-NMR (600 MHz; CD_3OD): δ= 7.43 (d, 1H, *J*= 1.92 Hz, *H*-2‘), 7.07-7.05 (dd, 1H, *J*= 8.31, 1.99 Hz, *H*-6‘), 6.86-6.84 (d, 1H, *J*= 8.24 Hz, *H*-5’) 6.34-6.33 (d, 1H, *J*= 2.28 Hz, *H*-6), 6.08 (d, 1H, *J*= 2.24 Hz, *H*-8), 5.51 (d, 1H, *J*= 3.6 Hz, *H*-1#), 5.35 (d, 1H, *J*= 3.64 Hz, *H*-1*), 4.88

(under Water) (1H, *H*-2), 4.20 (m, 1H, *H*-3), 3.91-3.71 (m, 7H, *H*-3#, *H*-3*, *H*-4*or#, 2*H*-6*, 2*H*-6#), 3.63-3.56 (m, 3H, *H*-2#, *H*-2*, *H*-4*or#,), 3.50-3.42 (m, 2H, *H*-5*, *H*-5#), 3.06-3.05 (dd, 1H, *J*= 17 Hz, 4.6 Hz *H*-4), 2.83-2.78 (dd, 1H, *J*= 17 Hz, 2.8 Hz *H*-4) ppm.

^{13}C-NMR (150 MHz; CD_3OD): δ= 157.9 (C7), 157.5 (C5), 157.1 (C9a), 148.2(C4‘), 146.1 (C3‘), 132.3 (C1’) 123.2 (C5’), 118.1 (C2’), 116.6 (C6’), 102.5 (C9b) 101.4 (C1*), 98.5 (C1#) 98.3 (C8), 96.8 (C6), 79.8 (C2), 75.0, 74.9, 74.4, 74.4 (C3#, C3*, C4#, C4*) 73.5, 73.4 (C2*, C2#) 71.5, 71.4, (C5*, C5#), 67.3 (C3), 62.5, 62.1 (C6*, C6#), 29.5 (C4) ppm.

7.2.7. Quercetin-3’-α-D-glucoside (36)



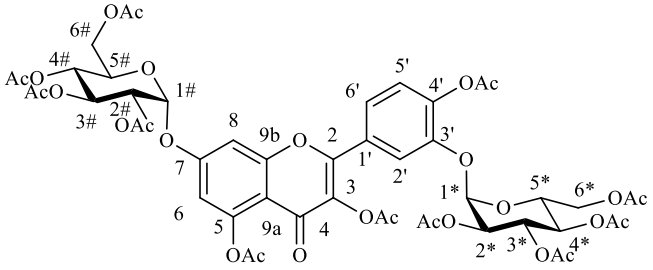
^{1}H-NMR (400 MHz; CD_3OD): δ= 8.16 (d, 1H, *J* = 2.1 Hz, *H*-2‘), 7.86-7.83 (dd, 1H, *J* = 8.5, 2.2 Hz, *H*-6‘), 6.98-6.96 (d, 1H, *J* = 8.7 Hz, *H*-5‘), 6.41 (d, 1H, *J* = 2.0 Hz, *H*-8), 6.18 (d, 1H, *J* = 2.1 Hz, *H*-6), 5.41-5.40 (d, 1H, *J* = 3.7 Hz, *H*-1*), 3.93-3.89 (dd, 1H, *J* = 9.0 Hz, *H*-3*), 3.88-3.79 (m 3H, *H*-5*, H*-6*), 3.65-3.61 (dd, 1H, *J* = 9.7, 3.7 Hz, *H*-2*), 3.54-3.49 (d, 1H, *J* = 9.0 Hz, *H*-4*) ppm.

^{13}C-NMR (100 MHz CD_3OD): δ= 177.4 (*C*-4), 165.6 (*C*-7), 162.5 (*C*-5), 158.2 (*C*-9b), 151.0 (*C*-4‘), 147.5 (*C*-2),146.3 (*C*-3‘), 137.4 (*C*-3), 125.3 (*C*-6‘), 124.3 (*C*-1‘), 119.1 (*C*-2‘), 117.2 (*C*-5‘), 104.5 (*C*-9a), 101.8 (C-1*), 99.3 (*C*-6), 94.6 (*C*-8), 74.9 (*C*-3*),74.6 (*C*-5*), 73.4 (*C*-2*), 71.1 (*C*-4*), 62.2 (*C*-6*) ppm.

MS (ESI positive):

Ion Formula:	$C_{21}H_{20}O_{12}Na^+$ $[M+Na]^+$
m/z calculated:	487.08470
m/z experimental:	487.08486
error [ppm]:	-0.34

7.2.8. Quercetin-3’,7-α-D-diglucoside-undecaacetate (39)

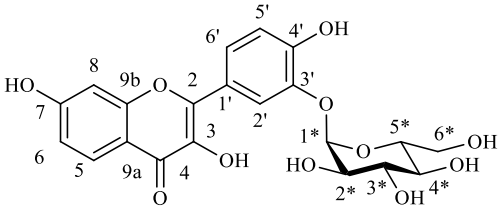


^{1}H-NMR (400 MHz; $CDCl_3$): δ= 7.73 (d, 1H, *J* = 2.0 Hz, *H*-2‘), 7.58 (dd, 1H, *J* = 8.4, 2.0 Hz, *H*-6‘), 7.22 (d, 1H, *J* = 8.4 Hz, *H*-5‘), 7.18 (d, 1H, *J* = 2.4 Hz, *H*-6), 6.84 (d, 1H, *J* = 2.4 Hz, *H*-8), 5.89 (d, 1H, *J* = 3.6 Hz, *H*-1#), 5.81 (d, 1H, *J* = 3.5 Hz, *H*-1*), 5.69-5.62 (m, 2H, *H*-3#, *H*-3*), 5.22-5.16 (m, 2H, *H*-4#, *H*-4*), 5.11-5.08 (dd, 1H, *J* = 10.3, 3.6 Hz, *H*-2#), 5.02 (dd, 1H, *J* = 10.4, 3.5 Hz, *H*-2*), 4.31-4.21 (m, 2H, *H*-6#, *H*-6*), 4.12-4.02 (m, 4H, *H*-5#, *H*-6#, *H*-5*, H*-6*), 2.43 (s, 3H, $COCH_3$), 2.43 (s, 3H, $COCH_3$), 2.33 (s, 3H, $COCH_3$), 2.09 (s, 3H, $COCH_3$), 2.08 (s, 3H, $COCH_3$), 2.06 (s, 3H, $COCH_3$), 2.055 (s, 3H, $COCH_3$), 2.05 (s, 3H, $COCH_3$), 2.04 (2s, 6H, $COCH_3$), 2.00 (s, 3H, $COCH_3$);

^{13}C-NMR (100 MHz; $CDCl_3$): δ =170.7, 170.6 170.3 170.2 170.2 170.2 170.0, 169.6, 169.6 169.6 168.8 167.9 (H_3CO, *C*-4), 159.9 (*C*-7), 157.8 (*C*-5), 153.6 (*C*-2), 151.1 (*C*-9b), 147.8 (*C*-3') 142.8 (*C*-4') 134.1 (*C*-3) 128.5 (*C*-1'), 123.5 123.4 (*C*-5' und *C*-6') 115.0 (*C*-2'), 112.7 (*C*-9a), 109.8 (*C*-8), 102.4 (*C*-6), 94.8 (*C*-1*) 94.4 (*C*-1#), 70.5 (*C*-2*), 70.0 (*C*-2#), 69.8, 69.5 (*C*-3*, *C*-3#), 69.0, 68,9 (*C*-5*, *C*-5#); 68.0, 67.7 (*C*-4*, *C*-4#), 61.4, 61.3 (*C*-6*, *C*-6#), 21.2, 20.8, 20.8 20.8, 20.7, 20.7, 20,7, 20,7 20,7, 20.7, 20.5 (CH_3CO) ppm.

MS (ESI positive, of the non-acetylated compound):
Ion Formula: $C_{27}H_{30}O_{17}Na^+$ $[M+Na]^+$
m/z calculated: 649.13752
m/z experimental: 649.13675
error [ppm]: 1.19

7.2.9. Fisetin-3'-α-D-glucoside (40)

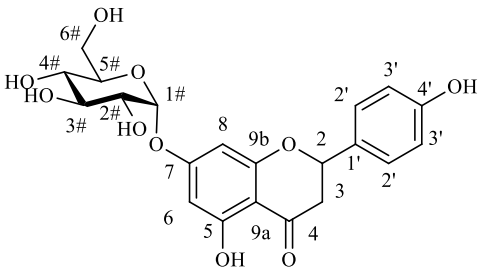


^{1}H-NMR (400 MHz; CD_3OD): δ= 8.21-8.20 (d, 1H, *J* = 2.1 Hz, *H*-2'), 8.00-7.97 (d, 1H, *J* = 8.5 Hz, *H*-5), 7.90-7.87 (dd, 1H, *J* = 8.6, 2.1 Hz, *H*-6'), 7.00-6.98 (d, 1H, *J* = 8.6 Hz, *H*-5'), 6.95-6.94 (d, 1H, *J* = 2.1 Hz, *H*-8) 6.93-6.90 (dd, 1H, *J* = 8.7, 2.1 Hz, *H*-6), 5.41-5.40 (d, 1H, *J* = 3.7 Hz, *H*-1*), 3.94-3.89 (dd, 1H, *J*= 9.3, 9.3 Hz, *H*-3*), 3.88-3.79 (m, 3H, *H*-5*, *H*-6*), 3.65-3.62 (dd, 1H, *J* = 9.5, 3.8 Hz, *H*-2*),3.53-3.49 (dd, 1H, *J* = 9.0, 9.0 Hz, *H*-4*) ppm.

^{13}C-NMR (100 MHz CD_3OD): δ= 174.4 (*C*-4), 164.3 (*C*-7), 158.6 (*C*-9b), 150.9 (*C*-4'), 147.0 (*C*-2), 146.3 (*C*-3'), 138.7 (*C*-3), 127.5 (*C*-5), 125.3 (*C*-6'), 124.5 (*C*-1') 119.2 (*C*-2'), 117.2 (*C*-5'), 116.0 (*C*-6), 115.5 (*C*-9a) 103.1 (*C*-8), 101.9 (*C*-1*), 74.9 (*C*-3*), 74.6 (*C*-5*), 73.5 (*C*-2*), 71.2 (*C*-4*), 62.2 (*C*-6*) ppm.

MS (ESI positive):
Ion Formula: $C_{21}H_{20}O_{11}Na^+$ $[M+Na]^+$
m/z calculated: 471.08978
m/z experimental: 471.08839
error [ppm]: 2.95

7.2.10. Naringenin-7-α-D-glucoside (41)



1:1: mixture of diastereomers

^{1}H-NMR (400 MHz; CD_3OD): δ= 7.33-7.31 (m, 2H, *H*-2'), 6.83-6.81 (m, 2H, *J* = 2.0 Hz, *H*-3'),6.30-6.27 (m, 2H, *H*-6, *H*-8), 5.59-5.56 (d, 1H, *J* = 3.5 Hz, *H*-1#), 5.40-5.36 (dd, 1H, *J* = 13.2, 2.9 Hz, *H*-2), 3.83-3.78 (dd, 1H, *J* = 9.3, 9.2 Hz, *H*-3#), 3.75-3.66 (m, 2H, *H*-6#), 3.58-

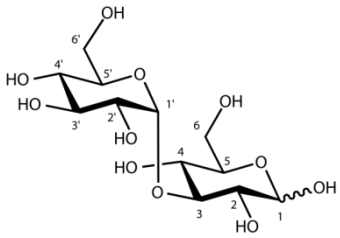
3.55 (dd, 1H, *J* = 9.7, 3.8 Hz, *H*-2#), 3.55-3.51 (m, 1H, *H*-5#) 3.45-3.40 (dd, 1H, *J* = 9.4, 9.5 Hz, *H*-4#) 3.20-3.13 (dd, 1H, *J*= 17.2, 13.3 Hz, *H*-3) 2.77-2.72 (m, 1H, *H*-3) ppm.

^{13}C-NMR (100 MHz; CD_3OD): δ= 198.6 (*C*-4), 166.6, 165.0 164.6 (*C*-5, *C*-7, *C*-9b), 159.1 (*C*-4‘), 131.8 (*C*-1‘), 129.1 (*C*-2‘), 116.3 (*C*-3‘), 104.9 (*C*-9a), 98.6 (*C*-1#), 98.1, 97.1 (*C*-6, *C*-8), 80.6 (*C*-2) 74.9 (*C*-5#), 74.7 (*C*-3#), 73.0 (*C*-2#), 71.1 (*C*-4#), 62.1 (*C*-6#), 44.1 (*C*-3) ppm.

MS (ESI positive):

Ion Formula:	$C_{21}H_{22}O_{10}Na^+$ $[M+Na]^+$
m/z calculated:	457.11052
m/z experimental:	457.11057
error [ppm]:	-0.11

7.2.11. Nigerose (23)



α-anomer (40%):

^{1}H NMR (600 MHz, D_2O): δ 5.37 (d, *J* = 3.9 Hz, 1H, H-1'), 5.23 (d, *J* = 3.8 Hz, 1H, H-1), 4.02 (m, 1H, H-5'), 3.87-3.81 (m, 4H, H-3, H-5, H-6a, H-6'a), 3.80-3.73 (m, 2H, H-6'b, H-6b), 3.75 (dd, *J* = 9.5, 9.5 Hz, 1H, H-3'), 3.64 (dd, *J* = 9.5, 7.0 Hz, 1H, H-4), 3.62 (dd, *J* = 9.8, 3.9 Hz, 1H, H-2), 3.57 (dd, *J* = 9.9, 5.1 Hz, 1H, H-2'), 3.46 (dd, *J* = 9.2, 10.2 Hz, 1H, H-4') ppm.

^{13}C NMR (150 MHz, D_2O): δ 100.25 (C-1'), 93.43 (C-1), 80.73 (C-3), 74.05 (C-3'), 72.93 (C-5'), 72.89 (C-2'), 72.38 (C-5), 71.29, 71.22 (C-2, C-4), 70.60 (C-4'), 61.56, 61.53 (C-6, C-6') ppm.

ß-anomer (60%):

^{1}H NMR (600 MHz, D_2O): δ 5.36 (d, *J* = 3.9 Hz, 1H, H-1'), 4.66 (d, *J* = 8.0 Hz, 1H, H-1), 4.02 (m, 1H, H-5'), 3.89 (dd, *J* = 2.2, 12.3 Hz, 1H, H-6a), 3.84 (dd, *J* = 12.7, 2.2 Hz, 1H, H-6'a), 3.78 (ddd, *J* = 12.7, 8.4, 4.0 Hz, 1H, H-6'b), 3.74 (dd, *J* = 9.6, 9.6 Hz, 1H, H-3'), 3.72 (dd, *J* = 12.4, 5.9 Hz, 1H, H-6b), 3.64 (dd, *J* = 9.5, 7.0 Hz, 1H, H-3), 3.63 (m, 1H, H-4), 3.56 (dd, *J* = 9.8, 4.1 Hz, 1H, H-2'), 3.47 (ddd, 9.7, 5.4, 2.8, 1H, H-5), 3.44 (dd, *J* = 10.6, 9.5 Hz, 1H, H-4'), 3.33 (dd, *J* = 9.3, 8.0 Hz, 1H, H-2) ppm.

^{13}C NMR (150 MHz, D_2O): δ 100.21 (C-1'), 97.16 (C-1), 83.25 (C-3), 76.84 (C-5), 74.08 (C-3'), 74.03 (C-2), 72.90 (C-5'), 72.82 (C-2'), 71.25 (C-4), 70.47 (C-4'), 61.74 (C-6), 61.37 (C-6') ppm.

7.3. A little tool to perform the direct linear plot analysis

Regrettably we did not come up with a smart sounding acronym.

Notes: The programm is written in Python 3.4.

The input file is a .txt file. Each line contains a pair of values separated by at least one space: The first value is the starting concentration, the second value the observed initial reaction velocity. The values need to contain a comma: eg.: 0,23 or 10,00. The tool is started by entering the directory of the input file, and generates an output file in the same directory with the median K_M and v_{max} value at the end.

```
import os
from shutil import copyfile
in_dir1 = os.path.normpath(input("Enter input file directory. column A [S0], column B [Vini]"))
# Make sure your numbers all contain a comma!! Else it crashes during float conversion
f_in1 = open(in_dir1, "r")
rename1 = in_dir1.split(".")
f_out1 = open(rename1[0]+"_out."+rename1[1], "a")

f_out1.write("[S0]          [V0]" + "\n")
ln = 1

raw = []
num = []

parameters =40
for line in f_in1:
        f_out1.write(line)
        raw = line.split()
        a = 0
        print (raw)

        for thing in raw:
                num = thing.split(",")
                if a == 0:
                        s_out = float(num[0]+"."+num[1])
                        a = 1
                else:
                        v_out = float(num[0]+"."+num[1])
        parameter = [ln, s_out, v_out]
        #print (parameter)
        parameters.update({ln : parameter})
        ln = ln + 1
print (parameters)
para_count = 0
for parameter_set in parameters:
        para_count = para_count +1
print (para_count)

count = 0
base = 0
s = 1
v = 2
list_of_all_KM = []
```

```
list_of_all_Vmax = []
while base < para_count:
	base = base + 1
	count = 0
	while count + base < para_count:
		count = count + 1
		set_a = parameters[base]
		bc = base + count
		#print (bc)
		set_b = parameters[bc]
		#print (set_a)
		#print (set_b)
		KM = (set_b[v]-set_a[v])/( (set_a[v]/set_a[s]) - (set_b[v]/set_b[s]) )
		Vmax = KM*(set_a[v]/set_a[s]) + set_a[v]
		print (str(base)+","+str(bc))
		print (KM)
		print (Vmax)
		f_out1.write(str(base)+","+str(bc)+"\n"+"KM:  "+str(KM)+"\n"+"Vmax:  "+str(Vmax)+"\n")
		list_of_all_KM.append(KM)
		list_of_all_Vmax.append(Vmax)
print ("KM-values:")
print (list_of_all_KM)
f_out1.write("KM-values:" + "\n")
for value in list_of_all_KM:
	f_out1.write(str(value)+"   ")
#f_out1.write(list_of_all_KM)
f_out1.write("  "+ "\n")
print ("Vmax-values:")
print (list_of_all_Vmax)
f_out1.write("Vmax-values:" +"\n")
for value in list_of_all_Vmax:
	f_out1.write(str(value)+"   ")

f_out1.write("  "+ "\n")
list_of_all_KM.sort()
print ("KM-values sorted:")
print (list_of_all_KM)
f_out1.write("KM-values sorted:" + "\n")
for value in list_of_all_KM:
	f_out1.write(str(value)+"   ")
f_out1.write("  "+ "\n")
list_of_all_Vmax.sort()
print ("Vmax-values sorted:")
print (list_of_all_Vmax)
f_out1.write("Vmax-values sorted:" +"\n")
for value in list_of_all_Vmax:
	f_out1.write(str(value)+"   ")

f_out1.write("  "+ "\n")

print (len(list_of_all_KM))
f_out1.write("  "+ "\n")

if len(list_of_all_KM)%2==0:
	print ("even")
	n = (len(list_of_all_KM))/2
	med_km = (list_of_all_KM[int(n)]+list_of_all_KM[int(n-1)])/2
	print ("Median KM:")
	print (med_km)
	f_out1.writelines("Median KM:"+"\n")
```

```
        f_out1.writelines(str(med_km)+"\n")

else :
        print ("odd")
        n = ((len(list_of_all_KM))+1)/2 - 1
        n = int(n)
        med_km = (list_of_all_KM[n])
        print ("Median KM:")
        print (med_km)
        f_out1.writelines("Median KM:"+"\n")

        f_out1.writelines(str(med_km)+"\n")

f_out1.write("  "+ "\n")

if len(list_of_all_Vmax)%2==0:
        print ("even")
        n = len(list_of_all_Vmax)/2
        med_Vmax = (list_of_all_Vmax[int(n)]+list_of_all_Vmax[int(n-1)])/2
        print ("Median Vmax:")
        print (med_Vmax)
        f_out1.writelines("Median Vmax:"+"\n")

        f_out1.writelines(str(med_Vmax)+"\n")

else :
        print ("odd")
        n = ((len(list_of_all_Vmax))+1)/2-1
        n = int(n)
        med_Vmax = (list_of_all_Vmax[n])
        print ("Median Vmax:")
        print (med_Vmax)
        f_out1.writelines("Median Vmax:"+"\n")

        f_out1.writelines(str(med_Vmax)+"\n")

f_out1.write("  "+ "\n")

f_in1.close()
f_out1.close()
```

7.4. Product profiles of BaSP Q345F with aromatic substrates

7.4.1. Product profile of BaSP Q345F with epicatechin as acceptor

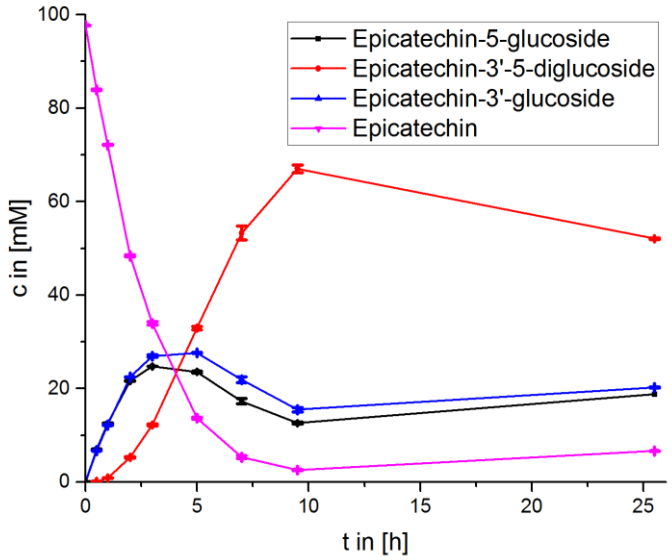
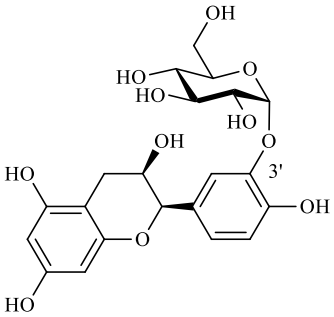
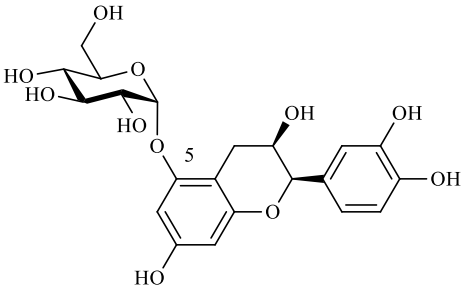


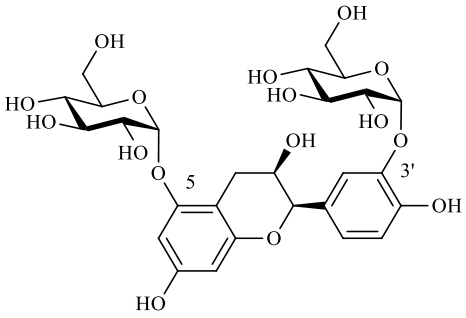
Figure 38 Product profile of BaSP Q345F with epicatechin as acceptor



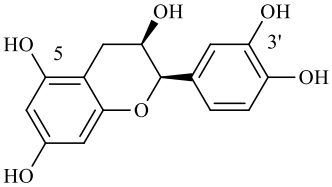
(-)-epicatechin-3'-glucosid **28**



(-)-epicatechin-5-glucosid **29**



(-)-epicatechin-3'-5-diglucosid **30**



(-)-epicatechin **32**

7.4.2. Product profile of BaSP Q345F with catechin as acceptor

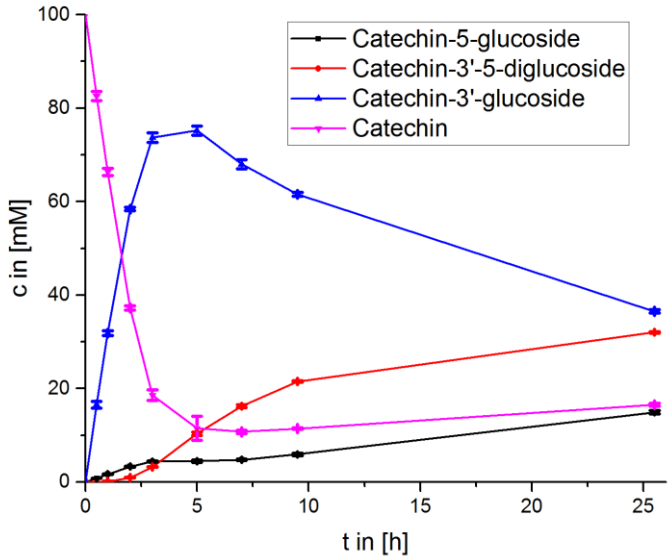
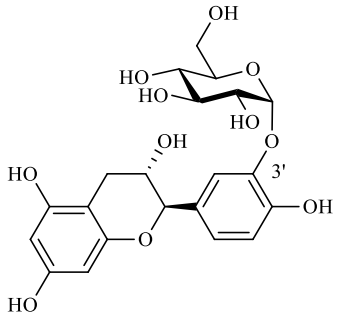
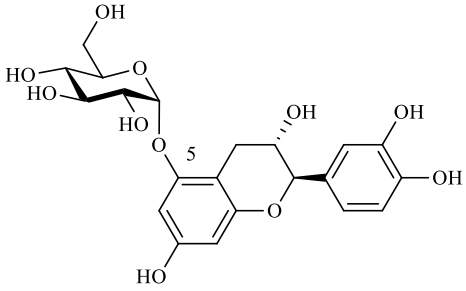


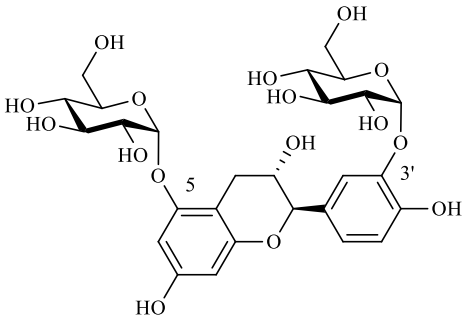
Figure 39 Product profile of BaSP Q345F with catechin as acceptor



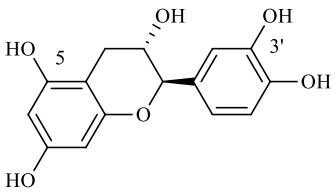
(+)-catechin-3'-glucosid **26**



(+)-catechin-5-glucosid **29**



(+)-catechin-3'-5-diglucosid **27**



(+)-catechin **17**

An peak of an uncharacterized product was detected and assigned to (+)-catechin-5-glucosid as it displays similar retention times as (-)-epicatechin-5-glucosid and the Q345F variant is known to produce (+)-catechin-3'-5-diglucosid.

7.4.3. Product profile of BaSP Q345F with quercetin as acceptor.

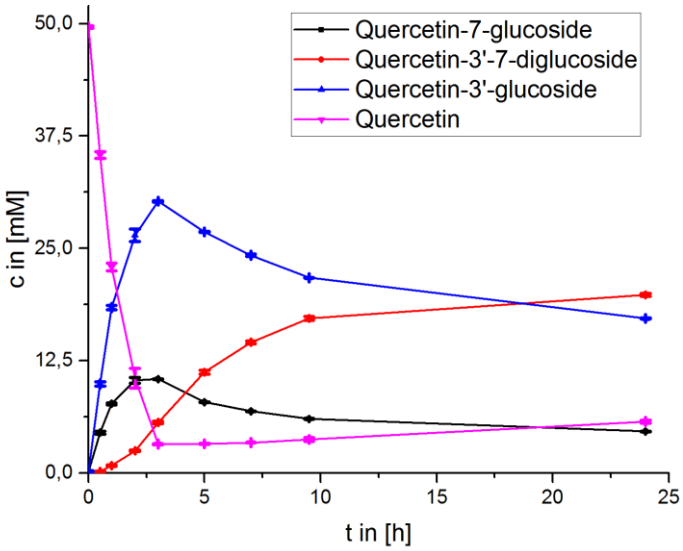
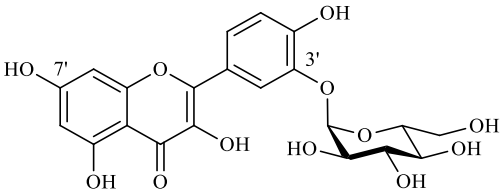
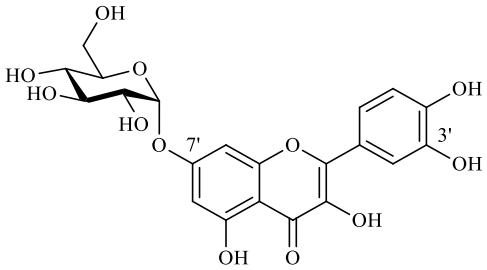


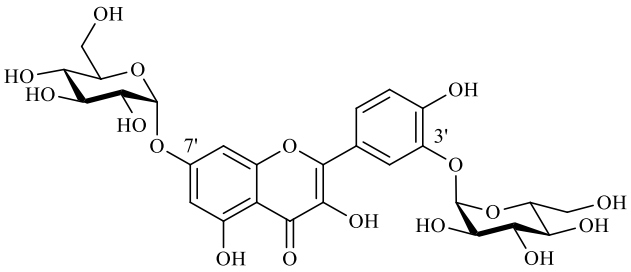
Figure 40 Product profile of BaSP Q345F with quercetin as acceptor



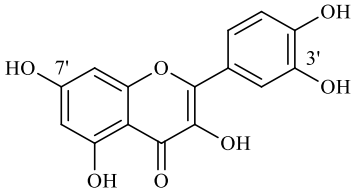
quercetin-3'-glucosid **36**



quercetin-7-glucosid **37**



quercetin-3'-7-diglucosid **38**



Quercetin **3**

Quercetin-3'-glucosid and quercetin-3'-7-diglucosid were isolated, the third component was not obtained in sufficient purity to allow NMR-characterisation. Its chromatographic behaviour corresponds to a monoglucoside and as the diglucoside is glucosylated in 3'- and 7-position we assumed the third product to be quercetin-7-glucosid.

7.4.4. Product profile of BaSP Q345F with fisetin as acceptor

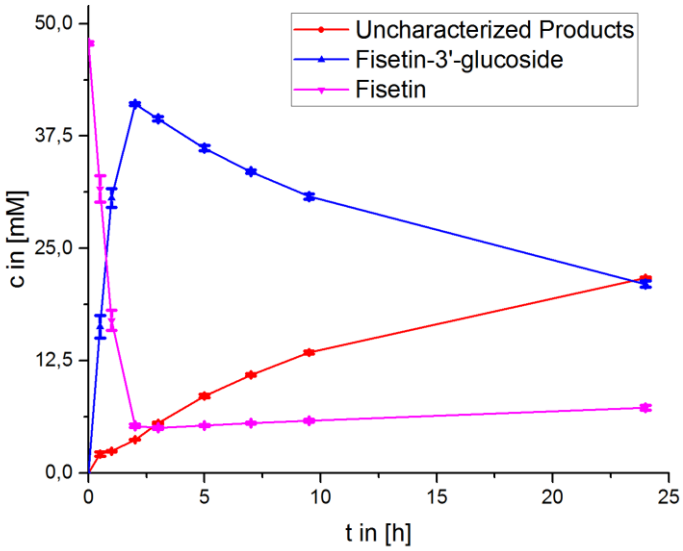
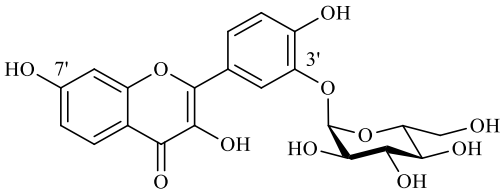
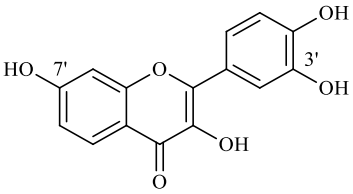


Figure 41 Product profile of BaSP Q345F with fisetin as acceptor



fisetin-3'-glucosid **40**



fisetin **33**

Two peaks of uncharacterized products were detected. Most likely this compounds are fisetin-7-glucoside and fisetin 3'-7-diglucoside similar to the reaction products found in quercetin.

7.4.5. Product profile of BaSP Q345F with naringenin as acceptor

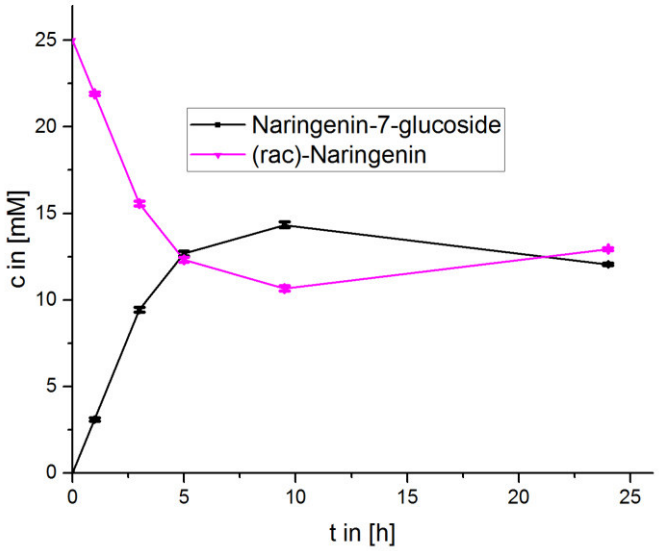
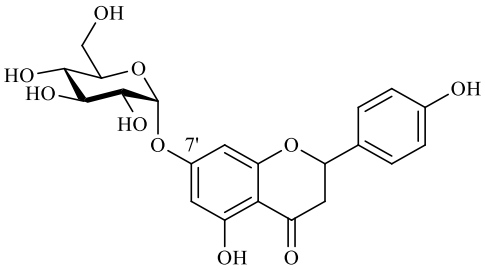
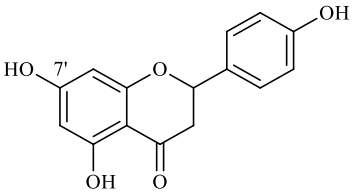


Figure 42 Product profile of BaSP Q345F with naringenin as acceptor



naringennin-7-glucosid **41**



naringenin **34**

7.4.6. Product profile of BaSP Q345F with resveratrol as acceptor

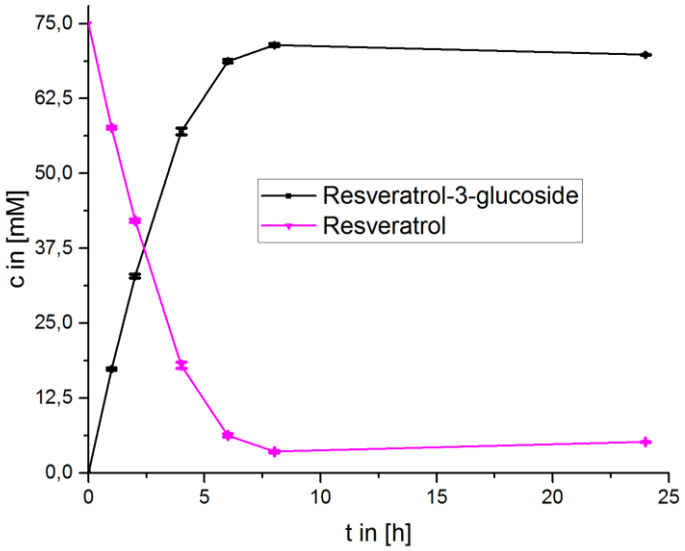
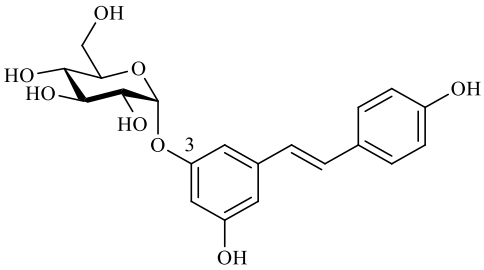
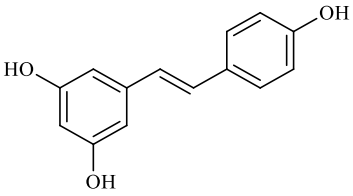


Figure 43 Product profile of BaSP Q345F with resveratrol as acceptor



resveratrol-3-glucosid **31**



resveratrol **2**

7.5. Crystal structure data collection and refinement statistics

Table 12 Data collection and refinement statistics Crystall structures discussed in chapter 2 *Highest resolution shell is shown in parenthesis. The high resolution cutoff and useful resolution range was determined according to the method described in [142].

	BaSP Q345F (PDB ID 5c8b)
Data collection	
Space group	*P 43 21 2*
Cell dimensions	
a, b, c (Å)	*82.5 82.5 156.4*
(°)	*90, 90, 90*
Resolution (Å)	*46.76 - 2.681 (2.776 - 2.681)**
Rmerge	*0.093 (1.546)*
I/σ(I)	*16.7 (1.6)*
Completeness (%)	*99.2 (92.8)*
Redundancy	*12.4 (12.7)*
Refinement	
Resolution (Å)	*46.76 - 2.681 (2.776 - 2.681)**
No. reflections (unique)	*15705 (1544)*
R_{work}/ R_{free}	
No. atoms, total	*3967*
Protein residues	*504*
Ligand atoms	*12*
Water molecules	*142*
B-factors, average (Å^2)	*95.5*
Protein	*96.4*
Ligand/ion	*79.4*
Water	*71.8*
R.m.s deviations	
Bond lengths (Å)	*0.016*
Bond angles (°)	*1.6*
Ramachandran	
favored (%)	*97*
outliers	*0*

Table 13 Data collection and refinement statistics Statistics Crystall structures discussed in chapter 4. The highest-resolution shell are shown in parentheses.

	Resveratrol	Nigerose	Apo
PDB ID	*5M9X*	*5MAN*	*5MB2*
Wavelength [Å]	*0.97625*	*0.97625*	*0.97625*
Resolution range	*47.06 - 2.349*	*43.75 - 2.04*	*43.57 - 1.752*
Space group	*P 43 21 2*	*P 43 21 2*	*P 43 21 2*
Unit cell (a, b, c, α, β, γ) [Å, Å, Å, °, °, °]	*83.1, 83.1, 157.2, 90.0. 90.0. 90.0*	*82.4, 82.4, 154.9, 90.0. 90.0. 90.0*	*82.9, 82.9, 153.6, 90.0. 90.0. 90.0*
Total reflections	*157373 (14977)*	*277167 (26255)*	*692483 (67202)*
Unique reflections	*23334 (2273)*	*34768 (3400)*	*54566 (5326)*
Multiplicity	*6.7 (6.6)*	*8.0 (7.7)*	*12.7 (12.6)*
Completeness [%]	*98.0 (97.0)*	*99.9 (99.9)*	*99.9 (99.0)*
Mean I/sigma(I)	*20.88 (0.96)*	*23.69 (1.60)*	*24.36 (1.69)*
Wilson B-factor	*67.88*	*48.91*	*30.56*
R-merge	*0.05327 (1.945)*	*0.04394 (1.366)*	*0.06463 (1.619)*
Reflections used in	*23328 (2273)*	*34766 (3400)*	*54556 (5323)*
R-work	*0.2575 (0.4161)*	*0.2349 (0.3726)*	*0.1974 (0.3375)*
R-free	*0.2847 (0.4017)*	*0.2616 (0.3764)*	*0.2323 (0.3679)*
Number of non-	*4038*	*4103*	*4309*
macromolecules	*3975*	*3966*	*3983*
Ligands	*28*	*23*	*6*
Solvent	*35*	*114*	*320*
Protein residues	*504*	*504*	*504*
RMS(bonds) [Å]	*0.003*	*0.003*	*0.004*
RMS(angles) [°]	*0.61*	*0.56*	*0.63*
Ramachandran			
favoured [%]	*97.0*	*96.2*	*97.4*
allowed [%]	*2.6*	*3.6*	*2*
outliers [%]	*0.4*	*0.2*	*0.6*
Clashscore	*2.2*	*1.2*	*4.0*
Average B-factor	*99.8*	*60.4*	*49.4*
Macromolecules	*100.1*	*60.6*	*49.8*
Ligands	*80.4*	*75.1*	*60.7*
Solvent	*88.6*	*49.9*	*44.5*

Table 14 Data collection and refinment statistics of the Crystall structures discussed in chapter 5

PDB ID: 6fme

Wavelength [Å]	0.97625
Resolution range	*42.33 - 1.51*
Space group	*P 21 21 21*
Unit cell (a, b, c, α, β, γ)[Å, Å, Å, °, °, °]	*76.01 101.925 152.569 90 90 90*
Total reflections	*665032 (66346)*
Unique reflections	*175921 (17788)*
Multiplicity	*3.8 (3.7)*
Completeness [%]	*95.0 (97.0)*
Mean I/sigma(I)	*10.32 (0.89)*
Wilson B-factor	*17.78*
R-merge	*0.08066 (1.239)*
Reflections used in refinement	*175552 (17783)*
R-work	*0.1670 (0.3180)*
R-free	*0.1879 (0.3347)*
Number of non-hydrogen atoms	*9200*
macromolecules	*8002*
Ligands	*42*
Solvent	*1156*
Protein residues	*1011*
RMS(bonds) [Å]	*0.008*
RMS(angles) [°]	*1.07*
Ramachandran favoured [%]	*98*
allowed [%]	*1.9*
outliers [%]	*0.099*
Clashscore	*4.49*
Average B-factor [$Å^2$]	*23.45*
Macromolecules	*21.68*
Ligands	*14.53*
Solvent	*36.02*

Statistics for the highest-resolution shell are shown in parentheses.

Chapter 8

LITERATURE

1. Movie: Mary Poppins, Director: R. Stevenson, W. D. Productions, 1964
2. D. R. Harris, S. G. McGeachin and H. H. Mills, *Tetrahedron Letters*, 1965, **6**, 679-685.
3. R. B. Woodward, E. Logusch, K. P. Nambiar, K. Sakan, D. E. Ward, B. W. Au-Yeung, P. Balaram, L. J. Browne, P. J. Card and C. H. Chen, *Journal of the American Chemical Society*, 1981, **103**, 3210-3213.
4. R. B. Woodward, B. W. Au-Yeung, P. Balaram, L. J. Browne, D. E. Ward, B. W. Au-Yeung, P. Balaram, L. J. Browne, P. J. Card and C. H. Chen, *Journal of the American Chemical Society*, 1981, **103**, 3213-3215.
5. R. B. Woodward, E. Logusch, K. P. Nambiar, K. Sakan, D. E. Ward, B. W. Au-Yeung, P. Balaram, L. J. Browne, P. J. Card and C. H. Chen, *Journal of the American Chemical Society*, 1981, **103**, 3215-3217.
6. K. C. Nicolaou, H. J. Mitchell, N. F. Jain, N. Winssinger, R. Hughes and T. Bando, *Angewandte Chemie International Edition*, 1999, **38**, 240-244.
7. J. D. Williams, C. A. Waltho, G. A. J. Ayliffe and E. J. L. Lowbury, *The Lancet*, 1967, **290**, 390-392.
8. G. Bonadonna, R. Zucali, S. Monfardini, M. de Lena and C. Uslenghi, *Cancer*, 1975, **36**, 252-259.
9. J. Golik, J. Clardy, G. Dubay, G. Groenewold, H. Kawaguchi, M. Konishi, B. Krishnan, H. Ohkuma, K. Saitoh and T. W. Doyle, *Journal of the American Chemical Society*, 1987, **109**, 3461-3462.
10. M. D. Lee, T. S. Dunne, M. M. Siegel, C. C. Chang, G. O. Morton and D. B. Borders, *Journal of the American Chemical Society*, 1987, **109**, 3464-3466.
11. D.-Z. Liu, S. Sinchaikul, P. V. G. Reddy, M.-Y. Chang and S.-T. Chen, *Bioorganic & Medicinal Chemistry Letters*, 2007, **17**, 617-620.
12. M. Jang, L. Cai, G. O. Udeani, K. V. Slowing, C. F. Thomas, C. W. Beecher, H. H. Fong, N. R. Farnsworth, A. D. Kinghorn, R. G. Mehta, R. C. Moon and J. M. Pezzuto, *Science*, 1997, **275**, 218-220.
13. S. Srivastava, R. R. Somasagara, M. Hegde, M. Nishana, S. K. Tadi, M. Srivastava, B. Choudhary and S. C. Raghavan, *Scientific Reports*, 2016, **6**, 24049.
14. M. Sajish and P. Schimmel, *Nature*, 2015, **519**, 370-373.
15. J. A. Baur, K. J. Pearson, N. L. Price, H. A. Jamieson, C. Lerin, A. Kalra, V. V. Prabhu, J. S. Allard, G. Lopez-Lluch, K. Lewis, P. J. Pistell, S. Poosala, K. G. Becker, O. Boss, D. Gwinn, M. Wang, S. Ramaswamy, K. W. Fishbein, R. G. Spencer, E. G. Lakatta, D. Le Couteur, R. J. Shaw, P. Navas, P. Puigserver, D. K. Ingram, R. de Cabo and D. A. Sinclair, *Nature*, 2006, **444**, 337-342.
16. K. T. Howitz, K. J. Bitterman, H. Y. Cohen, D. W. Lamming, S. Lavu, J. G. Wood, R. E. Zipkin, P. Chung, A. Kisielewski, L.-L. Zhang, B. Scherer and D. A. Sinclair, *Nature*, 2003, **425**, 191-196.
17. P. C. H. Hollman, M. N. C. P. Bijsman, Y. van Gameren, E. P. J. Cnossen, J. H. M. de Vries and M. B. Katan, *Free Radical Research*, 1999, **31**, 569-573.
18. L.-L. Yang, N. Xiao, X.-W. Li, Y. Fan, R. N. Alolga, X.-Y. Sun, S.-L. Wang, P. Li and L.-W. Qi, *Scientific Reports*, 2016, **6**, 35460.
19. H. J. Choi, J. H. Song, K. S. Park and D. H. Kwon, *European Journal of Pharmaceutical Sciences*, 2009, **37**, 329-333.
20. L. Sánchez-del-Campo, M. Sáez-Ayala, S. Chazarra, J. Cabezas-Herrera and J. N. Rodríguez-López, *International Journal of Molecular Sciences*, 2009, **10**, 5398-5410.

21. E. Fischer, *Berichte der deutschen chemischen Gesellschaft*, 1893, **26**, 2400-2412.
22. E. Fischer and L. Beensch, *Berichte der deutschen chemischen Gesellschaft*, 1894, **27**, 2478-2486.
23. M. Nagasaki, Y. Manabe, N. Minamoto, K. Tanaka, A. Silipo, A. Molinaro and K. Fukase, *The Journal of Organic Chemistry*, 2016, **81**, 10600-10616.
24. H. Paulsen, *Angewandte Chemie International Edition in English*, 1982, **21**, 155-173.
25. C.-C. Wang, J.-C. Lee, S.-Y. Luo, S. S. Kulkarni, Y.-W. Huang, C.-C. Lee, K.-L. Chang and S.-C. Hung, *Nature*, 2007, **446**, 896.
26. C. E. Cardini, L. F. Leloir and J. Chiriboga, *Journal of Biological Chemistry*, 1955, **214**, 149-155.
27. T. D. Hurley, S. Stout, E. Miner, J. Zhou and P. J. Roach, *Journal of Biological Chemistry*, 2005, **280**, 23892-23899.
28. R. Wild, J. Kowal, J. Eyring, E. M. Ngwa, M. Aebi and K. P. Locher, *Science*, 2018, **359**, 545-550.
29. S. Pfeffer, J. Dudek, M. Gogala, S. Schorr, J. Linxweiler, S. Lang, T. Becker, R. Beckmann, R. Zimmermann and F. Förster, *Nature Communications*, 2014, **5**, 3072.
30. T. Kubota, T. Shiba, S. Sugioka, S. Furukawa, H. Sawaki, R. Kato, S. Wakatsuki and H. Narimatsu, *Journal of Molecular Biology*, 2006, **359**, 708-727.
31. G. J. Williams, C. Zhang and J. S. Thorson, *Nature Chemical Biology*, 2007, **3**, 657.
32. K. D. Miller, V. Guyon, J. N. S. Evans, W. A. Shuttleworth and L. P. Taylor, *Journal of Biological Chemistry*, 1999, **274**, 34011-34019.
33. L. L. Lairson, B. Henrissat, G. J. Davies and S. G. Withers, *Annual Review of Biochemistry*, 2008, **77**, 521-555.
34. W. Z. Hassid, M. Doudoroff and H. A. Barker, *Journal of the American Chemical Society*, 1944, **66**, 1416-1419.
35. M. Doudoroff, N. Kaplan and W. Z. Hassid, *Journal of Biological Chemistry*, 1943, **148**, 67-75.
36. Y. Tsumuraya, C. F. Brewer and E. J. Hehre, *Archives of Biochemistry and Biophysics*, 1990, **281**, 58-65.
37. L. K. Skov, O. Mirza, A. Henriksen, G. P. De Montalk, M. Remaud-Simeon, P. Sarçabal, R.-M. Willemot, P. Monsan and M. Gajhede, *Journal of Biological Chemistry*, 2001, **276**, 25273-25278.
38. M. E. Ortiz-Soto, C. Possiel, J. Görl, A. Vogel, R. Schmiedel and J. Seibel, *Glycobiology*, 2017, **27**, 755-765.
39. B. Bissaro, P. Monsan, R. Fauré and Michael J. O'Donohue, *Biochemical Journal*, 2015, **467**, 17-35.
40. D. Aerts, T. F. Verhaeghe, B. I. Roman, C. V. Stevens, T. Desmet and W. Soetaert, *Carbohydrate Research*, 2011, **346**, 1860-1867.
41. D. O. Otieno, *Comprehensive Reviews in Food Science and Food Safety*, 2010, **9**, 471-482.
42. S. Kitao, T. Ariga, T. Matsudo and H. Sekine, *Bioscience, Biotechnology, and Biochemistry*, 1993, **57**, 2010-2015.
43. D. Sprogoe, L. A. van den Broek, O. Mirza, J. S. Kastrup, A. G. Voragen, M. Gajhede and L. K. Skov, *Biochemistry*, 2004, **43**, 1156-1162.
44. O. Mirza, L. K. Skov, D. Sprogoe, L. A. van den Broek, G. Beldman, J. S. Kastrup and M. Gajhede, *J Biol Chem*, 2006, **281**, 35576-35584.
45. V. Lombard, H. Golaconda Ramulu, E. Drula, P. M. Coutinho and B. Henrissat, *Nucleic acids research*, 2014, **42**, 490-495.
46. M. R. Stam, E. G. J. Danchin, C. Rancurel, P. M. Coutinho and B. Henrissat, *Protein Engineering Design and Selection*, 2006, **19**, 555-562.
47. D. E. Koshland, *Biological Reviews*, 1953, **28**, 416-436.
48. J. Voet and R. H. Abeles, *Journal of Biological Chemistry*, 1966, **241**, 2731-2732.
49. P. Wildberger, C. Luley-Goedl and B. Nidetzky, *FEBS Letters*, 2011, **585**, 499-504.
50. A. Levasseur, F. Piumi, P. M. Coutinho, C. Rancurel, M. Asther, M. Delattre, B. Henrissat, P. Pontarotti, M. Asther and E. Record, *Fungal genetics and biology : FG & B*, 2008, **45**, 638-645.

51. E. A. MacGregor, Š. Janeček and B. Svensson, *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, 2001, **1546**, 1-20.
52. A. M. Swistowska, S. Gronert, S. Wittrock, W. Collisi, H.-J. Hecht and B. Hofer, *FEBS Letters*, 2007, **581**, 4036-4042.
53. H. Leemhuis, T. Pijning, J. M. Dobruchowska, B. W. Dijkstra and L. Dijkhuizen, *Biocatalysis and Biotransformation*, 2012, **30**, 366-376.
54. G. J. Davies, K. S. Wilson and B. Henrissat, *Biochemical Journal*, 1997, **321**, 557-559.
55. H. Takata, T. Kuriki, S. Okada, Y. Takesada, M. Iizuka, N. Minamiura and T. Imanaka, *Journal of Biological Chemistry*, 1992, **267**, 18447-18452.
56. A. Schwarz and B. Nidetzky, *FEBS Letters*, 2006, **580**, 3905-3910.
57. A. Schwarz, L. Brecker and B. Nidetzky, *Biochemical Journal*, 2007, **403**, 441-449.
58. M. Mueller and B. Nidetzky, *FEBS Letters*, 2007, **581**, 1403-1408.
59. C. Goedl, A. Schwarz, M. Mueller, L. Brecker and B. Nidetzky, *Carbohydrate Research*, 2008, **343**, 2032-2040.
60. C. Luley-Goedl and B. Nidetzky, *Carbohydrate Research*, 2010, **345**, 1492-1496.
61. T. Verhaeghe, M. Diricks, D. Aerts, W. Soetaert and T. Desmet, *Journal of Molecular Catalysis B: Enzymatic*, 2013, **96**, 81-88.
62. M. Mueller and B. Nidetzky, *FEBS Letters*, 2007, **581**, 3814-3818.
63. P. Wildberger, G. A. Aish, D. L. Jakeman, L. Brecker and B. Nidetzky, *Biochemistry and Biophysics Reports*, 2015, **2**, 36-44.
64. C. Goedl, A. Schwarz, A. Minani and B. Nidetzky, *Journal of Biotechnology*, 2007, **129**, 77-86.
65. F. De Bruyn, B. De Paepe, J. Maertens, J. Beauprez, P. De Cocker, S. Mincke, C. Stevens and M. De Mey, *Biotechnology and Bioengineering*, 2015, **112**, 1594-1603.
66. F. De Bruyn, M. Van Brempt, J. Maertens, W. Van Bellegem, D. Duchi and M. De Mey, *Microbial Cell Factories*, 2015, **14**, 138.
67. C. Weyler and E. Heinzle, *Applied Biochemistry and Biotechnology*, 2015, **175**, 3729-3736.
68. J. Pei, P. Dong, T. Wu, L. Zhao, X. Fang, F. Cao, F. Tang and Y. Yue, *Journal of Agricultural and Food Chemistry*, 2016, **64**, 7966-7972.
69. M. Nishimoto and M. Kitaoka, *Carbohydrate Research*, 2009, **344**, 2573-2576.
70. L. A. M. van den Broek, E. L. van Boxtel, R. P. Kievit, R. Verhoef, G. Beldman and A. G. J. Voragen, *Applied Microbiology and Biotechnology*, 2004, **65**, 219-227.
71. S. Kitao, S. Yoshida, T. Horiuchi, H. Sekine and I. Kusakabe, *Biosci Biotech Bioch*, 1994, **58**, 790-791.
72. K. Morimoto, A. Yoshihara, T. Furumoto and G. Takata, *J Biosci Bioeng*, 2015, **119**, 652-656.
73. C. Goedl, T. Sawangwan, M. Mueller, A. Schwarz and B. Nidetzky, *Angewandte Chemie*, 2008, **47**, 10086-10089.
74. C. Luley-Goedl, T. Sawangwan, L. Brecker, P. Wildberger and B. Nidetzky, *Carbohydrate Research*, 2010, **345**, 1736-1740.
75. P. Wildberger, L. Brecker and B. Nidetzky, *Chemical Communications*, 2014, **50**, 436-438.
76. J. M. Bolivar, C. Luley-Goedl, E. Leitner, T. Sawangwan and B. Nidetzky, *Journal of Biotechnology*, 2017, **257**, 131-138.
77. K. De Winter, W. Soetaert and T. Desmet, *International Journal of Molecular Sciences*, 2012, **13**, 11333-11342.
78. S. Kitao and H. Sekine, *Bioscience, Biotechnology, and Biochemistry*, 1994, **58**, 38-42.
79. S. Kitao, T. Matsudo, T. Sasaki, T. Koga and M. Kawamura, *Bioscience, Biotechnology, and Biochemistry*, 2000, **64**, 134-141.
80. K. Nomura, K. Sugimoto, H. Nishiura, K. Ohdan, T. Nishimura, H. Hayashi and T. Kuriki, *Bioscience, Biotechnology, and Biochemistry*, 2008, **72**, 82-87.
81. L. Brecker, M. Mahut, A. Schwarz and B. Nidetzky, *Magnetic Resonance in Chemistry*, 2009, **47**, 328-332.

82. K. Sugimoto, K. Nomura, H. Nishiura, K. Ohdan, K. Ohdan, H. Hayashi and T. Kuriki, *J Biosci Bioeng*, 2007, **104**, 22-29.
83. T. Kwon, C. T. Kim and J.-H. Lee, *Biotechnology Letters*, 2007, **29**, 611.
84. R. K. Gudiminchi and B. Nidetzky, *ChemBioChem*, 2017, **18**, 1387-1390.
85. T. Verhaeghe, K. De Winter, M. Berland, R. De Vreese, M. D'Hooghe, B. Offmann and T. Desmet, *Chem Commun (Camb)*, 2016, **52**, 3687-3689.
86. P. Wildberger, A. Todea and B. Nidetzky, *Biocatalysis and Biotransformation*, 2012, **30**, 326-337.
87. P. Abad, J. Gouzy, J. M. Aury, P. Castagnone-Sereno, E. G. Danchin, E. Deleury, L. Perfus-Barbeoch, V. Anthouard, F. Artiguenave, V. C. Blok, M. C. Caillaud, P. M. Coutinho, C. Dasilva, F. De Luca, F. Deau, M. Esquibet, T. Flutre, J. V. Goldstone, N. Hamamouch, T. Hewezi, O. Jaillon, C. Jubin, P. Leonetti, M. Magliano, T. R. Maier, G. V. Markov, P. McVeigh, G. Pesole, J. Poulain, M. Robinson-Rechavi, E. Sallet, B. Segurens, D. Steinbach, T. Tytgat, E. Ugarte, C. van Ghelder, P. Veronico, T. J. Baum, M. Blaxter, T. Bleve-Zacheo, E. L. Davis, J. J. Ewbank, B. Favery, E. Grenier, B. Henrissat, J. T. Jones, V. Laudet, A. G. Maule, H. Quesneville, M. N. Rosso, T. Schiex, G. Smant, J. Weissenbach and P. Wincker, *Nature biotechnology*, 2008, **26**, 909-915.
88. K. Fujii, M. Iiboshi, M. Yanase, T. Takaha and T. Kuriki, *Journal of Applied Glycoscience*, 2006, **53**, 91-97.
89. A. Cerdobbel, K. De Winter, D. Aerts, R. Kuipers, H. J. Joosten, W. Soetaert and T. Desmet, *Protein Eng Des Sel*, 2011, **24**, 829-834.
90. C. Goedl and B. Nidetzky, *ChemBioChem*, 2009, **10**, 2333-2337.
91. D. H. M. E., V. Tom, D. W. Karel and D. Tom, *Angewandte Chemie International Edition*, 2015, **54**, 9289-9292.
92. M. Friedman, *Mol Nutr Food Res*, 2007, **51**, 116-134.
93. T. P. Cushnie and A. J. Lamb, *Int J Antimicrob Agents*, 2011, **38**, 99-107.
94. C. Ferreira, D. C. Soares, M. T. Nascimento, L. H. Pinto-da-Silva, C. G. Sarzedas, L. W. Tinoco and E. M. Saraiva, *Antimicrob Agents Chemother*, 2014, **58**, 6197-6208.
95. E. R. Kasala, L. N. Bodduluru, R. M. Madana, A. K. V, R. Gogoi and C. C. Barua, *Toxicol Lett*, 2015, **233**, 214-225.
96. K. S. Bhullar and B. P. Hubbard, *Biochim Biophys Acta*, 2015, **1852**, 1209-1218.
97. L. Biasutto and M. Zoratti, *Curr Drug Metab*, 2014, **15**, 77-95.
98. C. K. Singh, M. A. Ndiaye and N. Ahmad, *Biochim Biophys Acta*, 2015, **1852**, 1178-1185.
99. R. W. Gantt, P. Peltier-Pain and J. S. Thorson, *Natural Product Reports*, 2011, **28**, 1811-1853.
100. T. Desmet, W. Soetaert, P. Bojarova, V. Kren, L. Dijkhuizen, V. Eastwick-Field and A. Schiller, *Chemistry*, 2012, **18**, 10786-10801.
101. J. Gorl, M. Timm and J. Seibel, *Chembiochem*, 2012, **13**, 149-156.
102. M. Timm, J. Gorl, M. Kraus, S. Kralj, H. Hellmuth, R. Beine, K. Buchholz, L. Dijkhuizen and J. Seibel, *Chembiochem*, 2013, **14**, 2423-2426.
103. F. A. Shaikh and S. G. Withers, *Biochem Cell Biol*, 2008, **86**, 169-177.
104. Z. Sun, A. Ilie and M. T. Reetz, *Angewandte Chemie*, 2015, **54**, 9158-9160.
105. K. De Winter, K. Verlinden, V. Kren, L. Weignerova, W. Soetaert and T. Desmet, *Green Chemistry*, 2013, **15**, 1949-1955.
106. A. Cerdobbel, K. De Winter, T. Desmet and W. Soetaert, *Biotechnol J*, 2010, **5**, 1192-1197.
107. P. Aramsangtienchai, W. Chavasiri, K. Ito and P. Pongsawasdi, *Journal of Molecular Catalysis B: Enzymatic*, 2011, **73**, 27-34.
108. J. F. Hao-bin Hu, *Bulletin of the Korean Chemical Society*, 2009, **30**, 703-706.
109. W. Kabsch, *Acta crystallographica. Section D, Biological crystallography*, 2010, **66**, 125-132.
110. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *Journal of applied crystallography*, 2007, **40**, 658-674.

111. P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta crystallographica. Section D, Biological crystallography*, 2010, **66**, 486-501.
112. S. Murosaki, K. Muroyama, Y. Yamamoto, H. Kusaka, T. Liu and Y. Yoshikai, *Biosci Biotech Bioch*, 1999, **63**, 373-378.
113. A. W. Dox and R. E. Neidig, *Journal of Biological Chemistry*, 1914, **18**, 167-175.
114. S. A. Barker and T. R. Carrington, *J Chem Soc*, 1953, 3588-3593.
115. Y. Konishi and K. Shindo, *Biosci Biotech Bioch*, 1997, **61**, 439-442.
116. K. Matsuda, H. Watanabe, K. Fujimoto and K. Aso, *Nature*, 1961, **191**, 278.
117. F. M. Takanori Nihira, Kazuhiro Chiku, Mamoru Nishimoto, and K. i. O. a. H. N. Motomitsu Kitaoka, *J Appl Glycosci*, 2014, **61**, 75-80.
118. T. Yamamoto, T. Unno, M. Sugawaral and T. Goda, *Journal of Applied Glycoscience*, 1999, **46**, 475-482.
119. M. Kraus, C. Grimm and J. Seibel, *Chembiochem*, 2016, **17**, 33-36.
120. D. Sprogoe, L. A. M. van den Broek, O. Mirza, J. S. Kastrup, A. G. J. Voragen, M. Gajhede and L. K. Skov, *Biochemistry*, 2004, **43**, 1156-1162.
121. C. Goedl, T. Sawangwan, P. Wildberger and B. Nidetzky, *Biocatalysis and Biotransformation*, 2010, **28**, 10-21.
122. T. Verhaeghe, Thesis ISBN 9789059897465, Ghent University, 2014. .
123. A. Cerdobbel, T. Desmet, K. De Winter, J. Maertens and W. Soetaert, *J Biotechnol*, 2010, **150**, 125-130.
124. D. Goffin, P. Bystricky, A. S. Shashkov, M. Lynch, E. Hanon, M. Paquot and A. V. Savage, *Bulletin of the Korean Chemical Society*, 2009, **30**, 2535-2541.
125. O. Trott and A. J. Olson, *J Comput Chem*, 2010, **31**, 455-461.
126. N. Qureshi and D. V. Tamhane, *Applied Microbiology and Biotechnology*, 1985, **21**, 280-281.
127. M. F. Sanner, *J Mol Graph Model*, 1999, **17**, 57-61.
128. Y. Zhang, *Bmc Bioinformatics*, 2008, **9**.
129. C. Zhang, B. R. Griffith, Q. Fu, C. Albermann, X. Fu, I. K. Lee, L. Li and J. S. Thorson, *Science*, 2006, **313**, 1291-1294.
130. C. Liang, Y. Zhang, Y. Jia, W. Wenzhao, Y. Li, S. Lu, J.-M. Jin and S.-Y. Tang, *Scientific Reports*, 2016, **6**, 21051.
131. Y. Malbert, S. Pizzut-Serin, S. Massou, E. Cambon, S. Laguerre, P. Monsan, F. Lefoulon, S. Morel, I. André and M. Remaud-Simeon, *ChemCatChem*, 2014, **6**, 2282-2291.
132. M. Zhou, A. Hamza, C.-G. Zhan and J. S. Thorson, *Journal of Natural Products*, 2013, **76**, 279-286.
133. P. S. Coelho, E. M. Brustad, A. Kannan and F. H. Arnold, *Science*, 2013, **339**, 307-310.
134. J. B. Siegel, A. Zanghellini, H. M. Lovick, G. Kiss, A. R. Lambert, J. L. St.Clair, J. L. Gallaher, D. Hilvert, M. H. Gelb, B. L. Stoddard, K. N. Houk, F. E. Michael and D. Baker, *Science*, 2010, **329**, 309-313.
135. R. Obexer, A. Godina, X. Garrabou, P. R. E. Mittl, D. Baker, A. D. Griffiths and D. Hilvert, *Nat Chem*, 2017, **9**, 50-56.
136. C. K. Savile, J. M. Janey, E. C. Mundorff, J. C. Moore, S. Tam, W. R. Jarvis, J. C. Colbeck, A. Krebber, F. J. Fleitz, J. Brands, P. N. Devine, G. W. Huisman and G. J. Hughes, *Science*, 2010, **329**, 305-309.
137. M. Kraus, J. Gorl, M. Timm and J. Seibel, *Chem Commun (Camb)*, 2016, **52**, 4625-4627.
138. P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart, *Acta Crystallogr D*, 2010, **66**, 213-221.
139. R. Eisenthal and A. Cornish-Bowden, *Biochemical Journal*, 1974, **139**, 715-720.
140. M. Kraus, C. Grimm and J. Seibel, *Chem Commun (Camb)*, 2017, **53**, 12181-12184.

141. M. E. Dirks-Hofmeister, T. Verhaeghe, K. De Winter and T. Desmet, *Angew Chem Int Ed Engl*, 2015, **54**, 9289-9292.
142. P. A. Karplus and K. Diederichs, *Science*, 2012, **336**, 1030-1033.