Advanced Synthesis & Catalysis

Accepted Article

Title: How Site-Directed Mutagenesis Boosted Selectivity of a Promiscuous Enzyme

Authors: Pavlína Nekvasilová, Natallia Kulik, Nikola Rychlá, Helena Pelantova, Lucie Petraskova, Zuzana Bosáková, Josef Cvacka, Kristyna Slamova, Vladimir Kren, and Pavla Bojarová

To be cited as: *Adv. Synth. Catal.* 10.1002/adsc.202000604

Link to VoR: https://doi.org/10.1002/adsc.202000604

10.1002/adsc.202000604

FULL PAPER

DOI: 10.1002/adsc.202000604

How Site-Directed Mutagenesis Boosted Selectivity of a Promiscuous Enzyme

Pavlína Nekvasilová,[a,b,c] Natalia Kulik,[d] Nikola Rychlá,[a, e] Helena Pelantová,[a] Lucie Petrásková,[a] Zuzana Bosáková,[c] Josef Cvačka,[f] Kristýna Slámová,[a] Vladimír Křen,[a] and Pavla Bojarová[a,e,*]

[a] Institute of Microbiology, Czech Academy of Sciences, Vídeňská 1083, CZ-14220 Praha 4, Czech Republic
Phone: (+420) 296442360; Fax: (+420) 244471286; E-mail: bojarova@biomed.cas.cz
[b] Department of Genetics and Microbiology, Faculty of Science, Charles University, Viničná 5, CZ-12843, Praha 2, Czech Republic
[c] Department of Analytical Chemistry, Faculty of Science, Charles University, Hlavova 2030/8. CZ-12843 Praha 2, Czech Republic
[d] Center for Nanobiology and Structural Biology, Institute of Microbiology, Czech Academy of Sciences, Zámek 136, CZ-37333 Nové Hrady, Czech Republic
[e] Department of Health Care Disciplines and Population Protection, Faculty of Biomedical Engineering, Czech Technical University in Prague, Nám. Sítná 3105, CZ-27201 Kladno, Czech Republic
[f] Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Flemingovo nám. 2, CZ-16610 Praha 6, Czech Republic

Received: ((will be filled in by the editorial staff))

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/adsc.201######.((Please delete if not appropriate))

Abstract. β-*N*-Acetylhexosaminidases (GH20; EC 3.2.1.52) are *exo*-glycosidases with a dual activity for cleaving both *N*-acetylglucosamine (GlcNAc) and *N*-acetylgalactosamine (GalNAc) units from glycostructures. This substrate promiscuity is a hurdle in the selective synthesis of *N*-acetylhexosamine oligosaccharides combining both GlcNAc and GalNAc units since there are hardly any GalNAc transferring enzymes available for synthetic applications. We present here site-directed mutagenesis of a synthetically potent promiscuous β-*N*-acetylhexosaminidase from *Talaromyces flavus* (*Tf*Hex), which, as a wild type, exhibits a GalNAcase/GlcNAcase ratio of 1.2. On the basis of molecular modeling, we identified crucial amino acid residues responsible for its GalNAcase/GlcNAcase selectivity. Six site-directed mutants were prepared, heterologously expressed in *Pichia pastoris*, purified, and kinetically characterized. As a result, novel engineered enzymes with an up to 7-times higher selectivity for either GalNAc or GlcNAc substrates were obtained, preserving the favorable properties of the wild type *Tf*Hex, mainly its transglycosylation potential and tolerance to functional groups in the substrate molecule. The substrate selectivity and transglycosylation yield were further corroborated by reaction engineering. The new selective and synthetically capable enzymes were applied in the preparation of tailored *N*-acetylhexosamines.

Keywords: β-*N*-acetylhexosaminidase; glycosylation; *Talaromyces flavus*; *Pichia pastoris*; site-directed mutagenesis; substrate specificity

Introduction

Fungal β-*N*-acetylhexosaminidases (EC 3.2.1.52, GH20, http://www.cazy.org/) are *exo*-glycosidases of family GH20, which naturally catalyze the cleavage of *N*-acetylglucosamine (GlcNAc) and *N*-acetylgalactosamine (GalNAc) moieties from glycostructures.[1] However, under suitable reaction conditions, these enzymes can also catalyze the formation of the glycosidic bond *in vitro*.[2] This ability may be fortified by reducing the water activity, adding organic co-solvents[3] or inorganic salts[4] as well as using a high concentration of the glycosyl acceptor.

β-*N*-Acetylhexosaminidases utilize the substrate-assisted mechanism with the key catalytic conserved amino acid pair of Asp (stabilizing residue) and Glu (catalytic acid/base).[5] Furthermore, their active center comprises tyrosine, proposed to stabilize the oxazoline reaction intermediate.[6] A model β-*N*-acetylhexosaminidase from *Talaromyces flavus*

(*Tf*Hex), well known for its broad substrate specificity and high transglycosylation capability,[7-9] was studied for the effect of tyrosine substitution on the suppression of hydrolytic activity[10] and on the GalNAcase/GlcNAcase activity ratio.[11] This enzyme as a wild-type (WT) is a typical promiscuous β-*N*-acetylhexosaminidase, featuring the GalNAcase/GlcNAcase activity ratio of 1.2.[11] In general, the specificity of fungal β-*N*-acetylhexosaminidases varies from relatively selective GlcNAcases (enzymes from *Aspergillus versicolor* (*Av*Hex) with GalNAcase/GlcNAcase ratio 0.09-0.11[12] or from *Aspergillus oryzae* (*Ao*Hex) with 0.3-0.6)[13] to slightly prevalent GalNAcases (*e.g.*, *Penicillium oxalicum* β-*N*-acetylhexosaminidase *Po*Hex with GalNAcase/GlcNAcase ratio of up to 1.7).[14] Importantly, despite these notable changes in the substrate specificity, all these enzymes share quite a high homology. For example, *Tf*Hex exhibits a 62% homology with the *Ao*Hex[15] and 60% with *Po*Hex.[16] The *Tf*Hex protein is a tetramer consisting of two non-covalently associated subunits (each 65 kDa) and two non-covalently linked propeptides (each 15 kDa), accounting for the total of 160 kDa of the native enzyme.[17]

Whereas β-*N*-acetylhexosaminidases with a naturally increased preference for GlcNAc substrates occur among glycosidases, pure β-GalNAcase activity is quite rare in this class.[12] Described WT *N*-acetylgalactosaminyltransferases[18-19] are generally not applicable for the laboratory synthesis of oligosaccharides as they are strictly specific for particular acceptors (*e.g.*, polypeptides), unlike rather universal glycosidases. A more versatile GalNAc-transferring activity was found in a mutant galactosyltransferase from human placenta.[20] The so far described selective β-GalNAcases, namely from *Paenibacillus*,[21] *Clostridium*,[22] and *Bacteroides*[23] species, are of bacterial origin and have no documented transglycosylation capabilities.[24] Among GlcNAcases, many representatives also exhibit purely hydrolytic activity.[12, 25] Moreover, even in the very selective GlcNAcases, traces of GalNAcase activity may be found (*e.g.*, in *O*-GlcNAcase,[26] the GalNAcase/GlcNAcase ratio is 0.03). It is also noteworthy that although the main source of specificity of GalNAcase/GlcNAcase is encoded by the primary amino acid sequence of the β-*N*-acetylhexosaminidase, other factors, not yet fully understood, are undeniably involved, *e.g.*, the effect of posttranslational modifications or of heterologous expression in a different host.[11]

In the present work, original results were obtained disclosing the impact of a single point mutation in the amino acid sequence of *Tf*Hex on the GalNAcase/GlcNAcase ratio of the respective mutant enzyme. Based on homology model of *Tf*Hex WT,[27-28] we have found that Arg218 and Glu546 residues of *Tf*Hex are involved in direct hydrogen bond interaction with the substrate C4 hydroxyl (Figure 1), and thus are mutation hotspots for altering the substrate specificity of the enzyme both for hydrolysis and transglycosylation. Prospective point mutations were proposed based on a combination of *in silico* saturation mutagenesis and visualization of enzyme-substrate interactions in prospective mutants. A total of six single mutant variants of *Tf*Hex were studied concerning their specific activity, GalNAcase/GlcNAcase activity ratio and, first and foremost, their applicability as synthetic tools in the challenging preparation of combined *N*-acetylhexosamine oligosaccharides.

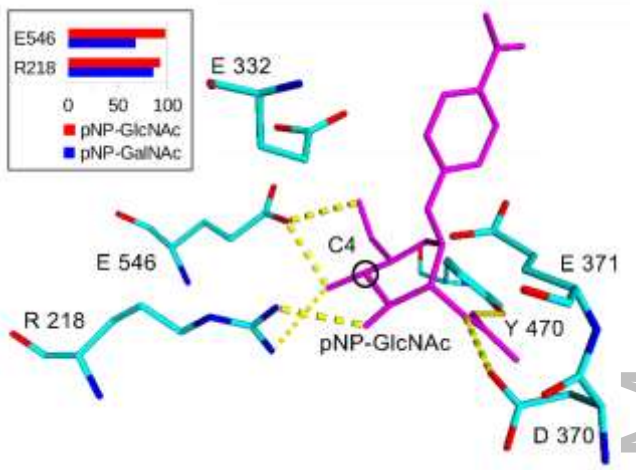


Figure 1. Molecular model of the active site of *Tf*Hex WT [27-28] with docked *p*NP-GlcNAc (**magenta**) after 5 ns of molecular dynamics simulation. For details please see the Experimental Section. Hydrogen bonds formed in the snapshot are shown in yellow dashed lines, other possible hydrogen bonds appearing with Glu546 and Arg218 residues during molecular dynamics are shown in yellow dotted lines. The inset in the upper left corner shows the percentage of snapshots during 2-20 ns of molecular dynamics where hydrogen bonds were formed with amino acid residues surrounding the C4 hydroxyl of *p*NP-GlcNAc and *p*NP-GalNAc. The C4 atom is encircled in **black**, hydrogens are hidden. The catalytic amino acid pair of *Tf*Hex are Asp370 (stabilizing residue) and Glu371 (acid/base).

Results and Discussion

Heterologous Expression, Purification and Characterization of *Tf*Hex Mutants

On the basis of molecular modelling described in detail in the next section, six single-point mutant variants of the *Tf*Hex enzyme were proposed, which should feature altered GalNAcase/GlcNAcase activity ratio. The genes of five *Tf*Hex mutant variants, namely Glu546Tyr, Glu546Gln, Glu546His, Arg218His and

Arg218Lys, were prepared commercially; the mutant variant of the Arg218Gln was prepared by PCR mutagenesis. The synthetic genes were cloned into the pPICZαA vector, and electroporated into *P. pastoris* competent cells KM71H. Expression vectors contained *Eco*RI and *Kpn*I restriction sites downstream of the α-factor that encodes extracellular targeting of the enzyme and resistance to zeocin. The expression of enzymes was screened at a small scale. For each mutant variant, sixteen colonies were inoculated in the complex medium upon induction by methanol. The mutant enzyme production by individual colonies of transformed *P. pastoris* was checked by SDS-PAGE and the enzyme activities were determined by a standard activity assay. Selected colonies were cryopreserved for later production. Two of the prepared *Tf*Hex mutants, Glu546Tyr and Arg218Gln, were successfully expressed and purified; however, they were found to possess no enzymatic activity. The other four mutant variants, Glu546Gln, Glu546His, Arg218His, Arg218Lys *Tf*Hex, were produced as active enzymes.

The preparative production of *Tf*Hex mutant variants was performed in minimal cultivation medium for easier purification. The extracellularly produced enzymes were purified to homogeneity by cation-exchange chromatography at pH 3.5 in a single purification step. The purity of the enzymes was confirmed by SDS-PAGE (see the Supporting Information, Figure S2). The purification yield, the GlcNAcase, and GalNAcase activities and their ratio were determined and compared to the recombinant wild-type *Tf*Hex (Table 1).

Table 1. Purification and yields of active *Tf*Hex mutant variants, and their specific GalNAcase and GlcNAcase activities

***Tf*Hex variant**	**Yield [b] [mg]**	**GlcNAcase [U mg^{-1}]**	**GalNAcase [U mg^{-1}]**	**GalNAcase/GlcNAcase ratio**
WT [a]	47 (58 %)	37	45	1.2
Glu546Gln	52 (71 %)	0.5	3.5	6.5
Glu546His	61 (66 %)	0.1	0.8	7.1
Arg218His	39 (52 %)	26	50	1.9
Arg218Lys	37 (52 %)	36	5.3	0.15

[a] Values were adopted from Bojarova *et. al*.[11]
[b] Protein yield from purification of 200 mL of cultivation medium.

The found specific activities clearly show that the enzyme GalNAcase/GlcNAcase ratio was strongly influenced by the point mutation introduced. While the wild-type strain exhibits a ratio of these activities of 1.2, in the Arg218Lys mutant this ratio significantly decreased to 0.15, affording a GlcNAcase. In other words, the introduction of Lys into the active center instead of Arg at position 218 caused an eight-fold increase in selectivity of the enzyme to *p*NP-GlcNAc (**1**) whereas the specific activity remained comparable to WT.

In contrast, by substituting Glu at position 546 for either Gln or His, the selectivity to *p*NP-GalNAc increased more than five times. Remarkably, both of these mutant GalNAcases showed a significant decrease in specific catalytic activity (13-times in the case of Glu546Gln, 56-times in the case of Glu546His *Tf*Hex). Apparently, the position of Glu546 is much more sensitive to changes in the amino acid sequence; this phenomenon was also observed in other functional mutants of glycosidases, namely mutant transglycosidases[11] and glycosynthases.[29] The specific catalytic activity in both of these groups of mutant enzymes is drastically reduced and, to perform the desired catalytic reaction, high protein amounts need to be used.

In the case of the last mutant variant, Arg218His *Tf*Hex, the introduction of His into the active center in position 218 caused only a slightly increased GalNAcase/GlcNAcase ratio, yielding a weakly specific GalNAcase. However, the Arg218His point mutation caused that the enzyme activity decreased completely during the week. This is in sharp contrast to the other enzymes used in this work, WT or mutant, which were active for more than one year at 4 °C. Basic biochemical parameters were measured for each of the four active mutant variants. pH Optima profiles are shown in the Supporting information, Figure S3. In the case of the selective variants Glu546Gln, Glu546His, and Arg218Lys *Tf*Hex, we analyzed just their major catalytic activity (GlcNAcase or GalNAcase). The selective mutant variants had their activity optimum at pH 5, which falls within the pH optima plateau found for *Tf*Hex WT.[11, 17] For the Arg218His mutant, a slightly pronounced GalNAcase (ratio 1.9), the pH optimum of both GalNAcase and GlcNAcase activities was flat in the range of 4.0-5.5. Interestingly, the GalNAcase activity of this variant was relatively stable under basic conditions (pH 7 to 12.5), contrary to its GlcNAcase activity, which rapidly decreased from pH 7, and from pH 9 it was so low that it fell behind the detection threshold of the

spectrophotometer. This phenomenon caused a rapid increase in the GalNAcase/GlcNAcase ratio of Arg218His *Tf*Hex at a higher pH. This phenomenon was observed before with other *Tf*Hex variants[11] and probably relates to changed protonation states of the key amino acid residues in the active site.

The dependence of the hydrolytic activity of all mutant variants on the temperature is shown in the Supporting Information, Figure S4. The graphs show that the temperature optimum of the mutant *Tf*Hex does not differ significantly from *Tf*Hex WT (65 °C).[17] For the Glu546Gln and Arg218Lys *Tf*Hex mutants, temperature optimum was 60 °C, and it was 55 °C for Glu546His *Tf*Hex. The Arg218His *Tf*Hex reported a rather flat temperature optimum between 55 °C and 65 °C for GalNAcase and 45-60 °C for GlcNAcase activities.

Furthermore, Michaelis-Menten parameters of the hydrolytic activity of all active mutants were acquired (Table 2, and Figure S5 in the Supporting Information).

Table 2. Kinetic parameters of *p*NP-GlcNAc (**1**) and *p*NP-GalNAc (**2**) hydrolysis by mutant *Tf*Hex variants

*Tf*Hex variant	Activity	K_M [mM]	k_{cat} [s^{-1}]	k_{cat}/K_M [s^{-1} mM^{-1}]
WT [a]	GlcNAcase	0.11 ± 0.02	47 ± 1	434
	GalNAcase	0.69 ± 0.09	104 ± 5	150
Glu546Gln [b]	GalNAcase	0.8 ± 0.2	4.9 ± 0.3	6.1
Glu546His [b]	GalNAcase	0.39 ± 0.02	0.6 ± 0.1	1.6
Arg218His	GlcNAcase	3.3 ± 0.5	12 ± 1	3.6
	GalNAcase	2.1 ± 0.7	6.6 ± 0.9	3.1
Arg218Lys [b]	GlcNAcase	0.9 ± 0.2	29 ± 1	32

[a] Values were adopted from Bojarová *et. al*.[11]
[b] Kinetic parameters were determined only for major activity

The K_M values indicate that in selective GalNAcases Glu546Gln and Glu546His *Tf*Hex, the affinity to *p*NP-GalNAc remained either comparable to WT or even increased. This was not the case of Arg218 mutants, featuring in all cases a decrease in affinity to the respective substrates (higher K_M). In the Arg218Lys GlcNAcase, the catalytic efficiency (k_{cat}/K_M) decreased approximately 13-times compared to WT. In Glu546Gln and Glu546His GalNAcases, the catalytic efficiency was reduced even more considerably, by 25-times and 100-times, respectively In the Arg218His mutants, we found 48-times decreased catalytic efficiency for *p*NP-GalNAc. This drastic reduction in catalytic activity was in all cases primarily caused by a considerable decrease in the turnover number (k_{cat}). Thus, despite a rather strong binding to the active site, further catalytic steps such as formation of the oxazoline reaction intermediate and product release are altered through the mutation. Due to the high yield and cost-effectiveness of the recombinant enzyme production in *P. pastoris* and its facile downstream processing compared to other available eukaryotic hosts, the decrease in specific activity is not a major problem for any synthetic applications since larger amounts of enzyme can easily be added to the reaction to compensate for this problem. As already mentioned above, this phenomenon is quite usual in functional glycosidase mutants such as transglycosidases or glycosynthases.[20] Supression of the undesired activity rather than enhancement of the desired activity is a general weakness of the rational engineering. Possibly, an alternative may be dosed random mutagenesis around the binding site to optimize beneficial structure-function relations.

Docking and Molecular Dynamics Simulations of *Tf*Hex Mutants

To investigate the structural properties and binding preferences of the prepared *Tf*Hex mutant variants, we performed docking of *p*NP-GlcNAc or *p*NP-GalNAc substrates into the active sites of energetically minimized models of all variants. The respective molecular models were prepared on the basis from our established molecular model of *Tf*Hex WT[27] corrected with respect of the crystal structure of β-*N*-acetylhexosaminidase from *A. oryzae* [28] (for details see the Experimental Section). Averaged binding scores calculated for equilibrated enzyme-substrate complexes from molecular dynamics simulations are shown in Table 3.

Table 3. Averaged binding scores of substrates in the active site of prepared mutant variants (kJ mol^{-1}), calculated for equilibrated complexes after molecular dynamics simulation.

Variant [a]	*p*NP-GlcNAc	*p*NP-GalNAc	Score ratio *p*NP-GalNAc/ *p*NP-GlcNA*c*
WT [b]	-38.6 (*-23.5*)	-34.9 (*-29.5*)	0.9
Glu546Gln	-35.4	-38.2	1.08
Glu546His	-34.5	-40.3	1.17
Arg218His	-30.6	-32.3	1.06
Arg218Lys	-36.4	-32.0	0.88
Glu546Tyr [c]	-35.2	-38.4	1.09
Arg218Gln	-27.0	-31.9	1.18

[a] Mutant GalNAcases are in **green**, mutant GlcNAcases in **blue**, inactive mutants in **grey**.
[b] Values were adopted from Bojarová *et. al*.[11] The numbers in *italics* indicate the binding score of the respective hydrolytic product (GlcNAc or GalNAc, respectively).
[c] After structure minimization, the Tyr546 residue was forced to move away from the ligand to fit it in a position that enabled docking by induced fit. However, this movement led to the destruction of many interactions found in WT. *In vivo*, docking is likely to be impossible as could be seen from molecular dynamics simulation of mutated enzyme without ligand.

The binding score calculated by Glide XP algorithm characterizes the affinity of enzyme to the substrate. The score ratio may be used as an indication of enzyme discrimination between the substrates. From the ratio we could see that WT and Arg218Lys *Tf*Hex have a higher affinity (*i.e*., lower binding score) to *p*NP-GlcNAc, which is in agreement with lower K_M value for *p*NP-GlcNAc in WT. In contrast, the other active variants (Arg218His, Glu546Gln, and Glu546His *Tf*Hex) prefer *p*NP-GalNAc, which also correlates with the kinetic results (Table 2, and Figure S5 in the Supporting Information). To fully assess the structure-function relationship, the geometry of substrate-enzyme complex should also be considered.

In the WT, Arg218 in the active site forms hydrogen bonds with C3/C4 hydroxyl of *p*NP-GalNAc and *p*NP-GlcNAc (Figure 1). In the Arg218Lys *Tf*Hex variant (GlcNAcase), interaction of C4 hydroxyl of substrates with Lys218 is not possible due to the sp^3 hybridization of nitrogen of Lys (Figure 2A, B). While Lys218 in the Arg218Lys *Tf*Hex-*p*NP-GlcNAc complex interacts with the C3 hydroxyl of the substrate (Figure 2A), the axial C4 hydroxyl in *p*NP-GalNAc sterically hinders the formation of this interaction. As a result, Lys218 interacts with the neighboring active site residue Glu546, which leads to the decrease in the number of hydrogen bonds formed with *p*NP-GalNAc in the Arg218Lys variant (Supporting Information, Figure S6A) and increase in the binding score (Table 3).

In the case of Arg218His *Tf*Hex variant (unpronounced GalNAcase), the conformation of the His218 side chain is different (nitrogen in sp^2 hybridization). Thus, His218 can still make hydrogen bonding with C4 of *p*NP-GlcNAc in the enzyme-substrate complex (present during 52% of the molecular dynamics simulation time). However, the interaction of both substrates with Glu332 is completely lost, leading to a higher binding score (Table 3, Figure 2C). Even though *p*NP-GalNAc does not form hydrogen bonding with His218 (Figure 2D), *p*NP-GalNAc is better stabilized in the active site of Arg218His variant due to the improved interaction with Glu546. The observed instability of Arg218His variant could result from a highly flexible interaction with catalytic residues. The distance between Glu371 and the oxygen of the glycosidic bond often changed (increased) during molecular dynamics simulation (Supporting Information, Figure S6B), and catalytic Asp370 formed hydrogen bond with the *N*-acetyl group of the substrate in less than 30% of the simulation time.

The docking of *p*NP-GalNAc and *p*NP-GlcNAc in Arg218Gln variant (inactive) showed a slight increase in binding scores for both substrates and a decrease in the number of hydrogen bonds (Supporting Information, Figure S6A). Moreover, the binding score of substrates is comparable to the score of the product of hydrolysis (Table 3).[11-12] This may lead to fast substrate unbinding. In the Arg218Gln complex with both substrates (Supporting Information, Figure S7), the substrates changed orientation with respect to WT, and unusual interactions were formed with Glu218. The main reason for enzyme inactivity may be the uncompensated negative charge located close to Gln218 in the active site, which leads to reorientation of substrates and of catalytic Glu371, and to the formation of unusual hydrogen bonding by protonated Glu371 and Glu332 (Supporting Information, Figure S7). The change in charge distribution in the active site might influence deprotonation of the catalytic Glu371, and impair proton donation to the glycosidic oxygen.

10.1002/adsc.202000604

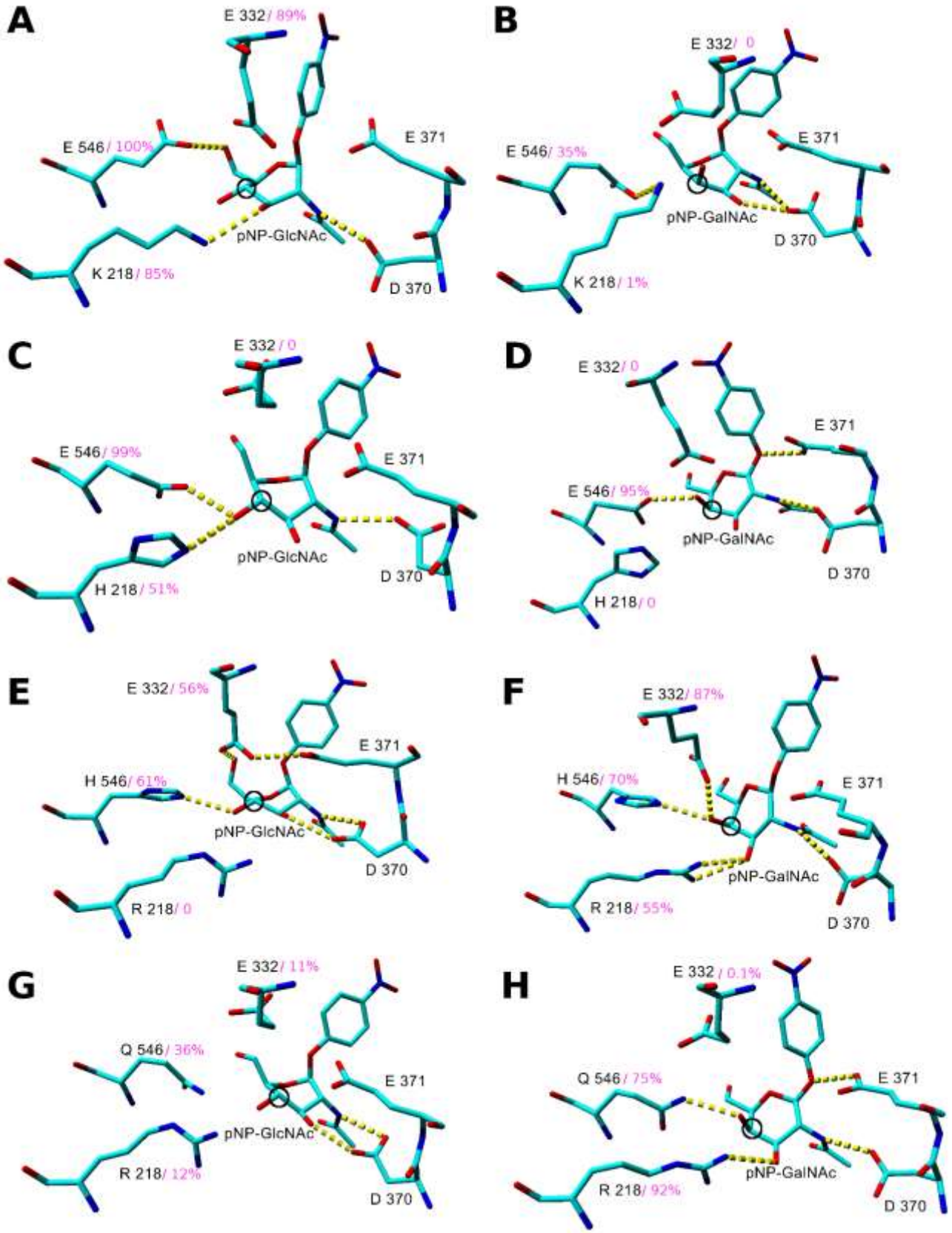


Figure 2. Orientation of docked substrates after 10 ns of molecular dynamics simulation. Active site amino acid residues, which could form hydrogen bond interaction with the substrate, are shown. Tyr470 and Trp509 residues are hidden for clarity. Amino acids and substrates are labeled in **black**; the probability of hydrogen bond formation with respective amino acid residues during molecular dynamics is indicated in **magenta**. Position of C4 carbon in substrates was marked in a black circle. (**A**), Arg218Lys *Tf*Hex-*p*NP-GlcNAc; (**B**), Arg218Lys *Tf*Hex-*p*NP-GalNAc; (**C**), Arg218His *Tf*Hex-*p*NP-GlcNAc; (**D**), Arg218His *Tf*Hex-*p*NP-GalNAc; (**E**), Glu546His *Tf*Hex-*p*NP-GlcNAc; (**F**), Glu546His *Tf*Hex-*p*NP-GalNAc; (**G**), Glu546Gln *Tf*Hex-*p*NP-GlcNAc; (**H**), Glu546Gln *Tf*Hex-*p*NP-GalNAc.

Another active-site residue close to the C4 hydroxyl, prospective for influencing the enzyme axial/equatorial selectivity at C4, is Glu546. In the WT this residue interacts with C4 or C6 hydroxyls of *p*NP-GlcNAc and with the C4 hydroxyl of *p*NP-GalNAc (Figure 1). Substitution of Glu546 residue by His or Gln led to a more favorable binding score compared to WT with a clear preference for *p*NP-GalNAc substrate. In the Glu546His variant (GalNAcase), His546 residue forms a hydrogen bond with *p*NP-GalNAc and reinforces interaction with Glu332 (higher probability of hydrogen bonding) by stabilizing the position of Glu332 closer to the substrate (Figure 2F). In the complex with *p*NP-GlcNAc, we observed interaction between Glu371 and Glu332 similar to the Arg218Gln variant (Supporting Information, Figure S7). This interaction is not so frequent here but still occurs in 68 % of molecular dynamics snapshots and leads to a partial inactivation of Glu371 for the glycosidic bond cleavage. It occurred as a result of changes in hydrogen bonding network: while Glu332 rotated and started to interact with the substrate C6 hydroxyl, His546 maintained its interaction with the substrate C4 hydroxyl but Arg218 completely lost interaction with the substrate (Figure 2E).

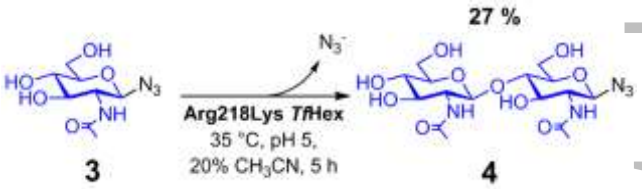
In the complex of Glu546Gln *Tf*Hex (GalNAcase) with *p*NP-GalNAc, the substrate is oriented closer to the catalytic Glu371 and it is stabilized by a great number of hydrogen bonds, including the hydrogen bond between the C4 hydroxyl and Gln546. Notably, *p*NP-GalNAc changed conformation from boat to semichair during molecular dynamics (Figure 2H). In contrast, *p*NP-GlcNAc lost interaction between its C4 hydroxyl and Gln546 and with Arg218 due to the bulkier side chain in position 546 upon Glu546Gln mutation (Figure 2G). The differences in hydrogen bond network influenced the orientation *p*NP-GlcNAc with respect to Glu371, leading to the formation of the inactivating hydrogen bond between Glu371 and Glu332 in 3 % of molecular dynamics snapshots, analogous to the Glu546His variant (Figure 2G).

The modeling of the Glu546Tyr mutation (inactive mutant variant) in the protein without substrate followed by molecular dynamics simulation showed that in the equilibrated model, Tyr546 could be in two distinct binding modes. Either it interacted with Trp509, occupying the place of the substrate C6 hydroxyl (mode **A**; Supporting Information, Figure S8A), or it interacted with Glu332 (mode **B**; Supporting Information, Figure S8B). This is probably caused by the high flexibility of Glu546Tyr *Tf*Hex and of its active site amino acid residues during molecular dynamics (Supporting Information, Figure S9). The orientation of Tyr546 before and after molecular dynamics simulation allowed docking of substrate only by induced fit because sterical hindrance with Tyr546 disallowed substrate binding.

Notably, in all mutants catalytic Asn370 interacts not only with the *N*-acetyl group of the substrate but also with its C3 hydroxyl (Figure 2B, E, G), which might also result in decreased catalytic activity.[30-31]

Transglycosylation Capability of Mutant Enzymes

Besides the biochemical characterization of prepared mutant enzymes, we aimed to address the question of their transglycosylation potential. As already mentioned, many native enzymes exhibiting a high selectivity to either GlcNAc[12] or GalNAc[21-23] moieties are useless for synthetic application due to the lack of transglycosylating activity. Moreover, we wanted to test if the remarkable substrate promiscuity valued in *Tf*Hex WT, *i.e*. its ability to process structurally modified substrates, had been maintained in the mutants. For this aim, we chose the unique transglycosylation reaction using GlcNAc azide substrate, which required the enzymatic hydrolysis of the C-N bond (Scheme 1). The Arg218Lys *Tf*Hex mutant easily cleaved the azido substrate and transferred the GlcNAc moiety to afford the corresponding disaccharide in 27% yield. Interestingly, the reaction was fully regioselective in contrast to the analogous reaction with *Tf*Hex WT,[32] with a comparable yield of the desired product (*cf*. the previously published 32% yield with the WT).[33]



Scheme 1. Enzymatic synthesis of functionalized disaccharide GlcNAcβ4GlcNAc-N_3 catalyzed by Arg218Lys *Tf*Hex.

We further hypothesized that the introduced mutation may have impact on the transglycosylation capability of the enzyme. For this aim, we analyzed the formation of *p*NP-functionalized chitooligomers (*N*,*N*´-diacetylchitobioside to *N*,*N*´,*N*´´,*N*´´´ tetraacetylchitotetraoside) under the catalysis of Arg218Lys *Tf*Hex mutant (Scheme 2) and compared it to an analogous reaction catalyzed by *Tf*Hex WT. We found that under the same reaction conditions, the Arg218Lys *Tf*Hex mutant forms longer chains of *p*NP-chitooligomers in higher yields. In total, we obtained 53 mg of *p*NP-chitooligomers in a ratio of 37: 23: 16 (**5**: **6**: **7**). The longer chain formation was previously observed only with mutant transglycosidases.[33] In contrast, *Tf*Hex WT yielded small amounts of chitobioside **5** and chitotrioside **6** (Supporting

10.1002/adsc.202000604

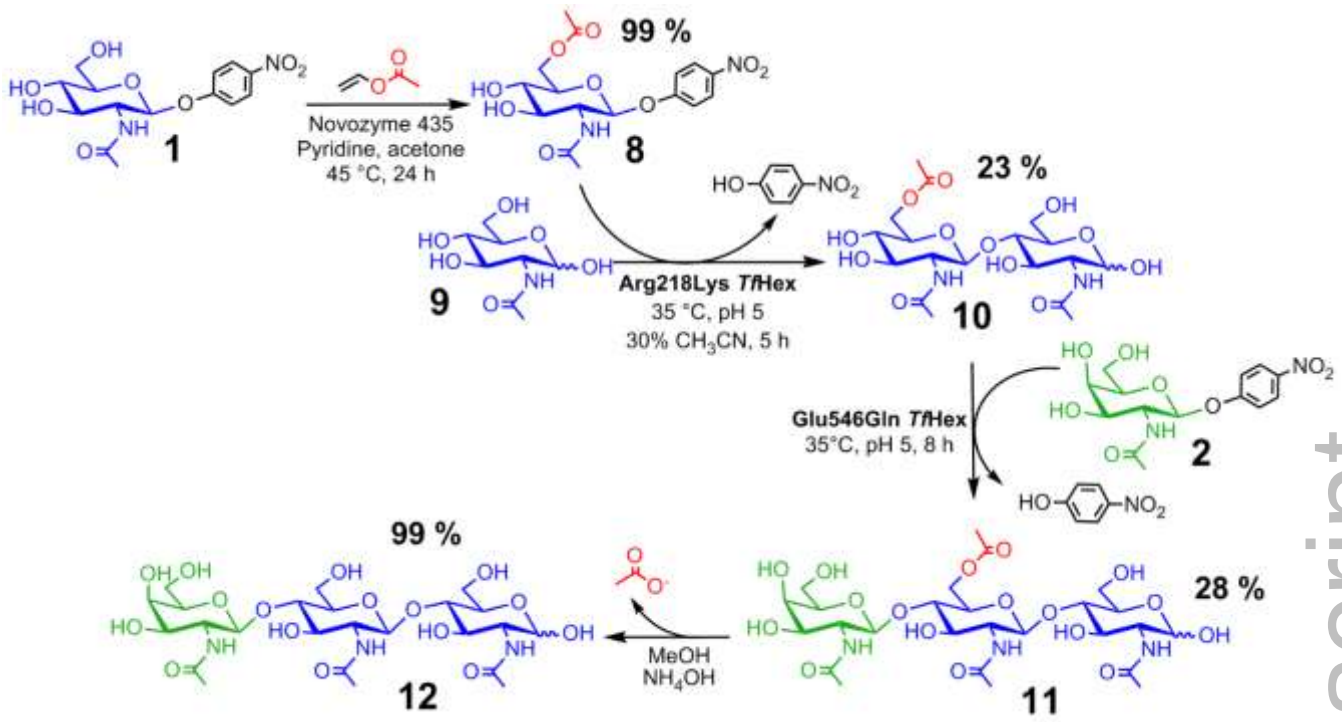
Information, Fig. S10). Structures of all prepared compounds were confirmed by NMR and MS (Supporting information, Figs. S12-14, Tables S3-5).

Scheme 2. Enzymatic synthesis of functionalized chitooligomers: *p*-nitrophenyl *N*,*N*´-diacetylchitobioside (GlcNAcβ4GlcNAc-*Op*NP), *N*,*N*´,*N*´´ -triacetylchitotrioside (GlcNAcβ4GlcNAcβ4GlcNAc-*Op*NP), and *N*,*N*´,*N*´´,*N*´´´ -tetraacetylchitotetraoside (GlcNAcβ4GlcNAcβ4GlcNAcβ4GlcNAc-*Op*NP) catalyzed by Arg218Lys *Tf*Hex.

Chemoenzymatic Synthesis of Complex Trisaccharides

We employed the prepared selective mutant enzyme variants for the synthesis of complex *N*-acetylhexosamine oligosaccharides to demonstrate their utility in this challenging synthetic task. *p*-Nitrophenyl glycosides are quite popular donors in preparative transglycosylation reactions. First, they are much cheaper than respective natural substrates (*e.g.*, *N*,*N*´-diacetylchitobiose). Secondly, they enable a more efficient separation of transglycosylation products from the remaining donor, e.g., by extraction of the aromatic portion on XAD-resin or into ethyl acetate. Our original intention to perform a selective GalNAc transfer to unprotected *N*,*N*'-diacetylchitobiose (GlcNAcβ4GlcNAc) with one of the new mutant GalNAcases, however, proved to be too challenging. Although the GalNAcase activity was strongly (more than 6-times) increased in both Glu546Gln and Glu546His mutant variants compared to WT, the residual GlcNAcase activity still caused a partial cleavage of the *N*,*N*'-diacetylchitobiose acceptor and the formation of a mixture of products. This is mainly because in glycosidase-catalyzed synthesis, the acceptor is used in a large molar excess (usually 5-10 times). Thus, to accomplish the desired synthesis of a complex oligosaccharide, we adopted a combined strategy employing the mutant enzyme along with a straightforward selective protection of *N*,*N*'-diacetylchitobiose acceptor at the C6 hydroxyl with an acetyl moiety. Though the C6 acetylated acceptor is still cleaved by the WT enzyme (with ca 6-fold reduced affinity), which is not sufficient for a selective transglycosylation reaction, the combination of substrate protection and the mutant enzyme Glu546Gln *Tf*Hex enabled to perform selective GalNAc transfer to the intact acceptor in a good yield. The influence of the 6-*O*-acetylation of the substrate on the affinity of mutant Glu546Gln *Tf*Hex and the WT enzyme is shown in the Supporting Information, Table S10 and Figure S19.

The C6 protected glycosyl donor *p*NP-6-*O*-Ac-β-D-GlcNAc (**8**, 99 % isolated yield) was prepared under the catalysis by Novozym 435 lipase.[34] It was then used as a glycosyl donor for the preparation of a selectively protected disaccharide **10** under the catalysis by the Arg218Lys *Tf*Hex mutant GlcNAcase (23 % isolated yield). Then, disaccharide **10** was used as an acceptor in the glycosylation by *p*NP-GalNAc donor under the catalysis by Glu546Gln *Tf*Hex, affording the 6'-*O*-acetylated trisaccharide **11** (28 % isolated yield), which was finally deprotected by ammonium hydroxide in methanol (1:10) to yield the desired complex trisaccharide GalNAcβ4GlcNAcβ4GlcNAc (**12**, 99 % isolated yield). The synthetic route is depicted in Scheme 3.



Scheme 3. Chemoenzymatic synthesis of the complex trisaccharide GalNAcβ4GlcNAcβ4GlcNAc catalyzed by a cascade reaction employing Novozym 435 lipase and two selective mutant *Tf*Hex variants.

Conclusion

By means of site-directed mutagenesis based on rational design following from *in silico* modeling, we were able to considerably enhance the substrate specificity of *Tf*Hex, a typically promiscuous β-*N*-acetylhexosaminidase with a comparable affinity to both GlcNAc- and GalNAc-terminated carbohydrates. Out of the six single-point mutants prepared, we identified two quite selective mutant GalNAcases and one GlcNAcase, with ca 7-times higher selectivity for either substrate. The enzymes were purified to homogeneity, biochemically characterized, and applied in a series of transglycosylation reactions to demonstrate their high synthetic potential. The enhanced selectivity of the prepared mutants, in combination with their synthetic potential, afford promising synthetic tools in the production of complex β-*N*-acetylhexosamines combining both GlcNAc and GalNAc units. These compounds are otherwise impossible to prepare by wild-type glycosidases.

Experimental Section

Materials

*p*NP-GlcNAc (**1**) and *p*NP-GalNAc (**2**) were obtained from Gold Biotechnology, USA. GlcNAc was obtained from Acros Organics, USA. Vinyl acetate was supplied by Sigma-Aldrich, Czech Republic, and Novozym 435 by Novozymes (Denmark). If not stated otherwise, all other chemicals were from VWR Chemicals or Lach-Ner (Czech Republic).

Preparation and Cloning of Mutant Genes

The pPICZαA expression constructs carrying the genes of mutant *Tf*Hex (Glu546Tyr, Glu546Gln, Glu546His, Arg218His, Arg218Lys *Tf*Hex) were prepared commercially (Generay, China). The genes were cloned downstream of the α-factor-encoding DNA segment for extracellular protein targeting and zeocin resistance using *Eco*RI and *Kpn*I restriction sites. The construct containing the gene of Arg218Gln *Tf*Hex was prepared by site-directed mutagenesis using *Tf*Hex WT[17] construct as a template (GenBank ID: JN601495), and the pair of primers (Generi Biotech, Czech Republic): Fw: 5'-gattgacacgggccaaaactttattactgt-3'; and
Re: 5'-acagtaataaagttttggcccgtgtcaatc-3'.

The PCR reaction was performed using NZYProof DNA polymerase (NZYTech, Portugal) in a T-Personal Thermal Cycler (Biometra, Germany) under the following conditions: initial denaturation at 95 °C for 1 min, followed by 35 cycles of denaturation at 95 °C for 30 s, primer annealing at 59 °C for 1 min, elongation at 72 °C for 6 min and final elongation at 72 °C for 10 min.

Then, 1 µL of *Dpn*I (New England Biolabs, USA) was added into the PCR mixture (50 µL), and the reaction mixture was incubated at 37 °C and 300 rpm overnight to cleave the methylated template. The PCR product was then transformed into *Escherichia coli* Top10 competent cells (Thermo Fisher Scientific, USA) by heat shock method, under the selection pressure of zeocin (100 µg mL^{-1}). The plasmids isolated from sixteen selected *E. coli* colonies using High Pure Plasmid Isolation Kit (Roche, Switzerland) were sequenced using the sequencing primer 5'-gactggttccaattgacaagc-3'. The plasmids containing the gene of Arg218Gln *Tf*Hex were isolated at a semi-preparative

scale and purified using the High Pure Plasmid Midi Isolation kit (Roche, DE) for transformation into *P. pastoris* as described further.

Transformation of Mutant Genes

The yeast expression vectors pPICZαA carrying individual genes of six *Tf*Hex mutant variants (Glu546Tyr, Glu546Gln, Glu546Tyr, Arg218His, Arg218Lys, Arg218Gln *Tf*Hex) were linearized (15 µg) using restriction endonuclease *Sac*I (New England Biolabs, USA). The vectors were then electroporated into *Pichia pastoris* KM71H competent cells (Invitrogen, Life Technologies, USA) using the standard protocol by Invitrogen (EasySelect Pichia Expression Kit; Invitrogen, Life Technologies, USA). The protein expression and enzyme activity were screened at small-scale production as described in the next section. Two to three colonies showing a good production of mutant enzyme were cryopreserved for long-term storage at -80 °C in 15% (*v/v*) glycerol.

Heterologous Expression of β-*N*-Acetylhexosaminidase in *Pichia pastoris*

Mutant *Tf*Hex variants were produced basically according to the manufacturer's instructions (EasySelect *Pichia* Expression Kit; Invitrogen, Life Technologies, USA). For a screening of the enzyme production in *Pichia pastoris* at a small-scale, we used a combination of BMGY medium (Buffered Glycerol complex Medium) and BMMY medium (Buffered Methanol complex Medium). The transformants (sixteen selected colonies from electroporation) were inoculated into 100 mL of BMGY medium and incubated at 28 °C and 220 rpm overnight. Cultures with grown cells were centrifuged (5000 rpm, 10 min, 12 °C) and the pellets were resuspended in 30 mL of BMMY medium, where the expression of the recombinant protein was induced by methanol (0.5 % *v/v*). The cultures were shaken at 28 °C and 220 rpm for three days. Another batch of methanol was added every 24 hours. On day 5, the cultures were tested for the presence of *Tf*Hex by SDS-PAGE and for enzyme activity.

Preparative production was performed in a combination of BMGH medium (Buffered Glycerol Minimal Medium) and BMMH medium (Buffered Methanol Minimal Medium). The cryopreserved cultures (100 µL) were inoculated into YPD medium (15 mL; 1 % yeast extract, 2 % peptone, 2 % D-glucose) and incubated at 28 °C and 220 rpm for 5 hours. This preculture was inoculated into 1 L of BMGH medium in 3 L Erlenmeyer flasks. On day 2, the cells were collected by centrifugation (5000 rpm, 10 min, and 12 °C) and resuspended in BMMH medium (200 mL) in 1 L Erlenmeyer flask. The culture was shaken at 28 °C and 220 rpm, the expression of the enzyme was induced by methanol (0.5 % *v/v*) every 24 h.

On day 5, the culture was centrifuged (5000 rpm, 10 min, 12 °C) and supernatant was purified to homogeneity by cation exchange chromatography (Fractogel EMD SO_3^-, Merck, D) on an Äkta Purifier system (Amersham Biosciences, Uppsala, Sweden). The column was equilibrated with 10 mM sodium citrate-phosphate buffer pH 3.5. The elution of the enzyme was performed with a linear gradient of 0-1 M NaCl (60 mL, 2 mL min^{-1}). The purity of the enzyme was verified by SDS-PAGE, the enzyme was concentrated by ultrafiltration in McIlvaine buffer (50 mM citric acid/100 mM Na_2HPO_4, pH 5.0), and stored at 4 °C.

Protein Characterization

The purity of native and deglycosylated *Tf*Hex mutants was verified by SDS-PAGE in 10% polyacrylamide gel. The protein concentration was determined by Bradford method[35] using Protein Assay Dye Reagent Concentrate (Bio-Rad, UK). Calibration was performed for bovine plasma γ-globulin (Bio-Rad, UK). Deglycosylation of *Tf*Hex variants was performed according to the manufacturer´s protocol using endo-glycosidase H (EndoH; New England Biolabs, USA). Each of the *Tf*Hex variants (20 µg) was combined with 10× Glycoprotein Denaturating Buffer (1 µl) supplied by the manufacturer and diluted with water to a total reaction volume of 10 µL. Then the protein was denatured by heating (100 °C; 10 min), mixed with 10× GlycoBuffer 3 (2 µL), EndoH (2 µL; 1000 U) and water up to a final volume of 20 µL and incubated overnight (37 °C, 300 rpm).

Standard Assay for Enzyme Activity

The determination of the hydrolytic activity of β-*N*-acetylhexosaminidases was performed using an end-point assay with spectrophotometric detection. The substrates *p*-nitrophenyl 2-acetamido-2-deoxy-β-D-glucopyranoside (**1**; *p*NP-GlcNAc) and *p*-nitrophenyl 2-acetamido-2-deoxy-β-D-galactopyranoside (**2**; *p*NP-GalNAc) were used. The reaction mixture in microtubes containing 2 mM *p*NP-GlcNAc (**1**) or *p*NP-GalNAc (**2**) substrate in McIlvaine buffer (50 mM citric acid/100 mM Na_2HPO_4, pH 5.0) was incubated at 35 °C in a thermoshaker, and reaction was started by adding 10 µL of diluted enzyme. The reaction ran for 10 min at 850 rpm and was quenched by adding 1 mL of 0.1 M sodium carbonate. The liberated *p*-nitrophenol, cleaved off the respective substrate, occurs in the basic medium in the form of the *p*-nitrophenolate anion that is yellow and absorbance can be measured at 420 nm. The measurement was performed in triplicate against a reference sample containing only substrate and buffer. One unit of enzymatic activity was defined as the amount of enzyme that cleaved 1 µmol of the respective substrate (GlcNAcase activity for cleaving *p*NP-GlcNAc, GalNAcase activity for cleaving *p*NP-GalNAc) per minute under the above conditions.

The activity of *Tf*Hex WT and mutant Glu546Gln *Tf*Hex with the protected donor *p*-nitrophenyl 6-*O*-acetyl-2-acetamido-2-deoxy-β-D-glucopyranoside (**8**) was performed analogously to the standard assay determination as described above and was evaluated as a per cent value related to the standard substrate *p*NP-GlcNAc (**1**); see also Table S10 in the Supporting Information.

The pH optimum of *Tf*Hex was determined in the universal Britton-Robinson buffer (0.04 M H_3PO_4, 0.04 M phenylacetic acid, 0.04 M H_3BO_3/ 0.2 M NaOH), which is used in the pH range of 2 to 12. The same buffer was used for respective enzyme dilutions. The enzyme activity was determined after 10 min at 35 °C and 850 rpm.

The temperature optimum was determined in McIlvaine buffer (50 mM citric acid/100 mM Na_2HPO_4, pH 5.0) in 5 °C intervals from 20 °C to 80 °C or until the enzyme was completely inactivated.

Michaelis-Menten parameters of hydrolysis of substrates **1** and **2** (0.1-3.0 mM of **2** for both GalNAcases, 0.1-7.0 mM of **1** for Arg218His *Tf*Hex and 0.1-4.5 mM of **1** for Arg218Lys *Tf*Hex) were measured by discontinuous kinetic assay. With fungal β-*N*-acetylhexosaminidases, which are generally acidophilic, the continuous assay shows a low sensitivity due to the low proportion of the ionized *p*-nitrophenolate in the reaction mixture The reaction mixture (400 µL), consisting of substrate **1** or **2**, McIlvaine buffer pH 5.0 and *Tf*Hex mutant enzyme, was incubated at 25 °C and 850 rpm for 7 min. Aliquots (50 µL) were taken every 1 min and the reaction was stopped by mixing with 1 M sodium carbonate (100 µL) in a microplate. Then absorbance at 420 nm was determined by Sunrise Microplate Reader (Tecan, Switzerland). All data were measured in triplicate. The kinetic parameters (K_M, k_{cat}) were extracted by non-linear regression using GraphPad Prism 7 (GraphPad Software, USA). Respective fitting curves are depicted in Figure S5 in the Supporting Information.

Molecular Modeling, Docking, and Molecular Dynamics

The molecular modeling of mutant variants was accomplished starting from the molecular model of *Tf*Hex WT[27] and modified according to the resolved structure of *Ao*Hex.[28] Briefly,[36] loops HL1 (residues 365-373), HL2 (residues 416-508) and the N-terminal propeptide were remodeled by restrained modeling in MODELLER 9.16[37] and minimized in YASARA[36] by standard protocol using YASARA2 force field in TIP3P water. The dimer was constructed by structural alignment in YASARA[36] based on the crystal of β-*N*-acetylhexosaminidase from *A. oryzae*[15] – (pdb code: 5oar). This model was used for docking and molecular dynamics simulations previously.[11, 36]

The mutation hotspots (Glu546 and Arg218) were identified on the basis of the difference between interactions of the WT enzyme with *p*NP-GalNAc and *p*NP-GlcNAc. In order to predict prospective mutations at these positions, we used *in silico* saturation mutagenesis. Both residues were substituted with all possible amino acids and the change in the binding energy for both *p*NP-GalNAc and *p*NP-GlcNAc substrates upon the respective mutation was calculated in YASARA (see the Supporting Information, Table S1). Before energy calculation, the models of the mutants were minimized in YASARA in water (by standard protocol, YASARA2 force field, with fixed residues further than 0.8 nm from the mutated residue). The binding energy was calculated by YASARA macro as in the MM/PBSA method (without the entropy cost) as the energy of separated compounds minus energy of their complex. As a parameter for selecting the prospective mutant variants we used the difference between the binding energies (similar to Δ*G*) of *p*NP-GalNAc and *p*NP-GlcNAc. Lower values correspond to the increase in affinity to *p*NP-GalNAc over *p*NP-GlcNAc compared to the WT. This pre-selection was followed by analyzing the substrate interaction with the theoretical mutants including the ability of glycan binding, product release and possible sterical conflicts in the active site, which resulted in the final selection of mutations.

Single point mutations were introduced in YASARA.[38] Protonation states of active-site residues was assigned based on Delphi[39] and Propka[40] (from Schrödinger Suite 2018-4, Build 12) prediction at pH 7; protonation state for catalytic residues was assigned manually. The conformation of Gln residue in mutants was selected based on the analysis of rotameric library during initial docking by Glide. More favorable conformation for substrate binding was selected. The models of mutant variants were minimized in YASARA to allow the exchanged residue to acquire favorable orientation and to optimize contacts with neighboring residues. Constrained minimization was done *in vacuo* using NOVA force field,[38] the last minimization run was done in TIP3P water model, periodic boundary condition, YASARA2 force field,[41] NPT ensemble. Substrate structures were built in YASARA and used for docking with Glide (part of Schrödinger suite).[40, 42-43]

Molecular dynamics simulations were run using the substrate-enzyme complexes minimized in YASARA; molecular dynamics was done in explicit TIP3P water at a temperature of 298 K and pH 7 for 20 ns, using YASARA2 force field. Sodium ions were used for system neutralization, Particle Mesh Ewald algorithm[41] was employed for long-range interaction modeling, and Berendsen thermostat was applied for temperature control. Molecular dynamics simulations were analyzed with YASARA tools.[38] The number of hydrogen bonds formed between substrate and enzyme was calculated for a stable period of molecular dynamics (2-20 ns), defined from root-mean square deviations of the Cα atoms of proteins (data not shown). The probability of hydrogen bond formation during molecular dynamics was calculated as a number of snapshots with formed hydrogen bonding *versus* total number of snapshots and expressed as per cent value. The complexes for calculation of docking scores were selected from equilibrated molecular dynamics simulations (2-20 ns) with the best parameters for hydrolysis of the glycosidic bond, namely the smallest distance between Glu371 (OE2 atom) and the substrate glycosidic oxygen, and the highest number of hydrogen bonds. Five selected structures from molecular dynamics simulation of each enzyme-substrate complex were used for short minimization with Protein Preparation Wizard (part of Schrödinger suite[40]) with OPLS3 force field[42] *in vacuo*. Minimized structures were used for rigid docking to calculate Glide XP binding scores.[43] Final reported binding score is calculated as an average value.

Analytical Methods

HPLC Analysis: HPLC analysis was used to monitor the course of transglycosylation reactions. The Shimadzu Prominence LC analytical system comprised Shimadzu CBM-20A system controller, Shimadzu LC-20AD binary HPLC pump, Shimadzu CTO-10AS column oven, Shimadzu SIL-20ACHT cooling autosampler, and Shimadzu SPD-20MA diode array detector (Shimadzu, JP). Analyses were performed on a TSK gel Amide-80 column (250 × 4.6 mm, 5 µm) preceded by TSKgel Amide-80 Guardgel (3.2 × 15 mm, Tosoh corp., Japan) in acetonitrile/water (4/1, *v*/*v*), with gradient elution as follows (A = acetonitrile, B = water): 22 % B for 0-7 min, 22-31 % B for

7-16 min, 31-22 % B for 16-17 min, and 22 % B for 17-22 min for column equilibration; flow rate of 1 mL min^{-1} at 25 °C; injection volume 1 µL: Detection was performed at 200 nm.

ESI-MS Analysis: Mass spectra were measured using LTQ Orbitrap XL hybrid mass spectrometer (Thermo Fisher Scientific, USA) equipped with an electrospray ion source. The mobile phase was methanol/water (4:1, *v*/*v*) at a flow rate of 100 µL min^{-1}. The samples were dissolved in methanol or methanol/ water and injected using a 5-µL loop into the mobile phase flow. For the negative ion mode, spray voltage, capillary voltage, tube lens voltage, and capillary temperature were 5.0 kV, -25 V, -125 V, and 275 °C, respectively. For the positive ion mode, spray voltage, capillary voltage, tube lens voltage, and capillary temperature were 5.0 kV, 9 V, 150 V, and 275 °C, respectively. The spectra were recorded at a resolution of 100,000.

NMR Analysis*:* NMR spectra were acquired on a Bruker Avance III 400 MHz (compound **8**), 600 MHz (compounds **4**, **10**, **11**, and **12**), and 700 MHz (compounds **5**, **6**, and **7**) spectrometers (Bruker BioSpin, Rheinstetten, Germany) in D_2O (100 atom % D, Sigma-Aldrich, Steinheim, Germany) at 30 °C. Residual signal of D_2O (δ_H 4.732 ppm) was used for the proton spectra reference; carbon spectra were referenced to the signal of acetone (δ_C 30.50 ppm). The individual monosaccharide units were assigned using COSY, HSQC, 1d-TOCSY, and HSQC-TOCSY experiments; their acetylation was proved by HMBC experiment. Glycosidic linkage (1→4) was confirmed by the HMBC correlation of carbon C4 with the adjacent anomeric proton.

Optical Rotation: Specific rotation of prepared compounds was measured with an Autopol IV polarimeter (Rudolph Research Analytical, USA, 2001). The $[\alpha]_D^{20}$ values were recorded in water, water/ acetonitrile or DMSO according to the solubility of compounds as specified further; the concentration units are g mL^{-1}.

Preparative Synthetic Reactions

***2-Acetamido-2-deoxy-β-D-glucopyranosyl- (1→4)-2-acetamido-2-deoxy-β-D-glucopyranosyl azide (4)*:** GlcNAc-N_3 (**3**, 183 mg, 0.7 mmol) and Arg218Lys *Tf*Hex (6.5 µL, 4.5 U, 0.2 mg) were incubated in McIlvaine buffer pH 5 (total reaction volume 1.5 mL) at 35 °C and 850 rpm. The reaction was monitored by TLC (propane-2-ol/water/NH_4OH aq., 7/2/1) and HPLC. After 5 h the reaction was stopped by boiling for 2 min. The reaction mixture was centrifuged (13,500 rpm, 10 min) and the supernatant was separated by gel permeation chromatography using Biogel P-2 column (26 × 1000 mm, Bio-Rad, UK) in water as a mobile phase at an elution rate of 7.2 mL h^{-1}, ambient temperature. The fractions containing the product were collected and lyophilized. The title compound **4** was obtained as a white solid (24 mg, 53.4 µmol, 27 % isolated yield). For structural data of compound **4,** see the Supporting information (Table S2, Figure S11). $[\alpha]_D^{20}$ = −20.6 (c 0.325, H_2O).

pNP-Functionalized Chitooligomers 5-7: *p*NP-GlcNAc (**1**, 390 mg, 1.1 mmol) and *Tf*Hex Arg218Lys (13 µL, 6 U, 0.2 mg) were incubated in 30% acetonitrile in McIlvaine buffer pH 5 (total reaction volume 6 mL) at 35 °C and 1000 rpm. After 2 h another portion of substrate **1** (180 mg, 0.5 mmol) was added. The reaction progress was monitored by TLC (propane-2-ol/water/NH_4OH aq., 7/2/1) and by HPLC. The reaction was stopped after 5.5 h by boiling for 2 min. Denatured enzyme was removed by centrifugation (13,500 rpm, 10 min) and the supernatant was loaded onto Biogel P-2 column (26 × 1000 mm, Bio-Rad, UK) with water as a mobile phase at an elution rate of 7.3 mL h^{-1}, ambient temperature. Fractions containing the products were collected and lyophilized. Traces of *p*-nitrophenol were removed by extraction into ethyl acetate. The fractions of products **5-7** still contained traces of non-substituted oligosaccharides, which were removed by extraction on Amberlite XAD-4 resin (BDH Chemicals, Ltd, UK). The samples were dissolved in water, loaded onto XAD-resin and after washing with water, pure *p*NP chitooligomers were eluted with MeOH. Compounds **5-7** were obtained as white solids: *p*-nitrophenyl 2-acetamido-2-deoxy-β-D-glucopyranosyl- (1→4)-2-acetamido-2-deoxy-β-D-glucopyranoside (**5**; 20 mg, 37 µmol), *p*-nitrophenyl 2-acetamido-2-deoxy-β-D-glucopyranosyl- (1→4)-2-acetamido-2-deoxy-β-D-glucopyranosyl- (1→4)-2-acetamido-2-deoxy-β-D-glucopyranoside (**6**; 17 mg, 23 µmol) and *p*-nitrophenyl 2-acetamido-2-deoxy-β-D-glucopyranosyl- (1→4)- 2-acetamido-2-deoxy-β-D-glucopyranosyl- (1→4)-2-acetamido-2-deoxy-β-D-glucopyranosyl- (1→4)-2-acetamido-2-deoxy-β-D-glucopyranoside (**7**; 16 mg, 16 µmol). For the structural analysis, see the Supporting information (Tables S3-5, Figures S12-14). For compound **5**, $[\alpha]_D^{20}$ = −20.7 (c 0.237, H_2O); for compound **6**, $[\alpha]_D^{20}$ = −20.2 (c 0.233, H_2O); for compound **7**, $[\alpha]_D^{20}$ = −15.4 (c 0.136, H_2O/CH_3CN, 3/2).

p-Nitrophenyl 6-O-acetyl-2-acetamido-2-deoxy-β-D-glucopyranoside (8): The selective protection of the C6 hydroxyl was carried out according to Simerská *et al*.[34] *p*NP-GlcNAc (**1**, 200 mg, 0.6 mmol) was dissolved in pyridine (15 mL) and then, acetone (21 mL) was added, followed by vinyl acetate (7.1 mL) and a catalytic amount of Novozym 435 (lipase from *Candida antarctica* immobilized on acrylic resin). The reaction mixture was filtered through a filter paper and co-evaporated with toluene *in vacuo* to dryness. The title compound **8** was obtained as a light brown solid (230 mg, 0.6 mmol, 99 % isolated yield). For the structural analysis, see the Supporting information (Table S6, Figure S15). $[\alpha]_D^{20}$ = −29.8 (c 0.225, DMSO).

6'-O-Acetyl-2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy-D-glucopyranose (10): Compound **8** (48 mg, 0.1 mmol) as a donor, GlcNAc (**9**, 444 mg, 2.0 mmol) as an acceptor and Arg218Lys *Tf*Hex (98 µL, 40 U, 2.7 mg) were incubated in 30% acetonitrile in McIlvaine buffer pH 5 (total reaction volume 4 mL) at 35 °C and 850 rpm. The progress of the reaction was monitored by TLC (propane-2-ol/water/NH_4OH aq., 7/2/1) and HPLC. The reaction was stopped after 5 h by boiling for 2 min. The mixture was centrifuged (13,500 rpm, 10 min) and the supernatant was separated by size exclusion chromatography using Biogel P-2 column (26 × 1000 mm,

Bio-Rad, UK) and water as a mobile phase at an elution rate of 7.3 mL h^{-1}, ambient temperature. The fractions containing the product were combined and lyophilized. The title compound **10** was obtained as a white solid (14 mg, 30 µmol, 23 % isolated yield). For the structural analysis, see the Supporting information (Table S7, Figure S16). $[\alpha]_D^{20}$ = +11.1 (c 0.215, H_2O).

2-Acetamido-2-deoxy-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranosyl-(1→4)-2-acetamido-2-deoxy-D-glucopyranose (11): *p*NP-GalNAc (**2**, 12 mg, 35 µmol) as a donor, compound **10** (52 mg, 0.1 mmol) as an acceptor and Glu546Gln *Tf*Hex (18 µL, 2.8 U, 0.7 mg) were incubated in McIlvaine buffer pH 5 (total reaction volume 1.1 mL) at 35 °C and 850 rpm. The reaction progress was monitored by TLC (propane-2-ol/water/NH_4OH aq., 7/2/1) and HPLC. The reaction was stopped after 8 hours by boiling for 2 min. The denatured enzyme was removed by centrifugation (13,500 rpm, 10 min) and the supernatant was separated by size exclusion chromatography. Biogel P-2 (26 × 1000 mm; Bio-Rad, UK) was used as a stationary phase, and water as a mobile phase at an elution rate of 7.5 mL h^{-1}. The 6´-*O*-acetylated compound **11** was obtained as a white solid (6.6 mg, 10 µmol, 28 % isolated yield). For the structural analysis, see the Supporting information (Table S8, Figure S17). $[\alpha]_D^{20}$ = +20.9 (c 0.215, H_2O). In the next step, it was deprotected by ammonium hydroxide in methanol (1:10) to afford the title compound **12** as a white solid (6.2 mg, 10 µmol, 99 % isolated yield). For the structural analysis, see the Supporting information (Table S9, Figure S18).

Acknowledgments

Support by the project 20-00215S by the Czech Science Foundation and by the mobility projects LTC18038 and LTC19035 by the Ministry of Education, Youth and Sports of the Czech Republic and is gratefully acknowledged. Computational resources are provided by the project "e-Infrastruktura CZ" (e-INFRA LM2018140), a part of the program Projects of Large Research, Development and Innovations Infrastructures. We thank the group of Dr. Š. Matějková, Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, for determining specific rotations.

References

[1] K. Slámová, P. Bojarová, L. Petrásková, V. Křen, *Biotechnol. Adv.* **2010**, *28*, 682-693.
[2] P. Bojarová, J. Bruthans, V. Křen, *Appl. Microbiol. Biotechnol.* **2019**, *103*, 7869-7881.
[3] P. Fialová, L. Weignerová, J. Rauvolfová, V. Přikrylová, A. Pišvejcová, R. Ettrich, M. Kuzma, P. Sedmera, V. Křen, *Tetrahedron* **2004**, *60*, 693-701.
[4] J. Rauvolfová, M. Kuzma, L. Weignerová, P. Fialová, V. Přikrylová, A. Pišvejcová, M. Macková, V. r. Křen, *J. Mol. Catal. B: Enzym.* **2004**, *29*, 233-239.
[5] B. L. Mark, M. N. G. James, *Can. J. Chem.* **2002**, *80*, 1064-1074.
[6] B. L. Mark, D. J. Vocadlo, S. Knapp, B. L. Triggs-Raine, S. G. Withers, M. N. G. James, *J. Biol. Chem.* **2001**, *276*, 10330-10337.
[7] K. Slámová, R. Gažák, P. Bojarová, N. Kulik, R. Ettrich, H. Pelantová, P. Sedmera, V. Křen, *Glycobiology* **2010**, *20*, 1002-1009.
[8] P. Bojarová, K. Slámová, K. Křenek, R. Gažák, N. Kulik, R. Ettrich, H. Pelantová, M. Kuzma, S. Riva, D. Adámek, K. Bezouška, V. Křen, *Adv. Synth. Catal.* **2011**, *353*, 2409-2420.
[9] P. Bojarová, K. Křenek, M. Kuzma, L. Petrásková, K. Bezouška, D.-J. Namdjou, L. Elling, V. Křen, *J. Mol. Catal. B: Enzym.* **2008**, *50*, 69-73.
[10] K. Slámová, J. Krejzová, P. Marhol, L. Kalachova, N. Kulik, H. Pelantová, J. Cvačka, V. Křen, *Adv. Synth. Catal.* **2015**, *357*, 1941-1950.
[11] P. Bojarová, N. Kulik, M. Hovorková, K. Slámová, H. Pelantová, V. Křen, *Molecules* **2019**, *24*, 599.
[12] P. Bojarová, N. Kulik, K. Slámová, M. Hubálek, M. Kotik, J. Cvačka, H. Pelantová, V. Křen, *Appl. Microbiol. Biotechnol.* **2019**, *103*, 1737-1753.
[13] L. Weignerová, P. Vavrušková, A. Pišvejcová, J. Thiem, V. r. Křen, *Carbohydr. Res.* **2003**, *338*, 1003-1008.
[14] P. Nekvasilová, I. Andreasová, L. Petrásková, H. Pelantová, V. Křen, P. Bojarová, *BBA - Proteins Proteom.* **2020**, *1868*, 140319.
[15] R. Ettrich, V. Kopecký, K. Hofbauerová, V. Baumruk, P. Novák, P. Pompach, P. Man, O. Plíhal, M. Kutý, N. Kulik, J. Sklenář, H. Ryšlavá, V. Křen, K. Bezouška, *BMC Struct. Biol.* **2007**, *7*, 32.
[16] H. Ryšlavá, A. Kalendová, V. Doubnerová, P. Skočdopol, V. Kumar, Z. Kukačka, P. Pompach, O. Vaněk, K. Slámová, P. Bojarová, N. Kulik, R. Ettrich, V. Křen, K. Bezouška, *FEBS J.* **2011**, *278*, 2469-2484.
[17] K. Slámová, P. Bojarová, D. Gerstorferová, B. Fliedrová, J. Hofmeisterová, M. Fiala, P. Pompach, V. Křen, *Protein Expression Purif.* **2012**, *82*, 212-217.
[18] K. Furukawa, K. Takamiya, K. Furukawa, *BBA - Gen. Subjects* **2002**, *1573*, 356-362.
[19] P. L. Deangelis, A. J. Padgett-Mccue, *J. Biol. Chem.* **2000**, *275*, 24124-24129.
[20] A. Šimonová, C. E. Kupper, S. Böcker, A. Müller, K. Hofbauerová, H. Pelantová, L. Elling, V. Křen, P. Bojarová, *J. Mol. Catal. B: Enzym.* **2014**, *101*, 47-55.
[21] T. Sumida, K. Fujimoto, M. Ito, *J. Biol. Chem.* **2011**, *286*, 14065-14072.
[22] I. Noach, B. Pluvinage, C. Laurie, K. T. Abe, M. G. Alteen, D. J. Vocadlo, A. B. Boraston, *J. Mol. Biol.* **2016**, *428*, 3253-3265.
[23] C. Roth, M. Petricevic, A. John, E. D. Goddard-Borger, G. J. Davies, S. J. Williams, *Chem. Commun.* **2016**, *52*, 11096-11099.
[24] K. Slámová, P. Bojarová, *BBA - Gen. Subjects* **2017**, *1861*, 2070-2087.
[25] M. S. Macauley, G. E. Whitworth, A. W. Debowski, D. Chin, D. J. Vocadlo, *J. Biol. Chem.* **2005**, *280*, 25313-25322.

[26] J. Krejzová, P. Šimon, L. Kalachova, N. Kulik, P. Bojarová, P. Marhol, H. Pelantová, J. Cvačka, R. Ettrich, K. Slámová, V. Křen, *Molecules* **2014**, *19*.
[27] N. Kulik, K. Slámová, R. Ettrich, V. Křen, *BMC Bioinformatics* **2015**, *16*, 28.
[28] J. Škerlová, J. Bláha, P. Pachl, K. Hofbauerová, Z. Kukačka, P. Man, P. Pompach, P. Novák, Z. Otwinowski, J. Brynda, O. Vaněk, P. Řezáčová, *FEBS J.* **2018**, *285*, 580-598.
[29] Y.-W. Kim, H. Chen, S. G. Withers, *Carbohydr. Res.* **2005**, *340*, 2735-2741.
[30] S. J. Williams, B. L. Mark, D. J. Vocadlo, M. N. G. James, S. G. Withers, *J. Biol. Chem.* **2002**, *277*, 40055-40065.
[31] E. S. Feldblum, I. T. Arkin, *Proc. Natl. Acad. Sci.* **2014**, *111*, 4085.
[32] P. Fialová, A. T. Carmona, I. Robina, R. Ettrich, P. Sedmera, V. Přikrylová, L. Petrásková-Hušáková, V. Křen, *Tetrahedron Lett.* **2005**, *46*, 8715-8718.
[33] D. Laaf, P. Bojarová, B. Mikulová, H. Pelantová, V. Křen, L. Elling, *Adv. Synth. Catal.* **2017**, *359*, 2101-2108.
[34] P. Simerská, A. Pišvejcová, M. Kuzma, P. Sedmera, V. r. Křen, S. Nicotra, S. Riva, *J. Mol. Catal. B: Enzym.* **2004**, *29*, 219-225.
[35] M. M. Bradford, *Anal. Biochem.* **1976**, *72*, 248-254.
[36] C. Garcia-Oliva, P. Hoyos, L. Petrásková, N. Kulik, H. Pelantová, A. H. Cabanillas, Á. Rumbero, V. Křen, M. J. Hernáiz, P. Bojarová, *Int. J. Mol. Sci.* **2019**, *20*, 6181.
[37] A. Šali, T. L. Blundell, *J. Mol. Biol.* **1993**, *234*, 779-815.
[38] E. Krieger, G. Koraimann, G. Vriend, *Proteins: Struct., Funct., Bioinf.* **2002**, *47*, 393-402.
[39] L. Wang, M. Zhang, E. Alexov, *Bioinformatics* **2015**, *32*, 614-615.
[40] G. Madhavi Sastry, M. Adzhigirey, T. Day, R. Annabhimoju, W. Sherman, *J. Comput. Aided Mol. Des.* **2013**, *27*, 221-234.
[41] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577-8593.
[42] E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight, J. W. Kaus, D. S. Cerutti, G. Krilov, W. L. Jorgensen, R. Abel, R. A. Friesner, *J. Chem. Theory Comput.* **2016**, *12*, 281-296.
[43] R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin, D. T. Mainz, *J. Med. Chem.* **2006**, *49*, 6177-6196.

10.1002/adsc.202000604

FULL PAPER

How Site-Directed Mutagenesis Boosted Selectivity of a Promiscuous Enzyme.

Adv. Synth. Catal. **Year**, *Volume*, Page – Page

Pavlína Nekvasilová, Natalia Kulik, Nikola Rychlá, Helena Pelantová, Lucie Petrásková, Zuzana Bosáková, Josef Cvačka, Kristýna Slámová, Vladimír Křen, Pavla Bojarová*