Applied Microbiology and Biotechnology
https://doi.org/10.1007/s00253-018-9534-z

BIOTECHNOLOGICALLY RELEVANT ENZYMES AND PROTEINS

Selective β-*N*-acetylhexosaminidase from *Aspergillus versicolor*—a tool for producing bioactive carbohydrates

Pavla Bojarová[1] · Natallia Kulik[2] · Kristýna Slámová[1] · Martin Hubálek[3] · Michael Kotik[1] · Josef Cvačka[3] · Helena Pelantová[1] · Vladimír Křen[1]

Received: 7 September 2018 / Revised: 5 November 2018 / Accepted: 17 November 2018

Abstract

β-*N*-Acetylhexosaminidases (EC 3.2.1.52) are typical of their dual activity encompassing both *N*-acetylglucosamine and *N*-acetylgalactosamine substrates. Here we present the isolation and characterization of a selective β-*N*-acetylhexosaminidase from the fungal strain of *Aspergillus versicolor*. The enzyme was recombinantly expressed in *Pichia pastoris* KM71H in a high yield and purified in a single step using anion-exchange chromatography. Homologous molecular modeling of this enzyme identified crucial differences in the enzyme active site that may be responsible for its high selectivity for *N*-acetylglucosamine substrates compared to fungal β-*N*-acetylhexosaminidases from other sources. The enzyme was used in a sequential reaction together with a mutant β-*N*-acetylhexosaminidase from *Talaromyces flavus* with an enhanced synthetic capability, affording a bioactive disaccharide bearing an azido functional group. The azido function enabled an elegant multivalent presentation of this disaccharide on an aromatic carrier. The resulting model glycoconjugate is applicable as a selective ligand of galectin-3 — a biomedically attractive human lectin. These results highlight the importance of a general availability of robust and well-defined carbohydrate-active enzymes with tailored catalytic properties for biotechnological and biomedical applications.

Keywords *Aspergillus versicolor* · β-*N*-Acetylhexosaminidase · Glycosidase · Homology modeling · Heterologous expression · *Pichia pastoris*

Introduction

Oligosaccharides composed of *N*-acetylglucosamine and *N*-acetylgalactosamine units are bioactive molecules of abundant biological activities (Slámová and Bojarová 2017). They are attractive ligands for targeting lectins — ubiquitous carbohydrate-binding proteins that read the glycocode on surfaces of living cells. As a result, synthetic conjugates decorated with specific *N*-acetylhexosamine ligands (Bojarová et al. 2017; Laaf et al. 2017a, b) may serve as potent agents in biomedical or biotechnological applications.

Glycosidases (*O*-glycoside hydrolases; EC 3.2.1.-) are destined to the in vivo cleavage of oligo- and polysaccharides by glycosyl transfer to water. They can form a glycosidic linkage if a hydroxyl moiety of a carbohydrate acceptor acts as a more efficient nucleophile than water. Such conditions can be achieved by a variety of strategies including reduction of water activity and use of glycosyl donors activated by a good leaving group. Through a kinetically controlled transglycosylation in an optimized setup, close-to-quantitative yields may be accomplished (Fialová et al. 2004). Enzymatic synthesis catalyzed by glycosidases is a biocompatible elegant alternative to the classical organic chemistry methods, which are burdened by a number of protection and deprotection steps. It offers absolute stereoselectivity, high regioselectivity, and glycosidic bond formation in one step. In contrast to glycosyltransferases (EC 2.4)

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00253-018-9534-z) contains supplementary material, which is available to authorized users.

✉ Pavla Bojarová
bojarova@biomed.cas.cz

[1] Laboratory of Biotransformation, Institute of Microbiology, Czech Academy of Sciences, Vídeňská 1083, CZ-14220 Prague 4, Czech Republic

[2] Laboratory of Structure and Function of Proteins, Institute of Microbiology, Czech Academy of Sciences, Zámek 136, CZ-37333 Nové Hrady, Czech Republic

[3] Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, Flemingovo náměstí 2, CZ-16610 Prague 6, Czech Republic

Published online: 02 January 2019

(Bojarová et al. 2013), glycosidases tolerate the presence of a wide choice of functional groups in the substrate molecule (Bojarová et al. 2008; Bojarová et al. 2011), and thus enable to efficiently prepare tailored oligosaccharides as a part of a chemoenzymatic synthetic pathway towards, e.g., bioactive glycomimetics.

β-*N*-Acetylhexosaminidases (EC 3.2.1.52) are exoglycosidases (glycoside hydrolase family GH20, http://www.cazy.org/) with a unique dual activity for cleaving both *N*-acetylglucosamine (GlcNAc) and *N*-acetylgalactosamine (GalNAc) carbohydrate moieties in a variety of substrates, such as chitooligomers, glycosphingolipids, glycoproteins, and other glycoconjugates (Slámová et al. 2010). Although GlcNAc and GalNAc units differ only in the equatorial versus axial hydroxyl at C-4, their chemical behavior and biological implications are vastly different. The dual activity of β-*N*-acetylhexosaminidases may be an obstacle in a selective synthesis of defined *N*-acetylhexosamine oligosaccharides, especially those carrying a combination of GlcNAc and GalNAc units. Though there is a separate family of *O*-GlcNAcases (GH84) (Gloster and Vocadlo 2010), these enzymes are specific for the cleavage of *O*-GlcNAc residues from proteins and their affinity to oligosaccharide substrates is very low. Moreover, these enzymes are not applicable for synthetic purposes as they do not exhibit any transglycosylation capabilities. Although specific glycosyltransferases exist for a selective introduction of GlcNAc moieties (Slámová and Bojarová 2017), these enzymes lack the substrate promiscuity and synthetic versatility of especially fungal glycosidases, which has been demonstrated in numerous previous applications (Bojarová and Křen 2009). In this article, we present the isolation and characterization of a selective β-*N*-acetylhexosaminidase from *Aspergillus versicolor* (*Av*Hex). The enzyme was recombinantly expressed in *Pichia pastoris* in a high yield and purified by one-step anion-exchange chromatography. The homology molecular model of this enzyme identified the Gln315 amino acid residue in the active site as potentially responsible for the enzyme high selectivity for GlcNAc substrates compared to fungal β-*N*-acetylhexosaminidases from other sources. The enzyme was used in a high-yielding sequential reaction together with the transglycosylating Tyr470His mutant of the β-*N*-acetylhexosaminidase from *Talaromyces flavus* to yield a bioactive disaccharide containing an azido functional group at C1. This azido moiety enabled its multivalent presentation on a simple aromatic carrier using click chemistry. The resulting glycoconjugate is a selective ligand of galectin-3, a human lectin with a high biomedical potential. This study demonstrates the potential of robust and selective carbohydrate-active enzymes with tailored catalytic capabilities in biotechnological and biomedical applications.

Materials and methods

Analytical methods

MS analysis of disaccharide 3 and conjugate 8

The electrospray mass spectra of disaccharide **3** and conjugate **8** were measured in the negative ion mode using the LTQ Orbitrap XL hybrid mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA). The mobile phase consisted of methanol/water (4/1, *v*/*v*) at a flow rate of 100 μL/min. The sample dissolved in methanol was injected into the mobile phase flow using a 2-μL loop. Spray voltage, capillary voltage, tube lens voltage, and capillary temperature were 4.6 kV, − 25 V, − 125 V, and 275 °C, respectively. The mass spectra were internally calibrated using deprotonated palmitic acid as lock mass.

MS analysis of compound 7

The EI spectrum of compound **7** was measured using orthogonal acceleration time-of-flight mass spectrometer GCT Premier (Waters, Milford, MA, USA). The sample was dissolved in chloroform, loaded into a quartz cup of the direct probe, and inserted into the ion source. The source temperature was 220 °C. Standard 70-eV spectra were recorded in the positive ion mode. For exact mass measurement, the spectrum was internally calibrated using perfluorotri-*n*-butylamine (heptacosa).

NMR analysis

NMR analysis was performed on Bruker AVANCE III 700 MHz (disaccharide **3**, compound **7**) and 400 MHz (conjugate **8**) spectrometers (Bruker BioSpin, Rheinstetten, Germany) in D_2O at 30 °C (**3** and **8**) or $CDCl_3$ at 20 °C (**7**). Residual signals of solvents ($CDCl_3$: δ_H 7.263 ppm, δ_C 77.01 ppm; D_2O: δ_H 4.732 ppm) served as internal standards; the carbon spectra in D_2O were referenced to the signal of acetone (δc 30.50). NMR experiments were acquired using standard manufacturers' software TopSpin 3.5. Two-parameter double-exponential Lorentz-Gauss function was applied for ^1H to improve resolution and line broadening (1 Hz) was applied to get better ^{13}C signal-to-noise ratio. Individual spin systems in compounds under study were assigned by COSY or 1d-TOCSY; the assignment was transferred to carbons by HSQC. The HMBC experiment enabled to assign quaternary carbons and join particular substructures together.

MS analysis of intact protein

The purified *Av*Hex protein in McIlvaine buffer pH 5.0 was diluted 1:1 with matrix (saturated solution of synapinic acid

dissolved in 50% acetonitrile/0.1% trifluoracetic acid/water) and spotted onto the MALDI target. The linear time-of-flight spectrum tuned to a mass range 20–50 kDa was recorded on a MALDI TOF instrument (Ultraflextreme, Bruker, Bremen, Germany).

Production of wild-type *Av*Hex in *A. versicolor*

The β-*N*-acetylhexosaminidase producing fungal strain of *A. versicolor* CCF 2491 (wild-type *Av*Hex) originated from the Culture Collection of Fungi (CCF), Department of Botany, Charles University in Prague (CZ) (Weignerová et al. 2003). The cultures were maintained on slants (g/L): mycological agar, 20.0; bacto-peptone, 5.0; and malt extract, 35.0. The peptone cultivation medium (M1) (Huňková et al. 1996) contained the following (g/L): yeast extract, 0.5; mycological peptone, 5.0; KH_2PO_4, 3.0; $NH_4H_2PO_4$, 5.0; casein acid hydrolyzate, 7.5; pH was adjusted to 6.0. The peptone cultivation medium induced with chitooligomers (M2) contained the following (g/L): yeast extract, 0.5; mycological peptone, 5.0; KH_2PO_4, 3.0; $NH_4H_2PO_4$, 5.0; chitin hydrolyzate, 7.5; pH was adjusted to 6.0. The mineral cultivation medium induced with GlcNAc (M3) (Slámová et al. 2012) contained the following (g/L): yeast extract, 0.5; KH_2PO_4, 3.0; $NH_4H_2PO_4$, 5.0; $(NH_4)_2SO_4$, 2.0; NaCl, 15.0; GlcNAc, 5.0 g/l); pH was adjusted to 4.0. After sterilization, sterile solution of $MgSO_4 \cdot 7H_2O$ (10% *w*/*v*) was added. Erlenmeyer flasks (500 mL) with medium (100 mL) were inoculated with 0.5 mL of a spore suspension in 0.1% (*v*/*v*) Tween 80 solution and cultivated on a rotary shaker at 200 rpm and 28 °C for 11–12 days.

The enzyme was purified from cultures grown in the GlcNAc-induced mineral medium (M3). After 12 days of cultivation, the mycelium was filtered off, and the medium was dialyzed against 10 mM citrate-phosphate buffer pH 3.5 for 4 h. Then, the medium (pH 3.5) was loaded on a cation-exchange chromatography column (Fractogel EMD SO_3^-; Merck, Darmstadt, Germany) equilibrated with 10 mM sodium citrate-phosphate buffer pH 3.5, using an Äkta Purifier protein chromatography system (Amersham Biosciences, Uppsala, Sweden). The enzyme was eluted with a linear gradient (80 mL, 2 mL/min) of 0–1 M NaCl. The pure fractions (as verified by SDS-PAGE) containing β-*N*-acetylhexosaminidase activity were combined and concentrated by ultrafiltration in McIlvaine buffer (50 mM citric acid/100 mM Na_2HPO_4) at pH 5.0.

Gene cloning and expression

The putative gene of the β-*N*-acetylhexosaminidase-encoding DNA (*Av*Hex; corresponding to the hypothetical protein ASPVEDRAFT_125959 (*A. versicolor* CBS 583.65, GenBank: OJI97839.1; 806 amino acid residues) was prepared synthetically (Generay, Shanghai, China) and cloned into the yeast expression vector pPICZαA (Invitrogen, Life Technologies, Carlsbad, CA, USA) downstream of the α-factor-encoding DNA segment for extracellular protein targeting and zeocin resistance using *Eco*RI and *Kpn*I restriction sites. The predicted signal propeptide sequence (residues 1–19) was not included in the cloned DNA segment (1770 base pairs), whose sequence was codon-optimized for the production in *P. pastoris* using Optigene™ Codon Optimization Analysis platform (Generay, Shanghai, China). It was deposited in GenBank under the accession number MH937562. The construct was linearized (15 μg) by restriction endonuclease *Sac*I (New England Biolabs, Ipswich, MA, USA), purified from an agarose gel using a JetQuick Gel Extraction Spin Kit (Genomed, Löhne, Germany) and used for electroporation of *P. pastoris* KM71H competent cells (Invitrogen, Life Technologies, Carlsbad, CA, USA) according to the manufacturer´s protocol (EasySelect Pichia Expression Kit; Invitrogen, Life Technologies, Carlsbad, CA, USA).

The resulting colonies were tested for *Av*Hex production as described below. Two colonies with the highest production of *Av*Hex in the highest purity as determined by SDS-PAGE electrophoresis and activity measurements were cryopreserved at − 80 °C in 15% (*v*/*v*) glycerol (OD_{600} = 50–100).

Production of recombinant *Av*Hex in *P. pastoris*

Recombinant *Av*Hex was produced essentially according to the manufacturer's instructions (EasySelect Pichia Expression Kit; Invitrogen, Life Technologies), which also shows the exact composition of cultivation media. For a small-scale production of the recombinant enzyme, the obtained transformants were inoculated into 100 mL of BMGY medium (Buffered Glycerol complex Medium) and incubated at 28 °C and 220 rpm overnight. Then the cells were collected by centrifugation (5000 rpm, 10 min, 20 °C) and the pellet was resuspended in 30 mL of BMMY medium (Buffered Methanol complex Medium). To induce expression of *Av*Hex, methanol (0.5% *v*/*v*) was added every 24 h to the cultures. The flasks were incubated on a rotary shaker at 28 °C and 220 rpm for 3 days. For the preparative production of recombinant *Av*Hex, 15 mL of YPD medium was inoculated with 100 μL of cryopreserved culture and incubated for 5 h at 28 °C and 220 rpm. Then, 1 L of BMGH medium (Buffered Minimal Glycerol medium) was inoculated with this preculture and incubated overnight under the same conditions. The cells were then collected and resuspended in 200 mL of BMMH medium (Buffered Minimal Methanol medium) and incubated at 28 °C on a rotary shaker with methanol addition (0.5% *v*/*v*) every 24 h for three days.

Recombinant *Av*Hex was purified from the culture supernatant as follows. The biomass was removed by centrifugation (5000 rpm, 10 min, 12 °C), the resulting supernatant was

diluted four times with water, and its pH was adjusted to 6.5. The enzyme was purified in a single step by anion-exchange chromatography (Q-Sepharose Fast Flow, GE Healthcare, Pittsburgh, PA, USA) on an Äkta Purifier protein chromatography system (Amersham Biosciences, Uppsala, Sweden) equilibrated with 10 mM sodium citrate-phosphate buffer pH 6.5. The enzyme was eluted with a linear gradient (60 mL, 2 mL/min) of 0–1 M NaCl. The pure fractions (as verified by SDS-PAGE) containing β-*N*-acetylhexosaminidase activity were combined and concentrated by ultrafiltration in McIlvaine buffer (50 mM citric acid/100 mM Na_2HPO_4) at pH 5.0.

Sequence determination by LC-MS/MS

Proteins in solution or in gel bands were reduced by dithiothreitol and alkylated by iodoacetamide prior to overnight digestion with trypsin or chymotrypsin in 50 mM ammonium acetate buffer pH 8.5. Digestion with pepsin was carried out under acidic conditions (0.04 M HCl) for 2 h. Peptides in solution were desalted on a C18 spin column and dried; peptides in gel bands were extracted by acetone and dried in Speed-Vac concentrator. Resulting peptides were analyzed on an UltiMate 3000 RSLCnano system (Thermo Scientific, Waltham, MA, USA) coupled to a TripleTOF 5600 mass spectrometer with a NanoSpray III source (Sciex, Darmstadt, Germany). The peptides were trapped and desalted with 2% acetonitrile in 0.1% formic acid at a flow rate of 5 μL/min using an Acclaim PepMap100 column (5 μm, 2 cm × 100 μm ID, Thermo Scientific, Waltham, MA, USA). Eluted peptides were separated on an Acclaim PepMap100 analytical column (3 μm, 25 cm × 75 μm ID, Thermo Scientific, Waltham, MA, USA). The 70-min elution gradient with a constant flow of 300 nL/min was set to 5% of phase B (0.1% formic acid in 99.9% acetonitrile; phase A, 0.1% formic acid) for the first 10 min; then, gradient elution took place with an increasing concentration of acetonitrile. The TOF-MS mass range was set to 350–1250 *m*/*z*; in MS/MS mode, the instrument acquired fragmentation spectra with *m*/*z* in a range of 100–1600. Protein Pilot 4.5 (Sciex, Darmstadt, Germany) was used for protein identification using the Uniprot database of available *Aspergillus* species protein sequences and the putative *Av*Hex sequence (hypothetical protein ASPVEDRAFT_125959, BLAST) (Altschul et al. 1990).

Protein characterization

Protein concentration was determined by the Bradford method (Bradford 1976) using Protein Assay Dye Reagent Concentrate (Bio-Rad, Watford, Hertfordshire, UK) calibrated for bovine plasma γ-globulin (IgG, BioRad, Watford, Hertfordshire, UK). The purity of *Av*Hex was verified by SDS-PAGE in 10% polyacrylamide gel. Native electrophoresis was performed in 7% polyacrylamide gel using 0.02 M HEPES/0.18 M glycine buffer pH 8.0.

Protein deglycosylation

The wild-type and recombinant *Av*Hex enzymes were *N*-deglycosylated by endoglycosidase Endo H (New England Biolabs, Ipswich, MA, USA), using the manufacturer's protocol. The buffer of the enzyme sample (containing 115 μg of protein) was exchanged for the supplier's Glyco-buffer (final volume 120 μL). Endo H (30 μL, 15,000 U) was added and the reaction mixture was incubated at 37 °C and 300 rpm overnight. The deglycosylation resulted in a shift the apparent molecular weights of the native and *N*-deglycosylated enzymes on SDS-PAGE gel.

Standard assay for enzyme activity

The β-*N*-acetylglucosaminidase activity of *Av*Hex was assayed in end-point experiments using *p*-nitrophenyl 2-acetamido-2-deoxy-β-D-glucopyranoside (**1**; *p*NP-GlcNAc; Sigma-Aldrich, St. Louis, MO, USA) as a substrate with a starting concentration of 2 mM. After incubation of the reaction mixture in McIlvaine buffer (50 mM citric acid/100 mM Na_2HPO_4 pH 5.0) at 35 °C and 850 rpm for 10 min, the liberated *p*-nitrophenol was determined spectrophotometrically at 420 nm under alkaline conditions (50 μL of the reaction mixture was mixed with 1 mL of 0.1 M Na_2CO_3). One unit of enzymatic activity was defined as the amount of enzyme releasing 1 μmol of *p*-nitrophenol per minute under the above conditions.

The effect of pH on *Av*Hex activity was measured at 35 °C as described above; Britton-Robinson buffer (0.04 M H_3PO_4, 0.04 M phenylacetic acid, 0.04 M H_3BO_3/0.2 M NaOH) was used for the pH range of 2–10 with the respective blanks void of enzyme. The temperature activity profile of *Av*Hex was determined as described above in McIlvaine buffer pH 5.0, in the range of 25–80 °C.

Michaelis–Menten parameters of hydrolysis of *p*NP-GlcNAc by *Av*Hex were determined in a discontinuous kinetic assay using *p*NP-GlcNAc as substrate at a concentration of 0.1–4.0 mM in McIlvaine buffer pH 5.0. At each substrate concentration, the reaction mixture (400 μL) was incubated at 25 °C and 850 rpm. The hydrolytic reaction lasted for ca 15 min. Aliquots (50 μL) were taken at regular time intervals and added to 100 μL of 1 M Na_2CO_3 in the wells of a styrene flat-bottom 96-well microtiter plate (Thermo Scientific, Waltham, MA, USA). The absorbance at 420 nm was measured using the Spectra Max Plus spectrophotometer (Molecular Devices, Sunnyvale, CA, USA). All data were measured in duplicates. The initial reaction velocity for each substrate concentration was determined by linear regression of respective time points, and the kinetic constants (K_m, k_{cat})

were determined from the initial velocities for each substrate concentration using GraphPad Prism 7 (GraphPad Software, La Jolla, CA, USA).

Homology modeling, docking, and molecular dynamics

We used the sequence of the hypothetical protein ASPVEDRAFT_125959 (GenBank database) in a BLAST search (Altschul et al. 1990) to identify protein homologs. Homology modeling was performed with Modeller 9.16 using the crystal structure of the *Aspergillus oryzae* β-*N*-acetylhexosaminidase (*Ao*Hex) as a template (Sali and Blundell 1993). Sequence alignment was constructed with T-coffee (Notredame et al. 2000) and manually modified. We employed several models for refinement (Raval et al. 2012) with YASARA (Krieger et al. 2002) using the YASARA force field. The most stable model was obtained from the refinement of the glycosylated dimeric model with *p*NP-GlcNAc substrate in the active site using constrained molecular dynamics simulation followed by a short (1 ns) free simulation of the system. Molprobity was used for validation of model geometry: 95.8% (548/572) residues of the best model were in favored areas of the Ramachandran plot and just 3 of 4375 covalent bonds had unfavorable geometry; all clashes were resolved (Davis et al. 2007). The averaged knowledge-based energy residues and *Z*-score (− 11.07 for 572 aa) estimated by ProSA were within favorable values (Wiederstein and Sippl 2007).

GlyProt (Bohne-Lang and von der Lieth 2005) and NGlycPred server (Chuang et al. 2012) were used for predicting the protein glycosylation, modeled in YASARA based on the glycosylation of the *Ao*Hex structure (pdb: 5OAR).

The initial orientation of substrates, docking, molecular dynamics simulation, and complex analysis was performed with YASARA (Krieger et al. 2002; Kulik et al. 2015) using YAMBER 2 force field. Carbohydrate parameters were used from Glycam force field incorporated in YASARA; non-standard residues were parameterized with AutoSMILES algorithm (Jakalian et al. 2002). Representative binding scores were estimated by Glide from Schrödinger package (Friesner et al. 2006) using values averaged over substrate–enzyme complexes. The Glide XP score (Friesner et al. 2006) approximates the ligand binding free energy and includes an electrostatic term, terms for van der Waals interactions, hydrogen bonds, π-stacking and π-cation interactions, contributions of hydrophobic interactions, strain energy (unfavorable contacts), and penalty upon desolvation of the substrate or protein.

2-Acetamido-2-deoxy-β-D-glucopyranosyl azide (2)

The title compound **2** was prepared from GlcNAc based on the procedures described previously (Bojarová et al. 2017). The ^{1}H and ^{13}C NMR data were consistent with the structure.

Sequential enzymatic synthesis of 2-acetamido-2-deoxy-β-D-galactopyranosyl-(1 → 4)-2-acetamido-2-deoxy-β-D-glucopyranosyl azide (3)

For the first synthetic step of the sequential reaction, we used the Tyr470His mutant of the β-*N*-acetylhexosaminidase from *T. flavus* (*Tf*Hex) with an enhanced synthetic capability. The enzyme was prepared as described previously (Slámová et al. 2015), by extracellular expression in *P. pastoris* KM71H (Invitrogen, Life Technologies, Carlsbad, CA, USA) followed by a single purification step of cation-exchange chromatography.

*p*NP-GalNAc (**1**; 41 mg, 0.12 mmol), acceptor **2** (118 mg, 0.48 mmol), and the Tyr470His mutant of *Tf*Hex (1.4 U, 12.6 mg, 450 μL) were incubated in a mixture of acetonitrile (10% *v*/*v*) and McIlvaine buffer pH 5.0 (total reaction volume 2.4 mL) at 45 °C and 1000 rpm. The reaction was monitored by HPLC and TLC (propan-2-ol/water/NH_4OH aq., 7:2:1). After 6.5 h, the reaction was stopped by boiling for 2 min, and, upon cooling down to 35 °C, the *Av*Hex enzyme was added (4.9 U, 0.12 mg, 30 μL) and the reaction was incubated at 35 °C and 850 rpm for 7 h to cleave 2-acetamido-2-deoxy-β-D-glucopyranosyl-(1 → 4)-2-acetamido-2-deoxy-β-D-glucopyranosyl azide (**4**) that originated in the transglycosylation reaction as an unwanted by-product. The reaction was monitored as above and then stopped by boiling for 2 min. After cooling down to room temperature, the reaction mixture was centrifuged (13,500 rpm; 10 min) and loaded onto a Biogel P-2 column (2.6 × 100 cm, Bio-Rad, Watford, Hertfordshire, UK) using water as a mobile phase at an elution rate of 7.5 mL/h. Pure acceptor **2** was partially recovered (28 mg). The fractions containing the product were combined and lyophilized; the title compound **3** was obtained as a white solid (25 mg, 0.056 mmol, 47% isolated yield). HRMS (ESI$^-$) found *m*/*z* 448.16819 (expected 448.16852 for $[M-H]^-$, $C_{16}H_{26}N_5O_{10}$). For the respective NMR and MS data, see the Supplementary Material (Table S1, Figs. S1, S2).

1,2,3-*tris*(Alynyloxy)benzene (7)

The title compound **7** was prepared according to a published procedure, with some modifications (Xie et al. 2008). To a solution of pyrogallol (**6**; 1.26 g, 10 mmol) in dry DMF (dimethylformamide, 30 mL), anhydrous K_2CO_3 (5.52 g, 40 mmol) was added and the solution was heated at 50 °C for 1 h under argon. 3-Bromo-1-propyne (4.31 mL, 40 mmol) was added dropwise and the reaction ran for 2 days at 65 °C, monitored by TLC (ether/hexane, 1:1). After cooling down, the reaction mixture was poured into water (300 mL), then chloroform (150 mL) was added, the phases were separated and the water phase was extracted with chlorofom (3 × 100 mL). The combined organic phases were concentrated in vacuo and repeatedly coevaporated with water to remove

the traces of DMF. The crude compound **7** was purified by column chromatography on silica gel using ether/hexane, 1:3, as a mobile phase. The title compound **7** was isolated as a yellowish solid (1.67 g, 7.0 mmol, 70% yield). HRMS (EI^+) found *m*/*z* 240.0785 (expected *m*/*z* 240.0786 for $M^{+\cdot}$, $C_{15}H_{12}O_3$). For the respective NMR and MS data, see the Supplementary Material (Table S2, Figs. S3, S4).

1,2,3-*tris*-[1-(2-Acetamido-2-deoxy-β-D-galactopyranosyl-(1 → 4)-2-acetamido-2-deoxy-β-D-glucopyranosyl) -1H-1,2,3-triazol-4-yl)-2-ethyloxy]benzene (8)

Compound **7** (5 mg, 0.021 mmol), sodium ascorbate (4.1 mg, 0.021 mmol), and disaccharide **3** (29 mg, 0.065 mmol) were dissolved in a mixture of water (300 μL) and methanol (735 μL). The reaction mixture was bubbled with argon and a solution of $CuSO_4 \cdot 5H_2O$ (5.2 mg, 0.021 mmol) in water (60 μL) was added. The reaction mixture was bubbled with argon again and left overnight at room temperature. After complete conversion was detected by TLC (propan-2-ol/H_2O/NH_4OH aq, 7:2:1; 3 times developed), PBS buffer containing 5% *w*/*v* EDTA (700 μL) was added. After 30 min, the reaction mixture was concentrated to one third of its volume to remove methanol and loaded onto a Biogel P2 column (2.6 × 100 cm, Bio-Rad, Watford, Hertfordshire, UK) with water as a mobile phase, at an elution rate of 7.3 mL/h. The fractions containing product **8** were combined and lyophilized. The title compound **8** was isolated as a white solid (17.5 mg, 0.011 mmol, 53%). HRMS (ESI^-) found *m*/*z* 1586.59778 (expected *m*/*z* 1586.59874 for $[M-H]^-$; $C_{63}H_{92}N_{15}O_{33}$). For the respective NMR and MS data, see the Supplementary Material (Table S3, Figs. S5, S6).

Competitive ELISA assay

Recombinant galectin-3 used in the assay was prepared as described previously (Bumba et al. 2018). The affinity of galectin-3 for conjugate **8** was determined by competitive ELISA assay essentially as reported previously (Laaf et al. 2017b; Bumba et al. 2018). Briefly, the F16 Maxisorp NUNC-Immuno Modules (Thermo Scientific, Hvidovre, Denmark) were coated overnight with asialofetuin (Sigma-Aldrich, Hamburg, Germany; 0.1 μM in PBS, 50 μL, 5 pmol/well). Then they were blocked with bovine serum albumin (2% *w*/*v*) dissolved in phosphate-buffered saline (PBS) for 1 h at room temperature. Afterwards, a mixture of the inhibitor compound in varying concentrations together with recombinant galectin-3 (total volume 50 μL; 6.5 μM final galectin-3 concentration) in EPBS buffer (50 mM NaH_2PO_4, 150 mM NaCl, 2 mM ethylendiaminetetraacetic acid, pH 7.5) was added and incubated for 2 h. In a control assay, this step was performed in the presence of 0.025% *v*/*v* Tween. Extensive rinsing of wells after every step with PBS containing Tween 20 (0.05% *v*/*v*) was performed. Detection of bound galectin-3 was achieved using anti-His_6-IgG1 antibody from mouse conjugated with horseradish peroxidase (Roche Diagnostics, Mannheim, Germany) diluted in PBS (1:1000, 50 μL, 1 h, room temperature). TMB One (Kem-En-Tec, Taastrup, Denmark) substrate solution (50 μL) was used to initiate the reaction with IgG-conjugated peroxidase. The reaction was stopped by adding 3 M hydrochloric acid (50 μL). The binding signal of bound galectin-3 was quantified colorimetrically at 450 nm using Sunrise absorbance microplate reader (Tecan Group Ltd, Männedorf. Switzerland). The obtained data were analyzed by non-linear regression (dose response-inhibition-variable slope) using GraphPad Prism 7 (GraphPad software, San Diego, CA, USA).

Results

Wild-type *Av*Hex production in *A. versicolor* and purification

*Av*Hex was identified as the most selective enzyme for *N*-acetylglucosamine (GlcNAc) substrates in our library of β-*N*-acetylhexosaminidases (> 100 enzymes) from filamentous fungi on the basis of our previous work (Weignerová et al. 2003). Its GlcNAcase/GalNAcase activity ratio ranges between ca 9 and 11 depending on the particular production batch. β-*N*-Acetylhexosaminidases are produced extracellularly by their native hosts into the cultivation medium and their production may be enhanced in the presence of various inducers. We compared three cultivation media and identified M3, mineral medium induced with submersed GlcNAc (5 g/L) (Slámová et al. 2012), as optimum for wild-type enzyme production (ca 0.4 U/ml; 11 U/mg protein). The time dependence of the secretion profile was similar to other fungi (Ryšlavá et al. 2011), with the maximum β-*N*-acetylglucosaminidase activity found after 10–12 days of cultivation at 28 °C (Fig. 1a).

The enzyme was purified to apparent homogeneity as determined by SDS-PAGE in a single chromatography step — cation-exchange chromatography at pH 3.5, preceded by dialysis against the loading chromatography buffer to reduce the medium conductivity. The overall purification yield reached 28% and the specific activity of purified *Av*Hex was 79 U/mg protein (Table 1). In total, an average of 0.13 mg protein could be isolated from 100 mL of cultivation medium. A small protein with a molecular weight of ca 14 kDa co-purified with the enzyme could be found in heavily overloaded samples (Fig. 1b, lanes 1, 2). On the basis of the previous research on fungal β-*N*-acetylhexosaminidases (Plíhal et al. 2007; Škerlová et al. 2018) and the sequence alignment (Fig. 2), it was assigned to be the *Av*Hex propeptide (aa 20–97). The size of the catalytic domain was estimated to be 66 kDa using SDS-PAGE; its *N*-

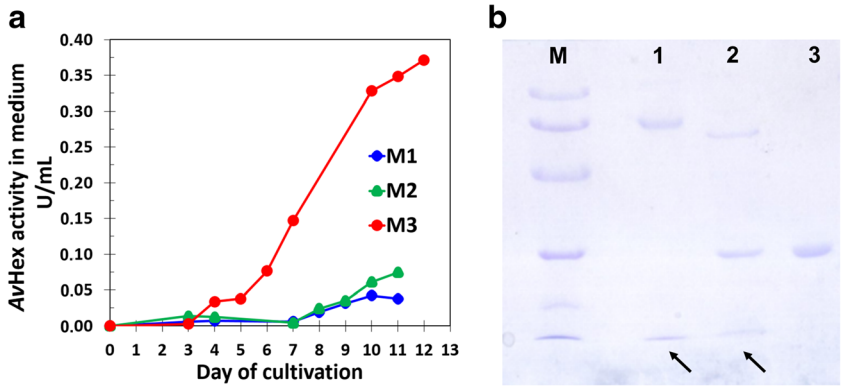


Fig. 1 **a** Time course of secretion of *Av*Hex in various cultivation media (M1–M3). The best production was achieved in M3 medium (yeast extract, 0.5; KH_2PO_4, 3.0; $NH_4H_2PO_4$, 5.0; $(NH_4)_2SO_4$, 2.0; NaCl, 15.0; GlcNAc, 5.0 g/L). **b** Purification and *N*-deglycosylation of *Av*Hex as analyzed by SDS-PAGE electrophoresis. Lane M, molecular mass markers consisting of phosphorylase b from rabbit muscle (97 kDa), bovine serum albumin (66 kDa), chicken egg white ovalbumin (45 kDa), carbonic anhydrase from bovine erythrocytes (30 kDa), trypsin inhibitor from soya beans (20.1 kDa), and α-lactalbumin from bovine milk (14.4 kDa); lane 1, wild-type *Av*Hex (66 kDa) with propeptide (14 kDa) purified by cation-exchange chromatography; lane 2, wild-type *Av*Hex after *N*-deglycosylation (58 kDa) with Endo H (29 kDa) and Endo H; lane 3, Endo H. The position of the putative propeptide co-purifying with the catalytic subunit is indicated with an arrow

glycosylation amounted to ca 8 kDa. This agrees well with the previously published data on related enzymes (Plíhal et al. 2007; Ryšlavá et al. 2011; Slámová et al. 2012) and with the theoretical calculated MW of the *Av*Hex catalytic domain of 56,872.29 kDa. In native electrophoresis, the *Av*Hex protein appears as a single band of MW roughly 140 kDa but this is just an approximate value influenced by protein shape and charge distribution. This is in agreement with the native form of this enzyme being a dimer with each monomer containing a glycoside hydrolase family 20 catalytic domain (http://www.cazy.org/) and a propeptide (a total of 160 kDa). The purified and concentrated enzyme was stable for ca 6 months when stored at 4 °C in McIlvaine buffer pH 5.0 conserved with 0.03% (*w*/*v*) NaN_3 (< 20% loss of activity).

Structural alignment and sequencing of the wild-type *Av*Hex

The complete amino acid sequence of *Av*Hex (ASPVEDRAFT_125959; GenBank: OJI97839.1) has more than 65% identity to many known fungal β-*N*-acetylhexosaminidases (Fig. 2). Namely, to *Ao*Hex (*A. oryzae*, GenBank: AY091636; Škerlová et al. 2018), *Po*Hex (*Penicillium oxalicum*, GenBank: EU189026; (Ryšlavá et al. 2011), and *Tf*Hex (*T. flavus*, GenBank: AY091636; Kulik et al. 2015) with 69, 66, and 61 identities over complete aligned sequences, respectively. Additionally, high sequence identities were observed in fungal β-*N*-acetylglucosaminidases (Table S4, Supplementary Material). In contrast, identities of less than 31% were determined in bacterial β-*N*-acetylhexosaminidase-encoding sequences (Fig. 2).

As we can see from Fig. 2, active-site amino acid residues in *Av*Hex are conserved except for Gln315 located in the loop HL1. This residue is quite frequently found as glutamate in β-*N*-acetylhexosaminidases. In contrast, in the GlcNAc-selective chitobiase from *Serratia marcescens* (*Sm*Chit; 1QBB), this residue is present as glutamine (Gln494). Gln is also found in a number of β-*N*-acetylglucosaminidases and other chitinases as shown in a multiple sequence alignment with fungal enzymes (Table S4, Supplementary Material).

In order to compare the putative amino acid sequence with the sequence of the purified wild-type *Av*Hex, peptide mapping was performed using various proteases (pepsin, trypsin, chymotrypsin). The enzyme was sequenced in both the native and deglycosylated forms, isolated from the SDS-PAGE gel or in solution. The *N*-terminal sequence of 19 amino acids (Fig. 3, in bold black letters) was not found in the wild-type *Av*Hex by any

Table 1 Purification of *Av*Hex from the culture medium of *A. versicolor* CCF 2491 and *P. pastoris* (800 mL)

Host	Purification step	Total protein (mg)	Total activity (U)	Specific activity (U/mg)	Purification factor (fold)	Yield (%)
A. versicolor	Culture medium	26.5	297.1	11.2	0	100
	Dialysis	19.6	237.0	12.1	1.1	80
	Cation exchange	1.05	83.4	79.4	7.1	28
P. pastoris	Culture medium	55.8	2000	36.8	0	100
	Anion exchange	18.4	1680	45.7	4.2	84

sequencing protocols. Therefore, we tentatively assigned it to the *A. versicolo*r signal sequence similar to other fungal β-*N*-acetylhexosaminidases (Slámová et al. 2012; Plíhal et al. 2007).

Molecular cloning, recombinant expression, and purification of *Av*Hex

Since the found protein amino acid sequence of the wild-type *Av*Hex and the translated amino acid sequence of the hypothetical protein ASPVEDRAFT_125959 were in good accordance, we used this sequence for the construction of the synthetic gene of *Av*Hex. The gene was synthesized commercially, and the respective nucleotide sequence was optimized for the expression in *P. pastoris* (GenBank: MH937562; see Fig. S7, Supplementary Material). The *Av*Hex sequence contained a single open reading frame of 1770 base pairs encoding 589 amino acids, comprising the propeptide and the catalytic domain (Slámová et al. 2012, Plíhal et al. 2007).

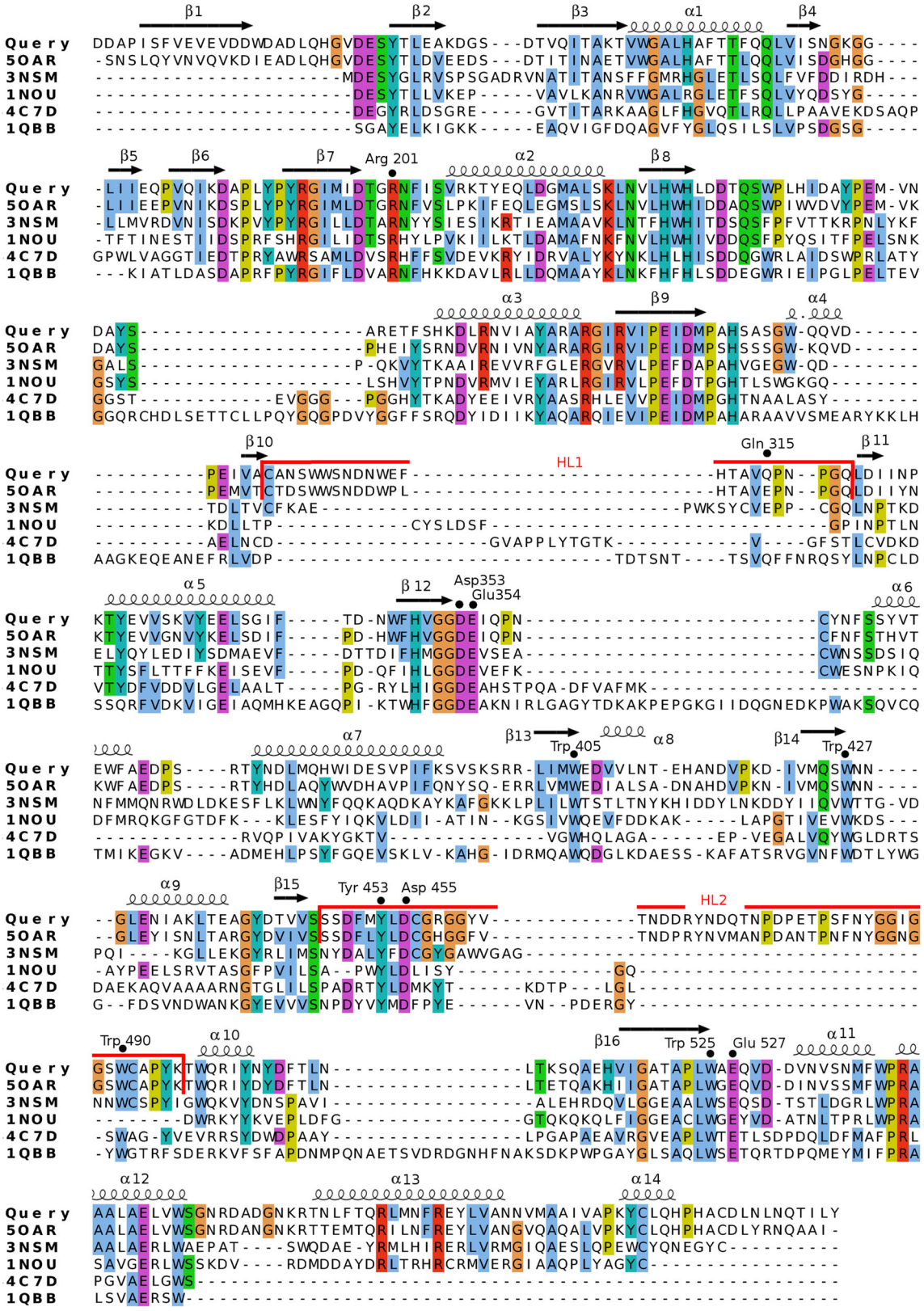
The pPICZαA-based construct enabled *Av*Hex to be secreted to the minimal cultivation medium using the *P. pastoris* KM71H strain, which previously proved to be an efficient host for the production of β-*N*-acetylhexosaminidases (Slámová et al. 2012, 2015). Upon induction by methanol, the highest activity (2.5 U/mL, 37 U/mg) was found after 3–4 days of cultivation. The recombinant *Av*Hex was purified from the medium in a single step by anion-exchange chromatography at pH 6.5, which was selected as the second option since the recombinant enzyme tended to bind irreversibly to the Fractogel resin used for the wild-type counterpart, resulting in purification yields ≤ 10%. By the selected method, though, we were able to isolate an average of 2.3 mg of recombinant *Av*Hex from 100 mL of cultivation medium with a specific activity of 46 U/mg (overall purification yield 84%; Table 1). Thus, we significantly increased and simplified the enzyme production compared to the tedious production in fungi (the cultivation time reduced by half, the enzyme production enhanced 20 times). The recombinant *Av*Hex appeared as a single band of ca 140 kDa on native electrophoresis (Fig. 4b), identical to the wild-type enzyme. Electrophoresis under reducing conditions (SDS-PAGE; Fig. 4a) showed two closely positioned bands (78 kDa and 64 kDa as determined by MALDI-MS; Fig. S8, Supplementary Material). The *N*-deglycosylation by EndoH resulted in a pronounced shift of the *Av*Hex protein bands on SDS-PAGE gel by ca 6 kDa, corresponding to the values of 72 kDa and 58 kDa, respectively, determined by MALDI-MS (Fig. S8, Supplementary Material). LC-MS/MS analysis of the protein bands confirmed that both protein bands correspond to *Av*Hex. Based on MALDI-MS data, we suggest that the catalytic domains of the recombinant *Av*Hex enzyme are produced with two different patterns of *O*-glycosylation, which have no measurable influence on *Av*Hex catalytic properties. Again, the *Av*Hex propeptide (ca 14 kDa) was visible as a weak band on a SDS-PAGE gel (Fig. 4b). The native active *Av*Hex is composed of two monomers, each containing two polypeptide chains (catalytic domain and propeptide), and it has a molecular weight of ca 170 kDa as calculated from MALDI-MS results, using an average mass of both catalytic domains for simplification. The recombinant enzyme was perfectly stable under long-term storage at 4 °C in McIlvaine buffer pH 5.0.

Fig. 2 Multiple sequence alignment of *Av*Hex catalytic domain ("Query") with solved crystal structures of catalytic domains of β-*N*-acetylhexosaminidases from various sources performed with Esprint 3.0 (Robert and Gouet 2014) and Jalview (Clamp et al. 2004). Active-site amino acids are marked by black dots and labeled according to *Av*Hex numeration. Identities of the aligned sequences are as follows: *Ao*Hex (pdb: 5OAR/chain B)—73% (99% coverage), β-*N*-acetylhexosaminidase from insect *Ostrinia furnacalis* (pdb: 3NSM/chain A)—31% (76% coverage), human lysosomal β-hexosaminidase isoform B (pdb: 1NOW/chain A)—29% (75% coverage), β-*N*-acetylhexosaminidase from bacteria *Streptomyces coelicolor* (pdb: 4C7D)—26% (68% coverage), *Sm*Chit (pdb: 1QBB)—24% (68% coverage). Residues with identities higher than 30% for a particular site are colored by ClustalX coloring scheme, which groups residues according to similar properties by selected color: G—**orange**; P—**yellow**; positively charged K, R—**red**; Y, H—**cyan**; negatively charged D, E—**magenta**; polar Q, N, S, T—**green**; hydrophobic L, V, M, A, I, W, C, F—**blue**. Secondary structure elements are based on the 5OAR structure and defined by symbols and letters as α-alpha helix or β-beta sheet. The long loops HL1 and HL2 conserved in fungal β-*N*-acetylhexosaminidases are labeled; their ends are marked by red lines

Biochemical properties of wild-type and recombinant *Av*Hex

The biochemical parameters of both the wild-type and recombinant *Av*Hex were characterized and compared to determine the effect of the host organism on the properties of the enzyme. Like other related β-*N*-acetylhexosaminidases (Ettrich et al. 2007; Ryšlavá et al. 2011), *Av*Hex was found to be a glycoprotein as also predicted by the presence of the glycosylation sites in its amino acid sequence (Fig. S9, Supplementary Material). Using the GlyProt (Bohne-Lang and von der Lieth 2005) and NGlycPred server (Chuang et al. 2012), four potential *N*-glycosylation sites were identified in the catalytic domain of *Av*Hex. Two of them (Asn361 and Asn508) correspond to the confirmed glycosylated sites in *Ao*Hex. The NGlycPred server predicted lower probability for the two remaining sites (Asn533 and Asn603) to be *N*-glycosylated in *Av*Hex (probability 52–51%). This is in accordance with the situation in *Ao*Hex where these respective two residues are not glycosylated either (one of them is buried inside the protein, and the other carries a proline in the glycosylation motif). *N*-Deglycosylation of both the wild-type and the recombinant enzymes (see also Fig. 4a) had no influence on their activities nor on GlcNAcase selectivities.

Using *p*NP-GlcNAc as a standard substrate, we determined the pH and temperature optima of the hydrolytic reaction catalyzed by *Av*Hex (Fig. 4c, d). Both the wild-type and


β1
β2
β3
α1
β4
Query DDAPISFVEVEVDDWDADLQHGVDESYTLEAKDGS---DTVQITAKTVWGALHAFTTFQQLVISNGKGG----
5OAR -SNSLQYVNVQVKDIEADLQHGVDESYTLDVEEDS---DTITINAETVWGALHAFTTLQQLVISDGHGG----
3NSM ----------MDESYGLRVSPSGADRVNATITANSFFGMRHGLETLSQLFVFDDIRDH---
1NOU ----------DESYTLLVKEP-----VAVLKANRVWGALRGLETFSQLVYQDSYG-----
4C7D ----------DEGYRLDSGRE-----GVTITARKAAGLFHGVQTLRQLLPAAVEKDSAQP
1QBB ----------SGAYELKIGKK-----EAQVIGFDQAGVFYGLQSILSLVPSDGSG-----
β5
β6
β7
Arg 201
α2
β8
Query -LIIEQPVQIKDAPLYPYRGIMIDTGRNFISVRKTYEQLDGMALSKLNVLHWHLDDTQSWPLHIDAYPEM-VN
5OAR -LIIEEPVNIKDSPLYPYRGIMLDTGRNFVSLPKIFEQLEGMSLSKLNVLHWHIDDAQSWPIWVDVYPEM-VK
3NSM -LLMVRDVNISDKPVYPYRGILLDTARNYYSIESIKRTIEAMAAVKLNTFHWHITDSQSFPFVTTKRPNLYKF
1NOU -TFTINESTIIDSPRFSHRGILIDTSRHYLPVKIILKTLDAMAFNKFNVLHWHIVDDQSFPYQSITFPELSNK
4C7D GPWLVAGGTIEDTPRYAWRSAMLDVSRHFFSVDEVKRYIDRVALYKYNKLHLHISDDQGWRLAIDSWPRLATY
1QBB ---KIATLDASDAPRFPYRGIFLDVARNFHKKDAVLRLLDQMAAYKLNKFHFHLSDDEGWRIEIPGLPELTEV
α3
β9
α4
Query DAYS----------ARETFSHKDLRNVIAYARARGIRVIPEIDMPAHSASGW-QQVD------
5OAR DAYS----------PHEIYSRNDVRNIVNYARARGIRVIPEIDMPSHSSSGW-KQVD------
3NSM GALS----------P-QKVYTKAAIREVVRFGLERGVRVLPEFDAPAHVGEGW-QD-------
1NOU GSYS----------LSHVYTPNDVRMVIEYARLRGIRVLPEFDTPGHTLSWGKGQ--------
4C7D GGST----------EVGGG---PGGHYTKADYEEIVRYAASRHLEVVPEIDMPGHTNAALASY--------
1QBB GGQRCHDLSETTCLLPQYGQGPDVYGGFFSRQDYIDIIKYAQARQIEVIPEIDMPAHARAAVVSMEARYKKLH
β10
HL1
Gln 315
β11
Query ----------PEIVACANSWWSNDNWEF----------HTAVQPN--PGQLDIINP
5OAR ----------PEMVTCTDSWWSNDDWPL----------HTAVEPN--PGQLDIIYN
3NSM ----------TDLTVCFKAE----------PWKSYCVEPP--CGQLNPTKD
1NOU ----------KDLLTP----------CYSLDSF----------GPINPTLN
4C7D ----------AELNCD----------GVAPPLYTGTK----------V----GFSTLCVDKD
1QBB AAGKEQEANEFRLVDP----------TDTSNT----TSVQFFNRQSYLNPCLD
α5
β12
Asp353
Glu354
α6
Query KTYEVVSKVYEELSGIF-----TD-NWFHVGGDEIQPN----------CYNFSSYVT
5OAR KTYEVVGNVYKELSDIF-----PD-HWFHVGGDEIQPN----------CFNFSTHVT
3NSM ELYQYLEDIYSDMAEVF-----DTTDIFHMGGDEVSEA----------CWNSSDSIQ
1NOU TTYSFLTTFFKEISEVF-----PD-QFIHLGGDEVEFK----------CWESNPKIQ
4C7D VTYDFVDDVLGELAALT-----PG-RYLHIGGDEAHSTPQA-DFVAFMK----------
1QBB SSQRFVDKVIGEIAQMHKEAGQPI-KTWHFGGDEAKNIRLGAGYTDKAKPEPGKGIIDQGNEDKPWAKSQVCQ
α7
β13
Trp 405
α8
β14
Trp 427
Query EWFAEDPS----RTYNDLMQHWIDESVPIFKSVSKSRR-LIMWEDVVLNT-EHANDVPKD-IVMQSWNN----
5OAR KWFAEDPS----RTYHDLAQYWVDHAVPIFQNYSQ-ERRLVMWEDIALSA-DNAHDVPKN-IVMQSWNN----
3NSM NFMMQNRWDLDKESFLKLWNYFQQKAQDKAYKAFGKKLPLILWTSTLTNYKHIDDYLNKDDYIIQVWTTG-VD
1NOU DFMRQKGFGTDFK---KLESFYIQKVLDII-ATIN--KGSIVWQEVFDDKAK-----LAPGTIVEVWKDS---
4C7D ----------RVQPIVAKYGKTV----------VGWHQLAGA------EP-VEGALVQYWGLDRTS
1QBB TMIKEGKV----ADMEHLPSYFGQEVSKLV-KAHG-IDRMQAWQDGLKDAESS-KAFATSRVGVNFWDTLYWG
α9
β15
Tyr 453
Asp 455
HL2
Query --GLENIAKLTEAGYDTVVSSSDFMYLDCGRGGYV----------TNDDRYNDQTNPDPETPSFNYGGIG
5OAR --GLEYISNLTARGYDVIVSSSDFLYLDCGHGGFV----------TNDPRYNVMANPDANTPNFNYGGNG
3NSM PQI----KGLLEKGYRLIMSNYDALYFDCGYGAWVGAG----------
1NOU -AYPEELSRVTASGFPVILSA--PWYLDLISY----------GQ----------
4C7D DAEKAQVAAAARNGTGLILSPADRTYLDMKYT------KDTP---LGL----------
1QBB G--FDSVNDWANKGYEVVSNPDYVYMDFPYE------VN--PDERGY----------
Trp 490
α10
β16
Trp 525
Glu 527
α11
Query GSWCAPYKTWQRIYNYDFTLN----------LTKSQAEHVIGATAPLWAEQVD-DVNVSNMFWPRA
5OAR GSWCAPYKTWQRIYDYDFTLN----------LTETQAKHIIGATAPLWGEQVD-DINVSSMFWPRA
3NSM NNWCSPYIGWQKVYDNSPAVI----------ALEHRDQVLGGEAALWSEQSD-TSTLDGRLWPRA
1NOU --------DWRKYYKVEPLDFG----------GTQKQKQLFIGGEACLWGEYVD-ATNLTPRLWPRA
4C7D -SWAG-YVEVRRSYDWDPAAY----------LPGAPAEAVRGVEAPLWTETLSDPDQLDFMAFPRL
1QBB -YWGTRFSDERKVFSFAPDNMPQNAETSVDRDGNHFNAKSDKPWPGAYGLSAQLWSETQRTDPQMEYMIFPRA
α12
α13
α14
Query AALAELVWSGNRDADGNKRTNLFTQRLMNFREYLVANNVMAAIVAPKYCLQHPHACDLNLNQTILY
5OAR AALAELVWSGNRDANGNKRTTEMTQRILNFREYLVANGVQAQALVPKYCLQHPHACDLYRNQAAI-
3NSM AALAERLWAEPAT----SWQDAE-YRMLHIRERLVRMGIQAESLQPEWCYQNEGYC----------
1NOU SAVGERLWSSKDV----RDMDDAYDRLTRHRCRMVERGIAAQPLYAGYC----------
4C7D PGVAELGWS----------
1QBB LSVAERSW----------

recombinant *Av*Hex displayed similar pH and temperature profiles. pH optimum was found to be in a broad plateau between pH 3.0 and 6.0 (Fig. 4c); the wild-type enzyme displayed a slightly higher activity up to pH 8.0, in contrast to the recombinant enzyme, which may be the impact of differences in glycosylation. Temperature activity profiles were almost identical, with the half-maximum and the highest enzyme activities at ca 40 and 55 °C, respectively (Fig. 4d). The enzyme was found relatively unstable at 50 °C (within 1 h its activity declined to ca 6%; Fig. S11, Supplementary Material). Therefore, 35 °C was selected as an optimum operating temperature for the enzymatic reactions, at which the enzyme showed perfect stability. Lower reaction temperatures are also recommendable in order to prevent spontaneous decomposition of sensitive compounds.

The Michaelis–Menten parameters of hydrolysis of *p*NP-GlcNAc substrate were determined both for wild-type and recombinant *Av*Hex as follows. For the wild-type *Av*Hex, $k_{cat} = 35.3 \pm 1.7\ s^{-1}$, $K_m = 0.22 \pm 0.06$ mM, and $k_{cat}/K_m = 160.5\ s^{-1}\ mM^{-1}$. For the recombinant *Av*Hex, $k_{cat} = 8.13 \pm 0.26\ s^{-1}$, $K_m = 0.18 \pm 0.03$ mM, and $k_{cat}/K_m = 45.2\ s^{-1}\ mM^{-1}$. The values of the Michaelis–Menten constants (K_m) of both enzymes were similar and in good agreement with the corresponding parameters published for other fungal β-*N*-acetylhexosaminidases (Plíhal et al. 2007; Ryšlavá et al. 2011; Slámová et al. 2012). Based on the k_{cat}/K_m values, the wild-type enzyme is a more efficient catalyst than its recombinant counterpart, which was also apparent from the values of specific activity (Table 1). This may be brought about by, e.g., differences in association of the catalytic domain with propeptide, which may depend on a particular cultivation and may slightly vary from batch to batch.

Additionally, sample transglycosylation reactions with *p*NP-GlcNAc substrate were performed with both wild-type and recombinant *Av*Hex, showing a similar reaction progress as monitored by TLC (data not shown). Obviously, *Av*Hex is able to perform transglycosylation reactions like many fungal β-*N*-acetylhexosaminidases (Bojarová and Křen 2011). In sum, we may conclude that the recombinant *Av*Hex produced in *P. pastoris* retained the properties of the wild-type fungal enzyme, and thus it is an equal synthetic alternative that offers the additional advantage of a high-yielding and elegant production.

Homology modeling and docking

For the sequence alignment used for model building, see Fig. S9 (Supplementary Material). The most stable protein structure is shown in Fig. 5. Both the propeptide and the loop HL2 showed reorientation already after 2 ns of refinement of monomeric structure while with dimer structures the initial conformation was preserved. This underlines the importance of dimer formation for maintaining the correct enzyme conformation during molecular dynamics and the positive impact of *N*-glycosylation in stabilizing the protein structure close to the active site.

The interdomain interactions after refinement are rather strong and comprise 23–24 hydrogen bonds. Loop HL1 amino acid residues participate just in one interdomain hydrogen bond interaction (Trp303 of one monomer and Asp599 of the other monomer). In contrast, the contribution of interdomain residues to loop HL2 stabilization appears as significant; its amino acid residues form 12–13 hydrogen bonds with the residues of the other domain.

The modeled structure of *Av*Hex consists of a propeptide (aa 19–92) followed by the catalytic domain, which includes a small zincin-like domain (aa 110 to approx.187) and a C-terminal $(\alpha/\beta)_8$ TIM-barrel (aa 188–608). The protein stretch containing aa 98–108 (PSSSVAAKAKR) corresponds to the segment cleaved during the activation process in the case of *Ao*Hex. This part could not be modeled properly because, presumably, the propeptide orientation changes after cleavage (Škerlová et al. 2018). Active-site amino acids include Asp353 and Glu354 (catalytic residues), hydrophobic Trp405 and Trp427, Trp525, and other residues forming hydrogen bonds with the substrate, i.e., Tyr453, Gln315, Arg201, Glu527, Asp455, and Trp490 (Fig. 6b, c).

In the refined model, the optimized orientation of the side chain of Gln315, the only non-conserved active-site residue, slightly differs from the position of the analogous residue Gln494 in *Sm*Chit (Fig. 5d). This may be due to a different orientation of HL1 in fungal β-*N*-acetylhexosaminidases (e.g., *Ao*Hex) compared to *Sm*Chit—the distance between C-α atoms of the respective residues Glu307 (*Ao*Hex) and Gln494 (*Sm*Chit (Tews et al. 1996)) is 0.86 Å. In fungal β-*N*-acetylhexosaminidases (Kulik et al. 2015; Škerlová et al. 2018), the analogous

```
1   MIYARLCTAL IAVASSVAAV KVNPLPAPRH IAWGSSGPRP LSDSVSLRCD QDTKDSILTN AWDRAWESIT SLEWVPAGIE
81  QPISSFEPFP TATASPTPSS SVAAKAKRTD DAPISFVEVE VDDWDADLQH GVDESYTLEA KDGSDTVQIT AKTVWGALHA
161 FTTFQQLVIS NGKGGLIIEQ PVQIKDAPLY PYRGIMIDTG RNFISVRKTY EQLDGMALSK LNVLHWHLDD TQSWPLHIDA
241 YPEMVNDAYS ARETFSHKDL RNVIAYARAR GIRVIPEIDM PAHSASGWQQ VDPEIVACAN SWWSNDNWEF HTAVQPNPGQ
321 LDIINPKTYE VVSKVYEELS GIFTDNWFHV GGDEIQPNCY NFSSYVTEWF AEDPSRTYND LMQHWIDESV PIFKSVSKSR
401 RLIMWEDVVL NTEHANDVPK DIVMQSWNNG LENIAKLTEA GYDTVVSSSD FMYLDCGRGG YVTNDDRYND QTNPDPETPS
481 FNYGGIGGSW CAPYKTWQRI YNYDFTLNLT KSQAEHVIGA TAPLWAEQVD DVNVSNMFWP RAAALAELVW SGNRDADGNK
561 RTNLFTQRLM NFREYLVANN VMAAIVAPKY CLQHPHACDL NLNQTILY
```

Fig 3 Sequence coverage of the hypothetical protein ASPVEDRAFT_125959 (GenBank: OJI97839.1) identified in the genome of *A. versicolor* CBS 583.65 by BLAST search in the GenBank Database as determined by LC-MS/MS analysis of the wild-type *Av*Hex. Peptide color code: grey, < 95% peptide identification confidence; green, ≥ 95% peptide identification confidence. The predicted *N*-terminal signal sequence of 19 amino acids is marked in bold black letters

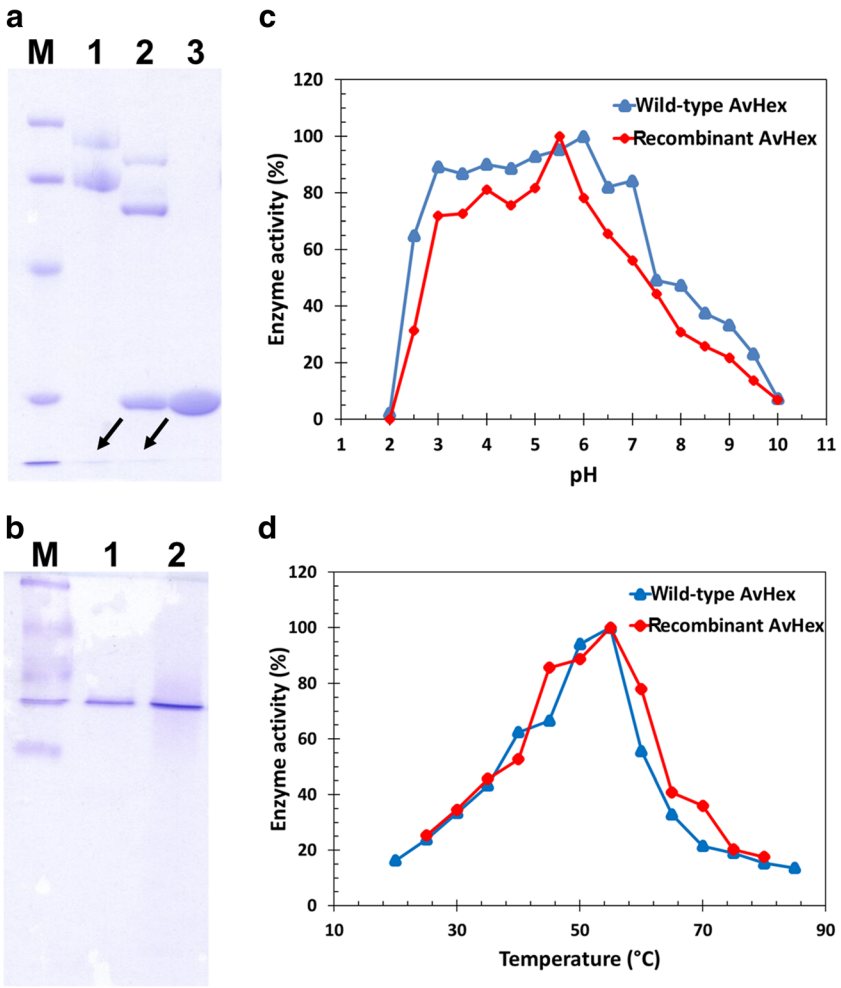


Fig. 4 **a** Purification and deglycosylation of recombinant *Av*Hex as analyzed by SDS-PAGE. Lane M, molecular mass markers consisting of phosphorylase b from rabbit muscle (97 kDa), bovine serum albumin (66 kDa), chicken egg white ovalbumin (45 kDa), carbonic anhydrase from bovine erythrocytes (30 kDa), trypsin inhibitor from soya beans (20.1 kDa), and α-lactalbumin from bovine milk (14.4 kDa); lane 1, recombinant *Av*Hex (two subunits of 78 and 64 kDa) with propeptide (14 kDa) purified by anion-exchange chromatography; lane 2, recombinant *Av*Hex after *N*-deglycosylation (two subunits of 72 and 58 kDa) with propeptide (14 kDa) and Endo H (29 kDa) and Endo H; lane 3, Endo H. The position of the putative propeptide co-purifying with the catalytic subunit is indicated with an arrow. **b** Native electrophoresis of *Av*Hex. Lane M, molecular mass markers consisting of porcine thyroid thyroglobulin (669 kDa), equine spleen ferritin (440 kDa), bovine liver catalase (232 kDa), bovine heart lactate dehydrogenase (140 kDa), and bovine serum albumin (66 kDa); lane 1, wild-type purified *Av*Hex; lane 2, recombinant purified *Av*Hex. **c** pH optimum of wild-type and recombinant *Av*Hex. **d** Temperature optimum of wild-type and recombinant *Av*Hex

Glu residue directly interacts with the carbohydrate substrate at the −1 position; in *Tf*Hex, it stabilizes the interaction with *p*NP-GalNAc substrate compared to GlcNAc by additional hydrogen bonding. Therefore, we hypothesize that the non-conserved residue Gln315 in *Av*Hex may contribute to its exceptional *N*-acetylglucosaminidase selectivity compared to other fungal β-*N*-acetylhexosaminidases that are quite promiscuous to both GlcNAc and GalNAc substrates.

Molecular docking

The refined model structure of *Av*Hex was used for analyzing the enzyme–substrate interactions by docking and molecular dynamics simulation. The results are shown in Table 2.

In the equilibrated models, we observed that all docked GlcNAc substrates, i.e., *p*NP-GlcNAc, GlcNAc-oxazoline (intermediate of substrate-assisted catalysis) (Slámová and Bojarová 2017), and GlcNAc monosaccharide, have a lower binding score than the respective GalNAc ligands, which corresponds to a more favorable interaction of GlcNAc with *Av*Hex than GalNAc. It is worthwhile to note that the hydrolytic product GlcNAc, which gets easily released from the active site, exhibits a similar binding score to *p*NP-GalNAc. This indicates that *p*NP-GalNAc has only weak binding interactions with the active-site residues. It is caused by the loss of hydrogen bond interactions with Arg201 and Glu527. On the other hand, a lower binding score for *p*NP-GlcNAc corresponds to strong binding interactions, which are brought about by a higher

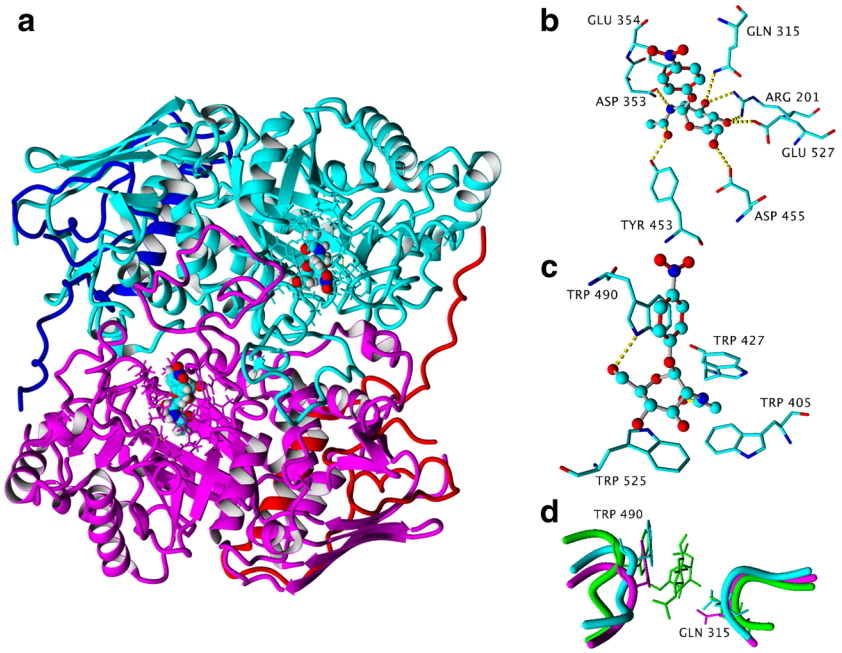


Fig. 5 Molecular model of *Av*Hex. **a** Protein dimer (cartoon representation) with *p*NP-GlcNAc substrate (stick representation) in the active site. Propeptide of monomer 1 is in **red**, catalytic domain of monomer 1 in **magenta**, propeptide of monomer 2 in **blue**, and catalytic domain of monomer 2 in **cyan**. **b**, **c** Active-site amino acid residues with docked *p*NP-GlcNAc after refinement (two different views). Trp residues forming Trp hydrophobic cavity for substrate binding are shown separately in panel **c** for clarity. Hydrogen bonds are depicted as yellow dotted lines. **d** Overlay of + 1 binding pocket segments in fungi (*Av*Hex in **cyan**, *Sm*Chit in **green**, and β-*N*-acetylhexosaminidase from *O. furnacalis* in **magenta**)

number of hydrogen bonds in the substrate–enzyme complex (namely, Arg201, Gln315, and Glu527, see Fig. 6).

To analyze the contribution of Gln315 to substrate binding, we calculated the Glide score for selected variants at the position of Gln315 (see Table S5, Supplementary Material). The binding score for *p*NP-GlcNAc bound in the active site of modeled Gln315Gly and Gln315Glu mutants significantly increased compared to the wild-type enzyme, which indicates the importance of Gln315 residue for the high affinity to *p*NP-GlcNAc. In contrast, in the case of *p*NP-GalNAc, the differences in binding affinity between the Gln315Gly and Gln315Glu mutants and the wild type were less significant.

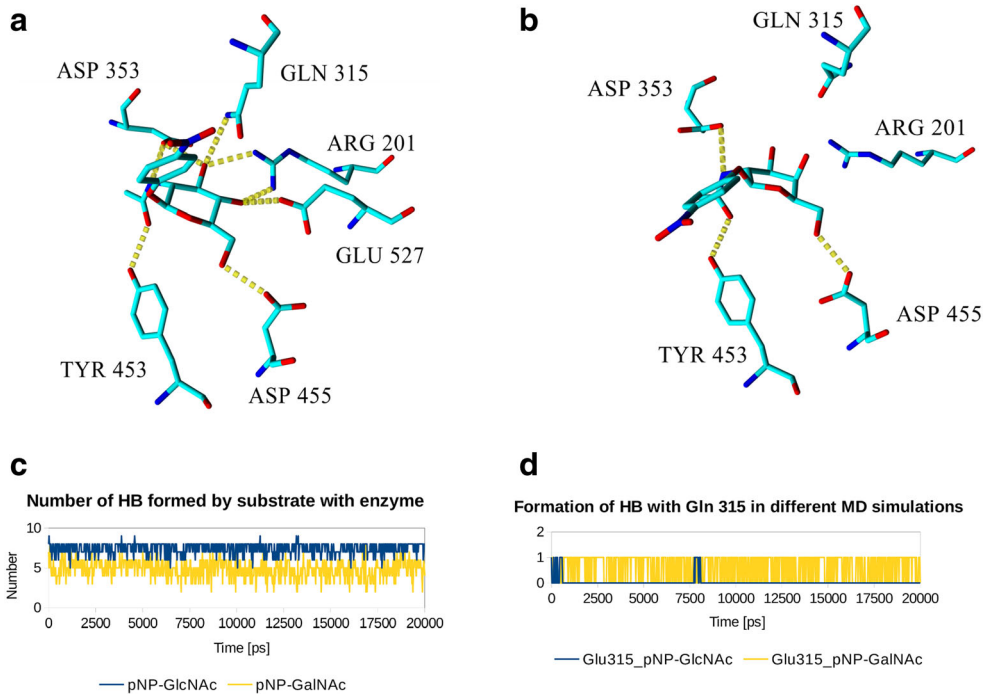


Fig. 6 **a** *p*NP-GlcNAc in the active site of *Av*Hex after 10 ns of molecular dynamics simulation. Hydrogen bonds are depicted as yellow dotted lines. **b** *p*NP-GalNAc in the active site of *Av*Hex after 10 ns of molecular dynamics simulation. **c** Number of hydrogen bonds over time formed with bound *p*NP-GlcNAc and *p*NP-GalNAc during the molecular dynamics simulation. **d** Time course of hydrogen bond formation between the substrate and Glu315; a value of 1 indicates the presence of a hydrogen bond; 0 indicates the absence of a hydrogen bond

Table 2 Averaged Glide XP binding scores for GlcNAc and GalNAc substrates in *Av*Hex

Substrate	Glide score (kJ/mol)[a]
*p*NP-GlcNAc	− 36.8
*p*NP-GalNAc	− 31.7
GlcNAc-oxazoline[b]	− 40.0
GalNAc-oxazoline[b]	− 35.9
GlcNAc	− 30.6
GalNAc	− 27.5

[a] Glide score after molecular dynamics simulation for equilibrated substrate–enzyme complex. Glide score represents the free energy of binding. The lower score, the stronger binding

[b] For the systematic names and structures of oxazolines, see Scheme S1, Supplementary Material

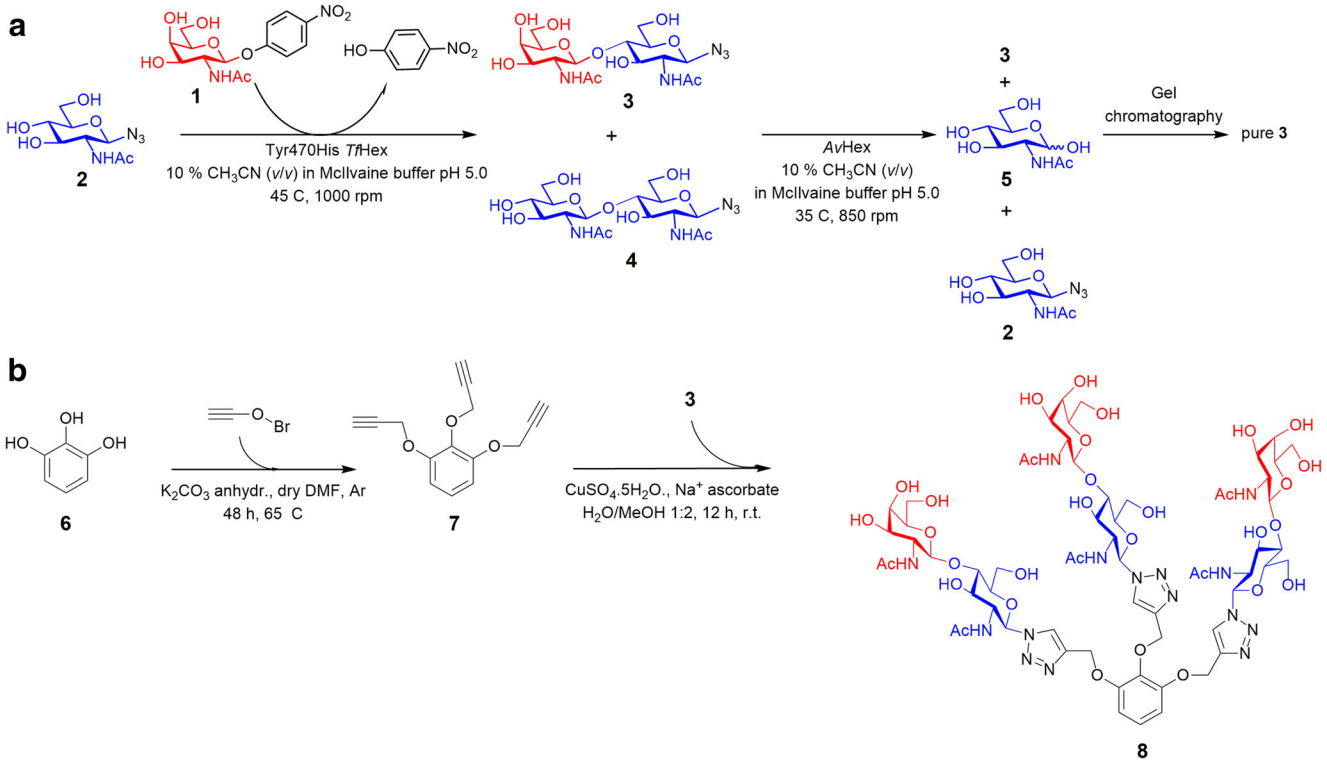
Further, molecular dynamics simulations with the Gln315Glu mutant showed that *p*NP-GalNAc lost its hydrogen bond with Arg201, which was compensated by the formation of a new hydrogen bond between the substrate C-4 hydroxyl and Glu315; a similar situation was previously observed in *Tf*Hex, which is a promiscuous β-*N*-acetylhexosaminidase with a GlcNAcase/GalNAcase ratio of 1/1.2. Therefore, we may suppose that the Gln315Glu mutation decreases the difference in binding affinities to GalNAc and GlcNAc substrates.

A notable issue is the presence of Tyr461 close to Gln315. In ca 85% of sequences of homologous fungal β-*N*-acetylhexosaminidases and β-*N*-acetylglucosaminidases containing Gln in the position corresponding to Gln315 in *Av*Hex, Tyr is found in the position corresponding to Tyr461 (see Table S4, Supplementary Material). The hydroxyl hydrogen of Tyr461 makes a hydrogen bond with the carboxyl oxygen of Gln315 and it may form a water-mediated hydrogen bond with the C-4 oxygen of GlcNAc-type substrate (equatorial hydroxyl), which may result in additional stabilization of this configuration.

Application of *Av*Hex in a sequential synthesis of a functionalized disaccharide

We employed the selective *Av*Hex enzyme in a sequential enzymatic synthesis of a bioactive disaccharide, β-D-GalNAc-(1 → 4)-β-D-GlcNAc-N_3 that carries an azido function at C-1 (Scheme 1a, compound **3**), using β-D-GlcNAc azide (**2**) as an acceptor (Fialová et al. 2005).

The azido-terminated disaccharide **3** is quite difficult to prepare using glycosyltransferases (Bojarová et al. 2013), which are sensitive to changes in both donor and acceptor substrate molecules; other linkers at C-1 are much more synthetically feasible (Bumba et al. 2018). This disaccharide may be prepared using site-directed mutants of *Tf*Hex (i.e., Tyr470His), which exhibit increased transglycosylation potential (Slámová et al. 2015). However, due to their dual affinity to both GalNAc and GlcNAc substrates combined with their broad substrate specificity, this reaction is complicated by the formation of



Scheme 1 **a** Sequential synthesis of functionalized disaccharide **3** employing Tyr470His *Tf*Hex and *Av*Hex. **b** Synthesis of trivalent glycoconjugate **8** using functionalized disaccharide **3** and trivalent carrier **7** by Cu(I)-catalyzed azide-alkyne cycloaddition

Table 3 Inhibition constants for glycoconjugate **8** and standard carbohydrates with galectin-3 determined by ELISA

Inhibitor	IC_{50} (μM)	Relative potency	Relative potency per glycan
Lactose (Galβ4Glc)	153 ± 24	0.36	0.36
GalNAcβ4GlcNAc	55 ± 11	1	1
Glycoconjugate **8**	9.6 ± 3.2	5.7	1.9

chitobiosyl azide (β-D-GlcNAc-(1 → 4)-β-D-GlcNAc-N_3; **4**) as an unwanted by-product. This compound originates in an enzymatic side reaction by processing the GlcNAc azide acceptor (**2**) and it may be contained in the ratio of up to 1:4 to the desired product **3**. Glycosyl azides have been shown to act as glycosyl donors with many glycosidases (Bojarová et al. 2007). Naturally, the separation of both disaccharides differing only in one axial-equatorial position is practically impossible by conventional methods such as HPLC or gel chromatography (Bojarová et al. 2008). As a result, the synthetic reaction occurs with a lower isolated yield and the product may have lower purity. The employment of a sequential reaction presented herein elegantly solves this synthetic problem since *Av*Hex efficiently cleaves the unwanted chitobiosyl azide **4**. Though *Av*Hex has some minor GalNAcase activity (GlcNAcase/GalNAcase ratio of ca 9–11), we did not observe any cleavage of our desired product **3** during the sequential reaction. In the present reaction setup, we were able to increase the yield of the desired disaccharide **3** by ca 30%, reaching a 47% overall isolated yield. The structure of disaccharide **3** was confirmed by HRMS and ^{1}H and ^{13}C NMR.

Though we tested other established β-*N*-acetylglucosaminidases for the selective hydrolysis of **4**, e.g., that from *Bacteroides thetaiotamicron* or recombinant human *O*-GlcNAcase, neither of the tested enzymes exhibited affinity to chitobiose-based oligosaccharides, apparently due to their selectivity for GlcNAc bound to a protein scaffold. Therefore, we consider *Av*Hex a very useful enzyme for biotechnological applications involving tailored carbohydrate synthesis.

Preparation of bioactive conjugate 8 by click chemistry and its binding to galectin-3

The functionalized GalNAcβ4GlcNAc epitope **3** was conjugated to a simple aromatic carrier **7** based on pyrogallol (**6**), carrying three propargyl moieties (Scheme 1b). Compound **7** was prepared according to a published procedure (Xie et al. 2008). The conjugation with carbohydrate **3** was performed by Cu(I)-catalyzed azide-alkyne cycloaddition in water/methanol mixture to yield glycoconjugate **8**. Though similar conjugates have been tested with galectins (Wang et al. 2012), glycoconjugate **8** is the first dendrimer-like multivalent glycoconjugate that carries the GalNAcβ4GlcNAc epitope selective for galectin-3. Compound **8** was isolated in 53% yield by gel chromatography; its binding affinity to galectin-3 was determined in a competitive ELISA-type assay using asialofetuin as a standard immobilized competitor (Laaf et al. 2017b) and compared to the free monovalent GalNAcβ4GlcNAc. The results are summarized in Table 3 and the respective dose-dependent inhibition curves are shown in Fig. S12 (Supplementary Material). We can see that the presentation of the selective epitope on the aromatic carrier via the triazole linker enhanced the conjugate affinity 5.7-fold compared to the monovalent disaccharide, and this mode of presentation has increased the relative inhibitory potency per glycan 1.9-fold. This interaction is purely specific and is not caused by potential non-specific hydrophobic forces as demonstrated in a control assay in the presence of 0.025% *v*/*v* Tween, giving the same IC_{50} values. In sum, conjugate **8** is a good leading example for the construction of glycomimetics based on aromatic scaffolds.

Discussion

GalNAc-terminated oligosaccharides are excellent ligands of galectin-3, a soluble monomeric human lectin of chimera type with a high biomedical relevance. Contrary to analogous epitopes terminated with β-D-galactopyranoside, they exhibit exclusive selectivity for galectin-3 compared to a similarly frequent galectin-1 (Šimonová et al. 2014), which is crucial for potential biomedical applications (Bojarová and Křen 2016). *N*-Acetylhexosamine carbohydrates with various functional groups in the molecule are advantageously synthesized using robust and stable glycosidases with a broad substrate specificity. If an azido moiety is used for conjugation to a carrier by click chemistry, the formed triazol supports binding to galectin-3 by additional interactions with its binding site (Bojarová et al. 2018). Moreover, since the enzymatic glycosylation of multivalent precursors by both glycosidases and glycosyltransferases does not proceed easily and mostly affords partially glycosylated by-products (Drozdová et al. 2011), a straightforward conjugation of functionalized oligosaccharides by click chemistry is a very attractive solution. Nevertheless, azido-functionalized GalNAc-terminated epitopes are quite challenging to prepare by glycosyltransferases due to their limited tolerance to structural modifications in the donor and acceptor molecule. Therefore, a high-yielding synthetic route to the selective azido-bearing GalNAcβ4GlcNAc epitope, using a combination of a previously reported Tyr470His variant of *Tf*Hex (Slámová et al. 2015) and a novel highly selective *Av*Hex, is an attractive option for the design of

galectin-3 inhibitors. The novel trivalent glycoconjugate **8** shows good affinity to galectin-3 and as such may serve as a leading structure for the construction of advanced multivalent selective galectin-3 inhibitors.

Since the dual activity towards both GlcNAc and GalNAc substrates is an intrinsic characteristic of β-*N*-acetylhexosaminidases, the selectivity of the present *Av*Hex is an exceptional catalytic feature. Notably, other GlcNAc-selective enzymes known, such as GH84 *O*-GlcNAcase (Macauley et al. 2005), are not applicable for the processing of *N*-acetylhexosamine oligosaccharides, nor do they show any transglycosylation capabilities. Both of these properties, together with a broad pH optimum, a good tolerance of functionalized carbohydrate substrates and excellent long-term storage stability, predestine *Av*Hex for biotechological applications.

Molecular modeling analysis and comparison to the existing β-*N*-acetylhexosaminidase counterparts, in particular *Ao*Hex, helped to decode the relations in *Av*Hex binding site and brought a possible explanation for its exceptional selectivity. Residue Gln315, possibly in combination with Tyr461, was identified as a probable source of the exceptional selectivity of *Av*Hex to substrates in GlcNAc configuration. It follows from the extensive alignment study of the existing homologous enzymes as well as from the computational modeling of mutants that this particular combination of amino acid residues may influence the GlcNAcase/GalNAcase activity ratio in β-*N*-acetylhexosaminidases. Recombinant expression in *P. pastoris*, recently identified as an optimum heterologous producer of fungal β-*N*-acetylhexosaminidases (Slámová et al. 2012), enabled to enhance the production of *Av*Hex ca 20-fold whereas its biocatalytic properties and selectivity were maintained. In conclusion, the recombinant *Av*Hex presented in this work is a highly useful tool for a range of biotechnological applications, especially involving the synthesis and/or processing of *N*-acetylhexosamine oligosaccharides that comprise a combination of both β-GlcNAc and β-GalNAc residues. An example of such carbohydrates is a selective ligand of galectins presented herein.

Acknowledgments P. B. and V. K. acknowledge support by mobility projects no. LTC18038 and LTC18041 (MEYS, the Ministry of Education, Youth and Sports of the Czech Republic). N. K. acknowledges access to the computing and storage facilities provided by the CESNET LM2015042 and the CERIT Scientific Cloud LM2015085 under the program "Projects of Large Research, Development, and Innovations Infrastructures."

Funding This study was funded by the Ministry of Education, Youth, and Sports of the Czech Republic mobility projects nos. LTC18038 and LTC18041, by CESNET (LM2015042), and by CERIT Scientific Cloud (LM2015085) under the program "Projects of Large Research, Development, and Innovations Infrastructures."

Compliance with ethical standards This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest The authors declare that they have no conflict of interest.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Bohne-Lang A, von der Lieth CW (2005) GlyProt: in silico glycosylation of proteins. Nucleic Acids Res 33:W214–W219. https://doi.org/10.1093/nar/gki385

Bojarová P, Křen V (2009) Glycosidases: a key to tailored carbohydrates. Trends Biotechnol 27:199–209. https://doi.org/10.1016/j.tibtech.2008.12.003

Bojarová P, Křen V (2011) Glycosidases in carbohydrate synthesis: when organic chemistry falls short. Chimia 65:65–70. https://doi.org/10.2533/chimia.2011.65

Bojarová P, Křen V (2016) Sugared biomaterial binding lectins: achievements and perspectives. Biomat Sci 4:1142–1160. https://doi.org/10.1039/c6bm00088f

Bojarová P, Petrásková L, Ferrandi E, Monti D, Pelantová H, Kuzma M, Simerská P, Křen V (2007) Glycosyl azides – an alternative way to disaccharides. Adv Synth Catal 349:1514–1520

Bojarová P, Křenek K, Kuzma M, Petrásková L, Bezouška K, Namdjou D-J, Elling L, Křen V (2008) *N*-Acetylhexosamine triad in one molecule: chemoenzymatic introduction of 2-acetamido-2-deoxy-β-D-galactopyranosyluronic acid residue into a complex oligosaccharide. J Mol Catal B Enzym 50:69–73. https://doi.org/10.1016/j.molcatb.2007.09.002

Bojarová P, Slámová K, Křenek K, Gažák R, Kulik N, Ettrich R, Pelantová H, Kuzma M, Riva S, Adámek D, Bezouška K, Křen V (2011) Charged hexosaminides as new substrates for β-*N*-scetylhexosaminidase-catalyzed synthesis of immunomodulatory disaccharides. Adv Synth Catal 353:2409–2420. https://doi.org/10.1002/adsc.201100371

Bojarová P, Rosencrantz RR, Elling L, Křen V (2013) Enzymatic glycosylation of multivalent scaffolds. Chem Soc Rev 42:4774–4797. https://doi.org/10.1039/c2cs35395d

Bojarová P, Chytil P, Mikulová B, Bumba L, Konefal R, Pelantová H, Krejzová J, Slámová K, Petrásková L, Kotrchová L, Cvačka J, Etrych T, Křen V (2017) Glycan-decorated HPMA copolymers as high-affinity lectin ligands. Polym Chem 8:2647–2658. https://doi.org/10.1039/C7PY00271H

Bojarová P, Tavares MR, Laaf D, Bumba L, Petrásková L, Konefał R, Bláhová M, Pelantová H, Elling L, Etrych T, Chytil P, Křen V (2018) Biocompatible glyconanomaterials based on HPMA-copolymer for specific targeting of galectin-3. J Nanobiotechnol 16:73. https://doi.org/10.1186/s12951-018-0399-1

Bradford M (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254. https://doi.org/10.1016/0003-2697(76)90527-3

Bumba L, Laaf D, Spiwok V, Elling L, Křen V, Bojarová P (2018) Poly-*N*-acetyllactosamine neo-glycoproteins as nanomolar ligands of

human galectin-3: binding kinetics and modeling. Int J Mol Sci 19: 372. https://doi.org/10.3390/ijms19020372

Chuang G-Y, Boyington JC, Joyce MG, Zhu J, Nabel GJ, Kwong PD, Georgiev I (2012) Computational prediction of *N*-linked glycosylation incorporating structural properties and patterns. Bioinformatics 28:2249–2255. https://doi.org/10.1093/bioinformatics/bts426

Clamp M, Cuff J, Searle SM, Barton GJ (2004) The Jalview Java alignment editor. Bioinformatics 20:426–427. https://doi.org/10.1093/bioinformatics/btg430

Davis IW, LeaverFay A, Chen VB, Block JN, Kapral GJ, Wang X, Murray LW, Arendall WB 3rd, Snoeyink J, Richardson JS, Richardson DC (2007) MolProbity: allatom contacts and structure validation. Nucleic Acids Res 35:W375–W383. https://doi.org/10.1093/nar/gkm216.

Drozdová A, Bojarová P, Křenek K, Weignerová L, Henßen B, Elling L, Christensen H, Jensen HH, Pelantová H, Kuzma M, Bezouška K, Krupová M, Adámek D, Slámová K, Křen V (2011) Enzymatic synthesis of dimeric glycomimetic ligands of NK cell activation receptors. Carbohydr Res 346:1599–1609. https://doi.org/10.1016/j.carres.2011.04.043

Ettrich R, Kopecký V Jr, Hofbauerová K, Baumruk V, Novák P, Pompach P, Man P, Plíhal O, Kutý M, Kulik N, Sklenář J, Ryšlavá H, Křen V, Bezouška K (2007) Structure of the dimeric *N*-glycosylated form of fungal β-*N*-acetylhexosaminidase revealed by computer modeling, vibrational spectroscopy, and biochemical studies. BMC Struct Biol 7:32–45. https://doi.org/10.1186/1472-6807-7-32

Fialová P, Weignerová L, Rauvolfová J, Přikrylová V, Pišvejcová A, Ettrich R, Kuzma M, Sedmera P, Křen V (2004) Hydrolytic and transglycosylation reactions of *N*-acyl modified substrates catalysed by β-*N*-acetylhexosaminidases. Tetrahedron 60:693–701. https://doi.org/10.1016/j.tet.2003.10.111

Fialová P, Carmona AT, Robina I, Ettrich R, Sedmera P, Přikrylová V, Petrásková-Hušáková L, Křen V (2005) Glycosyl azide - a novel substrate for enzymatic transglycosylations. Tetrahedron Lett 46: 8715–8718. https://doi.org/10.1016/j.tetlet.2005.10.040

Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC, Mainz DT (2006) Extra precision Glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. J Med Chem 49:6177–6196. https://doi.org/10.1021/jm051256o

Gloster TM, Vocadlo DJ (2010) Mechanism, structure, and inhibition of *O*-GlcNAc processing enzymes. Curr Signal Transduct Ther 5:74–91. https://doi.org/10.2174/157436210790226537

Huňková Z, Křen V, Ščigelová M, Weignerová L, Scheel O, Thiem J (1996) Induction of β-*N*-acetylhexosaminidase in *Aspergillus oryzae*. Biotechnol Lett 18:725–730. https://doi.org/10.1007/BF00130773

Jakalian A, Jack DB, Bayly CI (2002) Fast, efficient generation of high quality atomic charges. AM1BCC model: II. Parameterization and validation. J Comput Chem 23:1623–1641. https://doi.org/10.1002/jcc.10128

Krieger E, Koraimann G, Vriend G (2002) Increasing the precision of comparative models with YASARA NOVA a self parameterizing force field. Proteins 47:393–402. https://doi.org/10.1002/prot.10104

Kulik N, Slámová K, Ettrich R, Křen V (2015) Computational study of β-*N*-acetylhexosaminidase from *Talaromyces flavus*, a glycosidase with high substrate flexibility. BMC Bioinformatics 16:1–15. https://doi.org/10.1186/s12859-015-0465-8.

Laaf D, Bojarová P, Mikulová B, Pelantová H, Křen V, Elling L (2017a) Two-step enzymatic synthesis of β-D-*N*-acetylgalactosamine-(1-4)-D-*N*-acetylglucosamine (LacdiNAc) chitooligomers for deciphering galectin binding behavior. Adv Synth Catal 359:2101–2108. https://doi.org/10.1002/adsc.201700331

Laaf D, Bojarová P, Pelantová H, Křen V, Elling L (2017b) Tailored multivalent neo-glycoproteins: synthesis, evaluation, and application of a library of galectin-3-binding glycan ligands. Bioconjug Chem 28:2832–2840. https://doi.org/10.1021/acs.bioconjchem.7b00520

Macauley MS, Whitworth GE, Debowski AW, Chin D, Vocadlo DJ (2005) *O*-GlcNAcase uses substrate-assisted catalysis; kinetic analysis and development of highly selective mechanism-inspired inhibitors. J Biol Chem 280:25313–25,322. https://doi.org/10.1074/jbc.M413819200

Notredame C, Higgins DG, Heringa J (2000) T-coffee: a novel method for fast and accurate multiple sequence alignment. J Mol Biol 302:205–217. https://doi.org/10.1006/jmbi.2000.4042.

Plíhal O, Sklenář J, Hofbauerová K, Novák P, Man P, Pompach P, Ryšlavá H, Charvátová-Pišvejcová A, Křen V, Bezouška K (2007) Large propeptides of fungal β-*N*-acetylhexosaminidases are novel enzyme regulators that must be processed intracellularly to control activity, dimerization, and secretion into the extracellular environment. Biochemistry 46:2719–2734. https://doi.org/10.1021/bi061828m

Raval A, Piana S, Eastwood MP, Dror RO, Shaw DE (2012) Refinement of protein structure homology models via long, all-atom molecular dynamics simulations. Proteins 80:2071–2079. https://doi.org/10.1002/prot.24098.

Robert X, Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res 42:W320–W324. https://doi.org/10.1093/nar/gku316

Ryšlavá H, Kalendová A, Doubnerová V, Skočdopol P, Kumar V, Kukačka Z, Pompach P, Vaněk O, Slámová K, Bojarová P, Kulík N, Ettrich R, Křen V, Bezouška K (2011) Enzymatic characterization and molecular modeling of an evolutionarily interesting fungal β-*N*-acetylhexosaminidase. FEBS J 278:2469–2484. https://doi.org/10.1111/j.1742-4658.2011.08173.x

Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234:779–815. https://doi.org/10.1006/jmbi.1993.1626

Šimonová A, Kupper CE, Böcker S, Müller A, Hofbauerová K, Pelantová H, Elling L, Křen V, Bojarová P (2014) Chemoenzymatic synthesis of LacdiNAc dimers of varying length as novel galectin ligands. J Mol Catal B Enzym 101:47–55. https://doi.org/10.1016/j.molcatb.2013.12.018

Škerlová J, Bláha J, Pachl P, Hofbauerová K, Kukačka Z, Man P, Pompach P, Novák P, Otwinowski Z, Brynda J, Vaněk O, Řezáčová P (2018) Crystal structure of native β-*N*-acetylhexosaminidase isolated from *Aspergillus oryzae* sheds light onto its substrate specificity, high stability, and regulation by propeptide. FEBS J 285:580–598. https://doi.org/10.1111/febs.14360

Slámová K, Bojarová P (2017) Engineered *N*-acetylhexosamine-active enzymes in glycoscience. Biochim Biophys Acta Gen Subj 1861: 2070–2087. https://doi.org/10.1016/j.bbagen.2017.03.019

Slámová K, Bojarová P, Petrásková L, Křen V (2010) β-*N*-Acetylhexosaminidase: what's in a name...? Biotechnol Adv 28: 682–693. https://doi.org/10.1016/j.biotechadv.2010.04.004

Slámová K, Bojarová P, Gerstorferová D, Fliedrová B, Hofmeisterová J, Fiala M, Pompach P, Křen V (2012) Sequencing, cloning and high-yield expression of a fungal β-*N*-acetylhexosaminidase in *Pichia pastoris*. Prot Express Purif 82:212–217. https://doi.org/10.1016/j.pep.2012.01.004

Slámová K, Krejzová J, Marhol P, Kalachova L, Kulik N, Pelantová H, Cvačka J, Křen V (2015) Synthesis of derivatized chitooligomers using transglycosidases engineered from the fungal GH20 β-*N*-acetylhexosaminidase. Adv Synth Catal 357:1941–1950. https://doi.org/10.1002/adsc.201500075

Tews I, Perrakis A, Oppenheim A, Dauter Z, Wilson KS, Vorgias CE (1996) Bacterial chitobiase structure provides insight into catalytic mechanism and the basis of Tay-Sachs disease. Nat Struct Biol 3: 638–648. https://doi.org/10.1038/nsb0796-638

Wang GN, André S, Gabius HJ, Murphy PV (2012) Bi- to tetravalent glycoclusters: synthesis, structure-activity profiles as lectin inhibitors and impact of combining both valency and headgroup tailoring on selectivity. Org Biomol Chem 10:6893–6907. https://doi.org/10.1039/c2ob25870f

Weignerová L, Vavrušková P, Pišvejcová A, Thiem J, Křen V (2003) Fungal β-*N*-acetylhexosaminidases with high β-*N*-acetylgalactosaminidase activity and their use for synthesis of β-GalNAc-containing oligosaccharides. Carbohydr Res 338:1003–1008. https://doi.org/10.1016/S0008-6215(03)00044-2

Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35:W407–W410. https://doi.org/10.1093/nar/gkm290

Xie J, Hu L, Shi W, Deng X, Cao Z, Shen Q (2008) Synthesis and characterization of hyperbranched polytriazole via an 'A2 + B3' approach based on click chemistry. Polym Int 57:965–974. https://doi.org/10.1002/pi.2433