the FEBS Journal

Enzymatic characterization and molecular modeling of an evolutionarily interesting fungal β-*N*-acetylhexosaminidase

Helena Ryšlavá[1], Alžběta Kalendová[1], Veronika Doubnerová[1], Přemysl Skočdopol[1], Vinay Kumar[1,2], Zdeněk Kukačka[1], Petr Pompach[1,3], Ondřej Vaněk[1,3], Kristýna Slámová[3], Pavla Bojarová[3], Natallia Kulik[4], Rudiger Ettrich[4], Vladimír Křen[3] and Karel Bezouška[1,3]

1 Department of Biochemistry, Faculty of Science, Charles University Prague, Czech Republic
2 Department of Tropical Medicine, School of Public Health and Tropical Medicine, Tulane University, New Orleans, LA, USA
3 Institute of Microbiology, Academy of Sciences of the Czech Republic, Prague, Czech Republic
4 Department of Structure and Function of Proteins, Institute of Nanobiology and Structural Biology of GCRC, Academy of Sciences of the Czech Republic and Faculty of Sciences of the University of South Bohemia, Nové Hrady, Czech Republic

Keywords
deglycosylation; enzyme kinetics; hexosaminidase; molecular dynamics; molecular modeling

Correspondence
K. Bezouška, Department of Biochemistry, Faculty of Science, Charles University Prague, Czech Republic
Fax: +420 22195 1283
Tel: +420 2 2195 1272
E-mail: bezouska@biomed.cas.cz

(Received 21 December 2010, revised 29 April 2011, accepted 9 May 2011)

doi:10.1111/j.1742-4658.2011.08173.x

Fungal β-*N*-acetylhexosaminidases are inducible extracellular enzymes with many biotechnological applications. The enzyme from *Penicillium oxalicum* has unique enzymatic properties despite its close evolutionary relationship with other fungal hexosaminidases. It has high GalNAcase activity, tolerates substrates with the modified N-acyl group better and has some other unusual catalytic properties. In order to understand these features, we performed isolation, biochemical and enzymological characterization, molecular cloning and molecular modelling. The native enzyme is composed of two catalytic units (65 kDa each) and two propeptides (15 kDa each), yielding a molecular weight of 160 kDa. Enzyme deglycosylated by endoglycosidase H had comparable activity, but reduced stability. We have cloned and sequenced the gene coding for the entire hexosaminidase from *P. oxalicum*. Sufficient sequence identity of this hexosaminidase with the structurally solved enzymes from bacteria and humans with complete conservation of all catalytic residues allowed us to construct a molecular model of the enzyme. Results from molecular dynamics simulations and substrate docking supported the experimental kinetic and substrate specificity data and provided a molecular explanation for why the hexosaminidase from *P. oxalicum* is unique among the family of fungal hexosaminidases.

Enzymes
hexosaminidase, β-*N*-acetyl-D-hexosaminide *N*-acetylhexosaminhydrolase, EC 3.2.1.52

Introduction

β-*N*-Acetylhexosaminidases (EC 3.2.1.52, Hex) are enzymes that hydrolyse the terminal β-D-GlcNAc and β-D-GalNAc residues of oligosaccharide chains [1]. They have been extensively studied in higher vertebrates (including humans [2]) and bacteria [3,4]. Fungal Hex have recently attracted considerable attention because of their biology, architecture and biotechnological applications. These enzymes are involved in the binary chitinolytic system responsible for degrading chitooligomers and chitobiose [5], which is important

Abbreviations
CCF, Culture Collection of Fungi; Endo H, endoglycosidase H; Hex, hexosaminidase; HMM, hidden Markov model; MD, molecular dynamics; 4-MU-GlcNAc, 4-methylumbelliferyl 2-acetamido-2-deoxy-β-D-glucopyranoside; PNGase F, peptide:*N*-glycosidase F; *p*NP-GalNAc, *p*-nitrophenyl 2-acetamido-2-deoxy-β-D-galactopyranoside; *p*NP-GlcNAc, *p*-nitrophenyl 2-acetamido-2-deoxy-β-D-glucopyranoside; PoHex, hexosaminidase from *Penicillium oxalicum*.

for fungal cell wall regeneration and hyfae formation [6]. Biotechnologically, they have found use in the syntheses of new oligosaccharides by means of transglycosylation reactions [1,7,8]. Fungal Hex possess a notable enzyme architecture, in which catalytic subunits combine with large propeptides [9]. The propeptide in Hex from *Aspergillus oryzae* has recently been characterized and shown to represent a novel intracellular regulator that controls enzyme activity, dimerization and extracellular secretion [10].

Previous experiments were performed with crude ammonium sulfate precipitates of Hex from several strains of *Penicillium oxalicum* (PoHex). These preliminary studies revealed the unique properties of these enzymes. First, while many fungal Hex possess both β-*N*-acetylgalactosaminidase and β-*N*-acetylglucosaminidase activities, the *P. oxalicum* enzyme has the highest prevalence of the former activity in this entire enzyme family [11]. Second, this enzyme has the unique ability to readily cleave substrates that bear various chemical modifications such as N-acyls other than N-acetyl [12], substrates substituted at C6 [13,14] or even 4-deoxy substrates [15]. Third, the hexosaminidases from *Penicillium* species [16–19] have unusual pH stability and pH optima [18,19] and possess some other unique properties. The aims of the present study were (a) to verify these properties using highly purified enzymes devoid of the contaminants that could be present in crude enzyme preparations; (b) to study the details of enzyme kinetics not investigated previously; (c) to probe the molecular mechanisms behind these unique features; and (d) to correlate catalytic behaviour of the enzyme with the specific features of its three-dimensional structure and evolution.

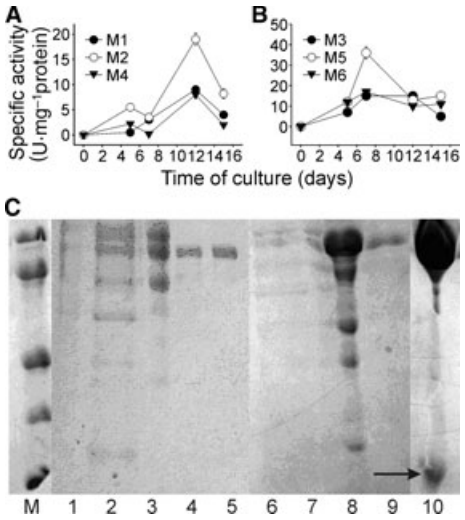

Fig. 1. Optimization of PoHex production and purification. (A), (B) Time course of secretion of PoHex from *Penicillium oxalicum* strains CCF 1959 and CCF 3438, respectively, in different media (M1–M6). The best production was achieved for the CCF 3438 strain cultivated in medium M5 made up of (per L) 0.2 g $NaNO_3$, 0.05 g KCl, 0.001 g $FeSO_4$, 0.1 g KH_2PO_4, 1.0 g GlcNAc, 0.5 g $MgSO_4$, pH 4.5. Other cultivations were performed as described in the experimental section. (C) Purification of PoHex from *P. oxalicum* strains CCF 1959 (lanes 1–5) and CCF 3438 (lanes 6–10) was monitored by SDS/PAGE. Lane M, molecular mass markers consisting of BSA (67 000), ovalbumin (45 000), trypsinogen (24 000), β-lactoglobulin (18 000) and lysozyme (14 000); lane 1, culture medium; lanes 2, 6, ammonium sulfate precipitate; lanes 3, 7, hexosaminidase purified by phenyl-Sepharose chromatography; lanes 4, 8, hexosaminidase purified by MonoQ chromatography; lanes 5, 9, final preparation after purification by gel filtration on Superdex 200; lane 10, same as lane 9 but 30 times as much protein loaded. The position of the putative propeptide co-purifying with the catalytic subunit is indicated by an arrow.

Results

Production, purification and characterization of PoHex from strains CCF 1959 and 3438

Secretion of Hex into media is typically a biphasic process [10,20]. For *P. oxalicum* Culture Collection of Fungi (CCF) 1959, the highest level of specific activity was achieved after 12 days of cultivation, whereas it was 7 days for the CCF 3438 strain (Fig. 1A,B). Homogeneous PoHex preparations were obtained by a combination of hydrophobic, anion exchange and size-exclusion chromatographies (Table S1). The purified enzyme was free of activities from other contaminating glycosidases (Table S2). It appeared to be homogeneous after SDS electrophoresis using the discontinuous buffer system of Laemmli [21] (Fig. 1C, lanes 5 and 9, respectively). Another protein band with a molecular weight of approximately 15 kDa that was co-purified with the enzyme could be found in heavily overloaded samples (Fig. 1C, lane 10) and based on data on other hexosaminidases [10] was tentatively assigned to the PoHex propeptide. Thirty cycles of N-terminal sequencing of the 65 kDa polypeptide yielded a sequence of DTAATAIHSVHLSVDAAXDLQHGVDESYTK. The analysis of the 15 kDa protein band provided a sequence of VKVNPLPAPRNITWGSSGPISITKPALHLE. These sequences were identical for both strains and displayed the highest homology with the N-terminal regions of the Hex precursors from filamentous fungi of *Aspergillus terreus*, *Penicillium chrysogenum* and *Aspergillus niger*. The native size of the PoHex was found to be approximately 160 kDa, as determined by gel filtration and native electrophoresis [22].

Despite their apparently identical primary structure, both hexosaminidases displayed vast differences in their specific activities after purification (10.8 and 35.6 U·mg^{-1} protein for the Hex from CCF 1959 and CCF 3438, respectively) (Table S1). The analysis of propeptide occurrence in the two preparations by Edman degradation [10] revealed that the CCF 3438 enzyme comprised an equimolar amount of the two polypeptides, indicating the presence of two propeptides per enzyme dimer. On the other hand, the propeptide content in the CCF 1959 enzyme (on a molar basis) was only about a third of that of the catalytic unit.

Enzymatic properties of the PoHex

The Hex from both strains of *P. oxalicum* (CCF 1959 and CCF 3438) displayed a broad pH optimum of 2–4, with a maximum at pH 3, using both *p*-nitrophenyl 2-acetamido-2-deoxy-β-D-glucopyranoside (*p*NP-GlcNAc) and *p*-nitrophenyl 2-acetamido-2-deoxy-β-D-galactopyranoside (*p*NP-GalNAc) substrates. The enzymes were more stable at neutral pH than at acidic pH. The activity of PoHex increased linearly between 15 and 50 °C (maximum); at higher temperatures, there was a rapid decrease in activity.

The PoHex activity was affected by salts. $(NH_4)_2SO_4$ and $MgSO_4$ decreased the β-GlcNAcase activity, but β-GalNAcase activity was slightly stimulated (Fig. 2A,B). $MgCl_2$ did not have the same effect (not shown).

The kinetics of PoHex with both substrates were studied in detail. Whereas the dependence of the reaction rate on *p*NP-GalNAc concentration was hyperbolic, when *p*NP-GlcNAc was used as substrate, inhibition due to excess of substrate was observed (Fig. 2C,D). Michaelis constants (K_m), substrate inhibition constants (K_{ss}), catalytic constants (k_{cat}) and catalytic efficiency (k_{cat}/K_m) for PoHex from both strains are given in Table 1. For comparison, the values of the kinetic constants of the Hex from *A. oryzae* are also provided. The K_m determined for PoHex was seven times higher with *p*NP-GalNAc than with *p*NP-GlcNAc. Inhibition by excess of the substrate was observed at concentrations exceeding 0.4 mM (Fig. 2C,D). The affinity of enzymes from both strains of *P. oxalicum* to the studied substrates was identical but significantly higher than for the *A. oryzae* Hex. Differences between PoHex from the two strains were found in reaction rates, maximal reaction rate and catalytic activity. The catalytic efficiency was more than three times higher for the PoHex from CCF 3438 than for that from the CCF 1959 strain, which correlates well with the differences in the propeptide content

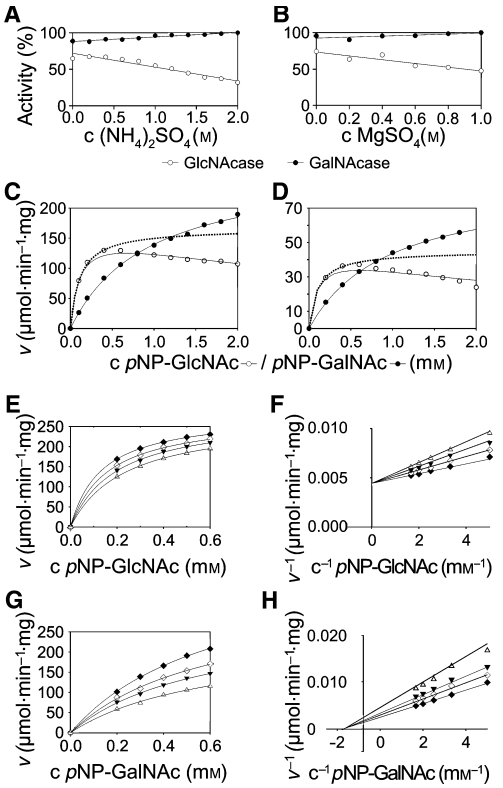

Fig. 2. Effect of salts, substrate and product concentrations on the reaction rate of PoHex from strain CCF 3438. (A), (B) The effect of $(NH_4)_2SO_4$ and $MgSO_4$, respectively, on Hex activity was measured for *p*NP-GlcNAc and *p*NP-GalNAc substrates. (C), (D) The effect of substrate concentrations (*p*NP-GlcNAc, *p*NP-GalNAc) on the initial reaction rate catalysed by PoHex (CCF 3438 and CCF 1959, respectively) was monitored. The experimental data were fitted to either the Michaelis–Menten equation or an equation describing substrate inhibition. The theoretical dependence according to Michaelis–Menten for *p*NP-GlcNAc is shown by the dotted line. (E)–(H) Inhibition of PoHex from *Penicillium oxalicum* strain CCF 3438 by the enzymatic product was monitored. Concentrations of the inhibitor used were (E), (F) 0 mM (filled diamonds), 5 mM (open diamonds), 10 mM (filled triangles) and 15 mM (open triangles) for GalNAc, and (G), (H) 0 mM (filled diamonds), 1 mM (open diamonds), 2 mM (filled triangles) and 5 mM (open triangles) for GlcNAc.

(three times lower in the CCF 1959 strain than the 3438 strain). The saturation kinetics of PoHex were also studied in the presence of 4-methylumbelliferyl 2-acetamido-2-deoxy-β-D-glucopyranoside (4-MU-GlcNAc) as substrate. The affinity of the enzymes for 4-MU-GlcNAc was higher and the reaction rate was lower than for *p*-nitrophenyl derivates, and substrate inhibition occurred for all hexosaminidases (Table 1).

The effect of the products (GlcNAc, GalNAc) on the reaction rate was studied in more detail (Table 2). Both compounds acted as inhibitors of the hydrolytic reaction catalysed by Hex; however, their behaviours

Table 1. Kinetic parameters of the β-hexosaminidases from *Penicillium oxalicum* strains CCF 1959 and 3438 with *p*NP-GlcNAc, *p*NP-GalNAc and 4MU-GlcNAc as substrates. n.i., not inhibiting.

	*p*NP-GlcNAc			*p*NP-GalNAc			4MU-GlcNAc		
β-hex	CCF 1959	CCF 3438	CCF 1066	CCF 1959	CCF 3438	CCF 1066	CCF 1959	CCF 3438	CCF 1066
K_m (mM)	0.14 ± 0.01	0.13 ± 0.01	1.10 ± 0.07	1.01 ± 0.02	1.04 ± 0.06	2.02 ± 0.22	0.07 ± 0.01	0.05 ± 0.01	0.14 ± 0.01
K_{ss} (mM)	0.41 ± 0.03	0.54 ± 0.04	n.i.	n.i.	n.i.	n.i.	0.06 ± 0.01	0.05 ± 0.01	0.75 ± 0.08
k_{cat} (s^{-1})	101 ± 2[a]	347 ± 13[a]	563 ± 17	227 ± 2	723 ± 21	419 ± 3	16 ± 1[a]	43 ± 3[a]	44 ± 2[a]
k_{cat}/K_m ($s^{-1}\cdot mM^{-1}$)	721[a]	2669[a]	511	224	695	207	229[a]	860[a]	314[a]

[a] The values were calculated from the highest reaction rate before inhibition due to excess substrate occurred.

Table 2. Inhibition constants (K_i) and the type of inhibition of the PoHex from *Penicillium oxalicum* strains CCF 1959 and 3438 for GlcNAc and GalNAc reaction products, and comparison with the Hex from *Aspergillus oryzae* (CCF 1066).

		GlcNAc		GalNAc	
Hex	Substrate	K_i (mM)	Type	K_i (mM)	Type
CCF 1959	*p*NP-GlcNAc	9	N	16	C
	*p*NP-GalNAc	10	N	22	C
CCF 3438	*p*NP-GlcNAc	4	N	13	C
	*p*NP-GalNAc	6	N	21	C
CCF 1066	*p*NP-GlcNAc	14	N	30	C
	*p*NP-GalNAc	16	N	22	C

were not identical. They differed not only in their inhibition constants, but also in the type of inhibition (Table 2, Fig. 2E–H). GlcNAc was a stronger inhibitor (non-competitive) than GalNAc (competitive). There was no significant difference in the inhibition of PoHex from the two fungal strains examined (CCF 1959 and 3438).

The ability of PoHex from strain CCF 3438 to hydrolyse substrates modified at their N-acetyl group (Fig. 3) was tested and compared with the enzyme from *A. oryzae* as a reference. PoHex cleaved these modified substrates significantly better than the *A. oryzae* enzyme, except for the trifluoroacetyl derivative which proved to be very resistant to hydrolysis by any enzyme preparation. The latter phenomenon is caused by the nature of the standard Hex hydrolytic mechanism via oxazoline intermediate [1]. Moreover, the crude *P. oxalicum* enzyme was more efficient at cleaving substrates with longer N-acyls such as N-glycolyl and N-propionyl (Fig. 3).

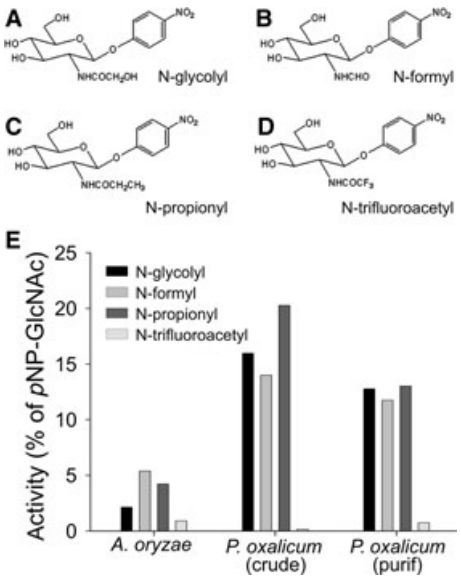

Fig. 3. Modified substrates cleaved by PoHex. *p*-Nitrophenyl 2-glycolylamido-2-deoxy-β-D-glucopyranoside (A), *p*-nitrophenyl 2-formamido-2-deoxy-β-D-glucopyranoside (B), *p*-nitrophenyl 2-propionamido2-deoxy-β-D-glucopyranoside (C) and *p*-nitrophenyl 2-trifluoroacetamido-2-deoxy-β-D-glucopyranoside (D) are shown at the top. (E) Cleavage of N-acyl modified substrates by Hex from various sources. The measured activities are compared with the activity obtained using the standard substrate, *p*NP-GlcNAc.

Molecular cloning and sequencing of PoHex

A detailed description of our molecular cloning strategy is described in supporting information (Data S1). The final DNA sequence containing the promoter-proximal region and the complete DNA sequence coding for the *PoHex* gene was deposited into GenBank (accession number EU189026). The sequences of the *PoHex* genes from both strains used were found to be entirely identical at the amino acid level; only three nucleotide differences causing no difference in the translated amino acid sequence were revealed. This indicates that little evolutionary drift occurred between the enzymes from the two available *P. oxalicum* strains. Promoter-proximal elements required for induction by GlcNAc [23] could

be identified approximately 300 bp upstream of the ATG triplet and were composed of two shorter ($_{-355}$CCAA$_{-352}$ and $_{-327}$AGGG$_{-324}$) elements and one extended regulatory sequence ($_{-312}$CCAGTGATC-ATTGCGCT-ACCCGTCTGGCCCT$_{-280}$). Although we could not identify the classical TATA box sequences, a promoter region containing two TATA box-like sequences (TAAATA and TAAATT) is located approximately 100 bp upstream from the ATG initiation codon. The start of transcription could also be identified as C$_{-70}$.

The sequence *PoHex* gene contains a single open reading frame of 1803 bp coding for 601 amino acids with no consensus intron sequences. The structure of the PoHex protein is closely related to that described previously for *A. oryzae* [10]. The entire protein is composed of the signal sequence, the propeptide and the catalytic domain including the C-terminal segment [10].

The catalytic subunit (Asp100–Pro601) is 501 amino acids long and contains several interesting structural determinants (Fig. 4). First, although there is no cysteine residue in the propeptide of the PoHex, there are six conserved cysteine residues in the catalytic subunits (marked by red dots in Fig. 4) that form three disulfide bridges supporting the structure of the catalytic subunit [24]. The arrangement of these disulfide bridges (similarly to the *A. oryzae* enzyme) is consecutive, i.e. Cys290 pairs with Cys351, Cys449 binds Cys483, and Cys583 forms a bridge with Cys590, as has been published in detail elsewhere [24]. Second, the substrate-hydrolysing and substrate-binding amino acids in the catalytic site of the enzyme that are conserved throughout the entire glycoside hydrolase family 20 [25] are also conserved in the *P. oxalicum* enzyme (Fig. 4, marked by black dots). Third, similarly to other fungal Hex the *P. oxalicum* enzyme is heavily N-glycosylated. In total, five classical and one non-canonical N-glycosylation sites have been detected in the sequence (Fig. 4, experimentally confirmed sites marked by rectangles). Experimental analysis of these individual N-glycosylation sites revealed that all the classical AsnXxxThr/Ser sequons are actually used and bear attached oligosaccharides, while the non-canonical site Asn339Asn-Cys341 was shown not to be used (P. Pompach, unpublished results).

N-Glycosylation influences the stability of the enzyme molecule

In order to study the role of N-glycosylation, we performed enzymatic deglycosylations using the commonly available enzymes peptide:*N*-glycosidase F (PNGase F) and endoglycosidase H (Endo H). When PoHex was deglycosylated using PNGase F under denaturing conditions, the molecular weight of its catalytic subunit shifted from approximately 65 to 56 kDa, i.e. its molecular weight was reduced by approximately 9 kDa (Fig. 5A, lanes 5 and 6). The theoretical molecular weight of the catalytic subunit calculated from its amino acid sequence should be 56 293 Da, indicating successful and complete deglycosylation by the enzyme. In order to further verify the extent of deglycosylation described above, we checked the glycosylation status of PoHex using mass spectrometry. The occupancy of individual sites of glycosylation clearly indicates that one site localized in the predicted propeptide and five classical sites in the catalytic subunits were all used for the attachment of glycans with average mass of $Man_8GlcNAc_2$ high-mannose oligosaccharides (Table S3). Thus, five sites of glycosylation containing glycans with averaged mass 1721 Da amount to an 8605 Da mass difference, corresponding well to experimental data.

Unfortunately, the use of PNGase F-treated enzymes for follow-up studies on enzymatic activity proved to be impossible, since deglycosylation only occurred efficiently under denaturing conditions. Using Endo H, however, it was possible to deglycosylate the enzyme under mild conditions at pH 5.5 without denaturation (Fig. 5A, lanes 2–4). Upon Endo H treatment, efficient removal of the majority of the carbohydrate (other than the core GlcNAc) occurred, causing a notable reduction in the size of the native enzyme (Fig. 5B, lane 2). α-Mannosidase, an exoglycosidase expected to digest most high-mannose type oligosaccharides, proved to be less efficient at reducing the molecular weight (Fig. 5B, lane 3). Sedimentation velocity measurements [26] in an analytical ultracentrifuge (Fig. 5C,D) indicated that there was a significant reduction in the value of the calculated sedimentation coefficient upon deglycosylation (7.8 S for the native and 7.2 S for the deglycosylated enzyme).

Similar to the *A. oryzae* enzyme, the Endo H-treated PoHex was less stable under acidic pH than the native enzyme (Fig. 5E). Interestingly, however, the stability was also significantly reduced at alkaline pH, and there were dramatic differences in stability between the deglycosylated and native enzyme at pH ~9. Nevertheless, the enzymatic activity of the Endo H-treated PoHex was very similar to that of the native enzyme, not only for the standard substrates (*p*NP-GlcNAc and *p*NP-GalNAc) but also for modified substrates (Fig. 5F; compare with Fig. 3E).

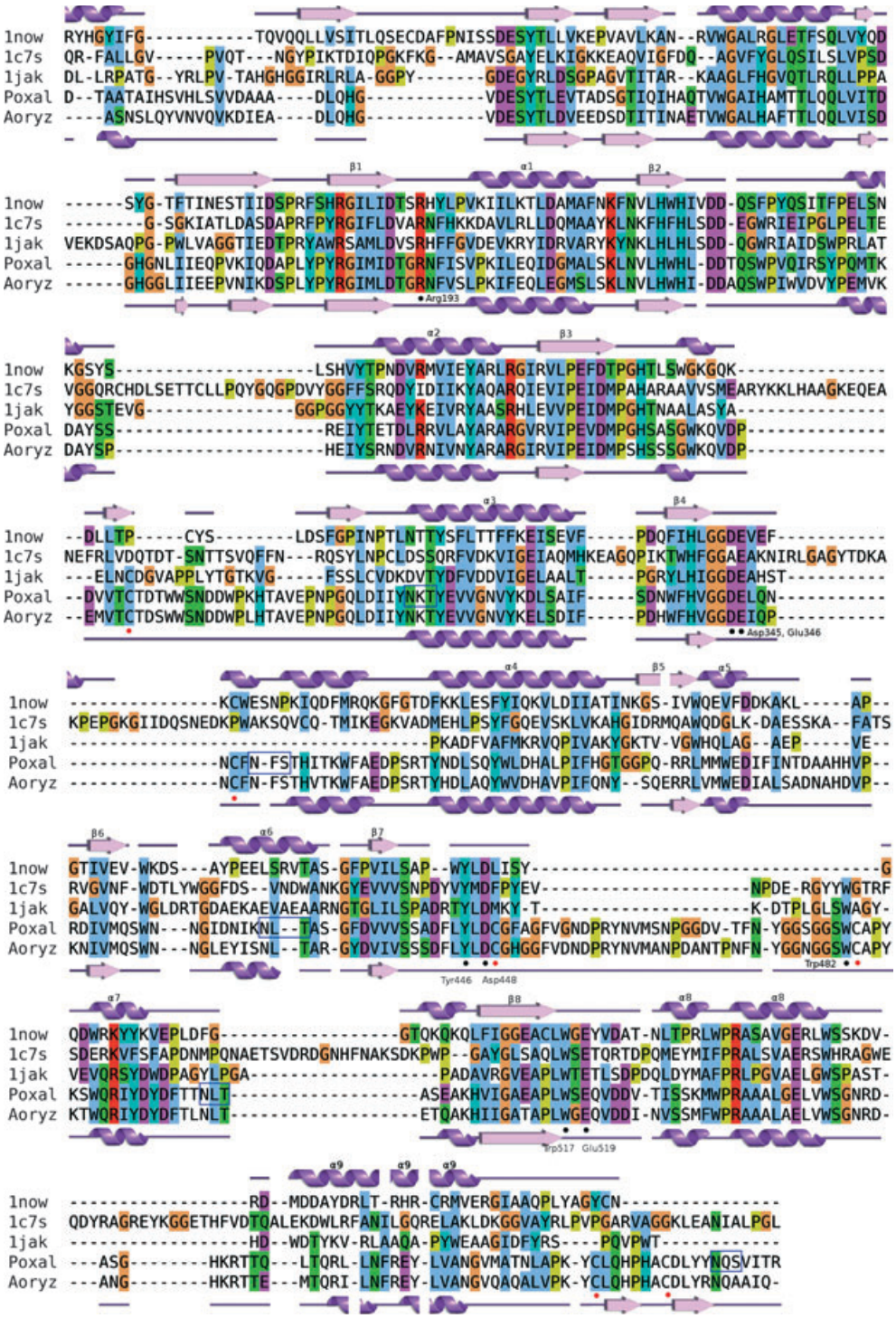

Fig. 4. Multiple structure-based sequence alignment of the catalytic unit of hexosaminidases from *Penicillium oxalicum*, *Aspergillus oryzae*, human (1NOW), bacteria (*Streptomyces plicatus* – 1JAK and *Serratia marcescens* – 1C7S). CLUSTALX colouring scheme is used. Secondary structure elements are shown for 1NOW above the aligned sequences (assigned by PROCHECK) and for PoHex below the aligned sequences (from the model). Numbering of the secondary structure elements of the catalytic domain is done according to Prag *et al.* [30]. The N-glycosylation sites confirmed by MS analyses of the enzyme are enclosed by blue rectangles. Active site amino acids are indicated by black dots. Cysteines that form disulfide bridges in the model of PoHex are identified by red dots.

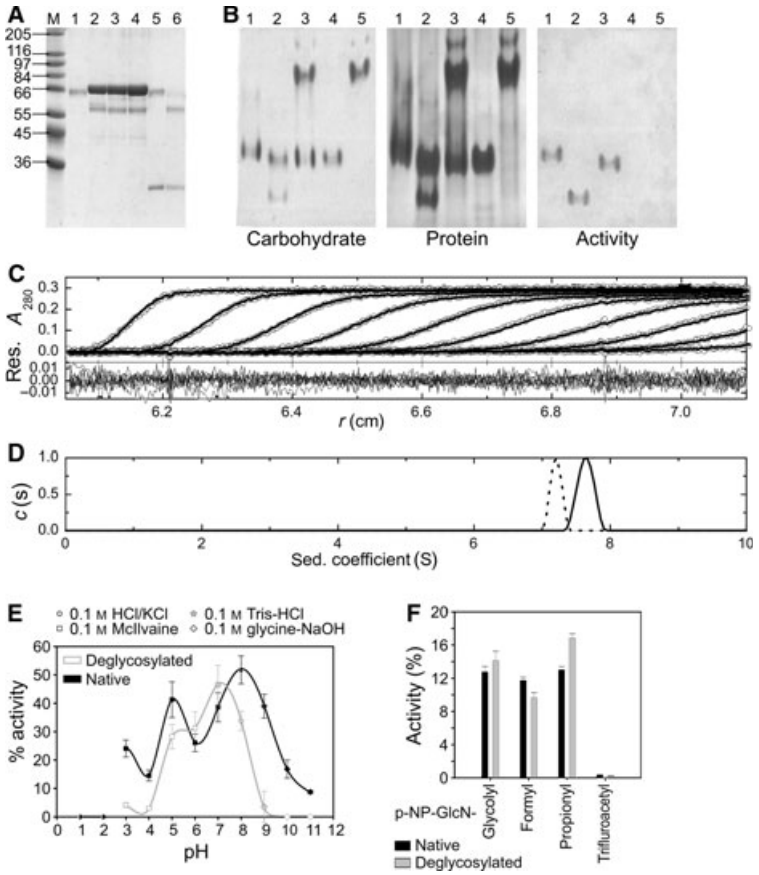

Fig. 5. Effect of deglycosylation of PoHex on stability and activity. (A) Deglycosylation of the hexosaminidase from *Penicillium oxalicum* CCF 1959 using Endo H and PNGase F. Lane 1, native Hex; lane 2, Hex deglycosylated by Endo H (buffer only); lane 3, PoHex deglycosylated by Endo H (buffer + SDS, no boiling); lane 4, PoHex deglycosylated by Endo H (buffer + SDS, boiled); lane 5, PoHex with PNGase F (no denaturation); lane 6, deglycosylated PoHex with PNGase F (after denaturation). Molecular weight marker is on the left. (B) Native electrophoretograms were stained for protein-linked carbohydrates (left panel), protein (middle panel) and for enzymatic activity (right panel). Lane 1, PoHex; lane 2, PoHex plus Endo H; lane 3, PoHex plus α-mannosidase; lane 4, Endo H; lane 5, α-mannosidase. (C) Sedimentation velocity analysis of native PoHex (CCF 3438): the fitted data (upper panel) with residual plot (bottom panel) showing goodness of fit are shown. (D) Calculated continuous size distribution *c*(*s*) of sedimenting species for native (full line) and deglycosylated (dashed line) PoHex. (E) Effect of deglycosylation on the pH stability of PoHex (CCF 3438). (F) Activity of deglycosylated PoHex (CCF 3438) and Hex from *Aspergillus oryzae* for modified substrates.

Molecular model of the PoHex

The similarity of the primary sequence to the structurally solved enzymes from bacteria *Serratia marcescens*, *Streptomyces plicatus* and humans allowed the computation of a three-dimensional homology model of PoHex. The multiple alignment shown in Fig. 4 used for the modelling of *P. oxalicum* was refined and adjusted, taking into account an older alignment [25] as well as secondary structure prediction using a hidden Markov model (HMM) and multiple structure-based sequence alignments (Fig. S4). Loops encompassing amino acids 300–312 and 454–472 were remodelled by MODLOOP [27]. Minor changes occurred in the secondary structure of the model after 2 ns of refinement: 0.4% of the turn structure was remodelled to β-sheets and the percentage of α-helical elements increased by 0.5%. The positions of the C-alpha atoms of Ala452-Asn462 and Gly469-Thr472 from the long loop moved by more than 0.3 nm.

Most (83.5%) of the amino acid residues are plotted in the favourable regions of the Ramachandran plot. The deviation of geometrical parameters from ideal values (G-factors) is higher than −0.5, characterizing an acceptable model. The overall average G-factor estimated by PROCHECK is −0.25 [25]. After refinement, the *Z*-score improved from −7.43 to −7.9, primarily due to the improvement of the model region from

amino acid 280 to amino acid 330 (Fig. S1). The BLAST algorithm identified two protein domains: the catalytic domain characteristic of the glycosyl hydrolase family 20 (GH20), which is represented by a TIM barrel fold, and a zincin-like domain (GH20b). The structure of the catalytic domain is very similar in all selected templates. The zincin-like fold of the obtained model consists of four antiparallel β-sheets and one α-helix (Fig. 6A, bottom right). The long loop between β-sheet 7 and α-helix 7 is characteristic of fungal Hex [25] (Fig. 6B). Bacteria have a significantly shorter loop in the corresponding place in their three-dimensional structure (Fig. 4), and the human enzyme has an even shorter turn.

Catalytic amino acids are highly conserved within the glycosyl hydrolase 20 family, at least within its clade A or 'subfamily 2' part into which the fungal hexosaminidases cluster [28,29] (Fig. 4, marked by black dots; only five of the seven indicated amino acids, namely Asp345, Glu346, Tyr446, Asp448 and Trp517, appear to be also conserved in clade B or 'subfamily 1' hexosaminidase related to *Caenorhabditis elegans* enzymes). Considering the clear differences in substrate specificity, there were surprisingly small variations in the primary structure of the Hex from *P. oxalicum* and *A. oryzae* (Fig. 4). There are two amino acid sequences close to the active site of the enzyme, however, where it appears that a distinct evolutionary rearrangement occurred. First, the sequence Gln387AsnTyrSerGln391 in the *A. oryzae* enzyme, which encompasses one of the N-glycosylation sites, is substantially different from the Gly387ThrGlyGly-Pro391 sequence found at the same location in the primary structure of PoHex (Fig. 4, loop between α-helix 4 and β-sheet 5). Second, the sequence Asp468Ala-AsnThrProAsn473, forming the lid of the substrate binding pocket in the *A. oryzae* enzyme fixed by the middle disulfide bridge, was replaced by a shorter Gly468GlyAspValThrPhe473 sequence (Fig. 4, loop between β-sheet 7 and α-helix 7). Thus, the smaller lid may allow better access and easier passage of larger substrates into the binding site of the enzyme.

Since N-linked glycans may significantly influence the surface characteristics of the enzyme as well as access of the substrate to the catalytic site, we decided to complete the molecular model by adding one of the most common glycans, $Man_5GlcNAc_2$, to all five possible sites. Demonstrating the spatial flexibility typical

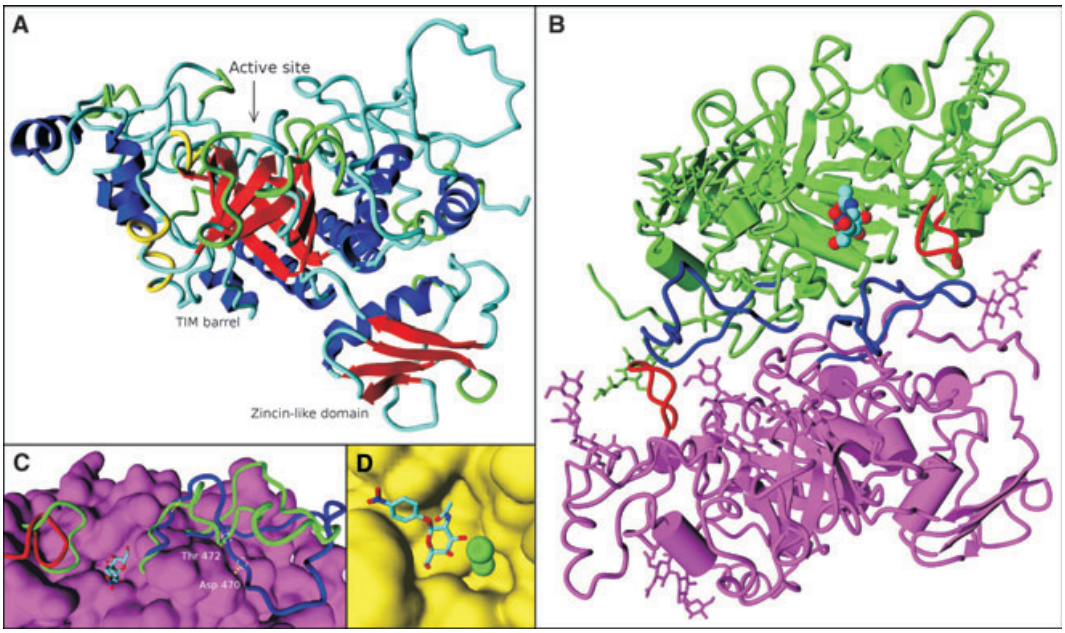

Fig. 6. Molecular model of the hexosaminidase from *Penicillium oxalicum*. (A) Molecular model. (B) Dimeric structure of the enzyme with *p*NP-GlcNAc docked at one monomer. The long loop specific for fungi is coloured in blue. Loop Pro301-Pro310 is coloured in red. (C) Overlay of the hexosaminidases from *P. oxalicum* and *Aspergillus oryzae* with *p*NP-GlcNAc docked at the active site. Loops Pro301-Pro310 (red) and 'lid' loop (blue) of the PoHex of one monomer (magenta). Corresponding loops of the *A. oryzae* Hex (green) are depicted with tubes, and the rest of the protein is represented by molecular surface (magenta). Hydrogen bonds created by loops and amino acids are coloured in yellow. The positions of the C-alpha atom of Asp470 and Thr472 residues are marked by white circles. (D) Surface of the active site with bound *p*NP-GalNAc after 5 ns of MD with Cl ions (green balls).

for this type of surface glycan, two of the N-linked sugar chains indeed appear to be in the proximity of the active site β-barrel and may thus influence the diffusion of the substrate to the binding site. Moreover, since this Hex is arranged as a dimeric enzyme under native conditions, we modelled the dimeric structure of PoHex as shown in Fig. 6B.

Visual analysis of the monomeric and dimeric models of both hexosaminidases confirmed that the most important difference between the two structures is the lid-forming loop (see above) [30]. Hydrogen bond pairs Asp470-Lys487 and Thr472-Tyr486 keep the 'lid' closer to its own monomer in PoHex (in contrast to *A. oryzae*), resulting in an active site that is more solvent-exposed (Fig. 6C). Sequence difference might provide an additional explanation for the lid bending back to its own monomer. Furthermore, differences are also evident in the spatially adjacent loop (Pro301-Pro310, Fig 6B,C), which is characterized by the presence of a hydrophilic positively charged lysine residue instead of the hydrophobic leucine in the *A. oryzae* enzyme (Fig. 4). In sum, these observations lead to the conclusion that the smaller and more flexible amino acids in the lid may allow better access and easier passage of the larger (modified) substrates into the binding site of the *P. oxalicum* enzyme.

Molecular dynamics (MD) simulations of the enzyme–substrate complex at various pH values revealed a stronger fluctuation of residues at pH 3 (Fig. S2). The lower stability of the protein at pH 3 could be explained by the protonation of Glu, Asp and His residues and by the loss of some stabilizing interactions (the total charge of the enzyme changes from −12 at pH 7 to +55 at pH 3). A strong distortion close to the active site of the enzyme was observed in the simulations at pH 3 when chloride ions (Fig. 6D) were used as counter-ions to reach simulated cell neutrality; these ions penetrated deep into the protein structure.

Docking of substrates and substrate analogues into the active site

Docking of *p*NP-GlcNAc and *p*NP-GalNAc substrates into the active site of the refined model of the PoHex, followed by MD simulations, revealed the atomic details of the substrate–enzyme interactions (Fig. 7A,B). The protein showed stable behaviour after only 1.5 ns of simulation (Fig. S3), so we used a 4-ns simulation for substrate–enzyme complex analysis to have at least 2 ns of equilibrated data for analysis. Whereas *p*NP-GlcNAc was bound with a total of eight hydrogen bonds, only five bonds could be identified for *p*NP-GalNAc binding. In particular, the C4 position (Fig. 7A,B, in yellow and magenta, respectively) seems to play a key role in the specificity of these interactions. For the *p*NP-GlcNAc substrate, the C4 hydroxyl hydrogen bonds to both Arg193 and Glu520, whereas for *p*NP-GalNAc only a single, non-persistent, hydrogen bond to Arg193 could be observed. This difference in binding is also reflected in the monitored interaction energies during the MD simulations. The standard MD simulations at pH 7 in water show an average value of the interaction energy for the equilibrated production phase of the simulation of 345 kJ·mol^{-1} for *p*NP-GlcNAc and 334 kJ·mol^{-1} for *p*NP-GalNAc.

Inhibition by excess substrate that was observed experimentally with *p*NP-GlcNAc (although not with *p*NP-GalNAc) (Fig. 2C,D) may be caused by the existence of additional binding sites for this compound. To test this hypothesis, a blind docking experiment was designed to screen the protein surface for additional potential binding sites. The docking experiment did indeed reveal the existence of one 'secondary' binding site (Fig. 7C) in close proximity to the active site of the enzyme. The interaction score of −21.5 kJ·mol^{-1} given by the scoring function of AUTODOCK [31,32] is comparable to the value measured for the substrate docked into the active site (−21.1 kJ·mol^{-1}). Hydrogen bonds were observed between the oxygen at the C4 position of *p*NP-GlcNAc and residues Arg491 and Asp443. The Asp425 residue was found to be within 0.3 nm of the docked inhibitor, a distance favourable for electrostatic interaction; in *A. oryzae* Hex, this residue is substituted by Glu424 and thus has a longer side chain (see Fig. 7). The PoHex residue Asp443 belongs to the same turn as the active site residue Tyr446, participating in the formation of a substrate–enzyme intermediate [4]. AUTODOCK was able to dock *p*NP-GalNAc into the enzyme active site (−18.0 kJ·mol^{-1}) but was unable to identify an additional binding site for it with favourable binding energy.

A similar procedure was used to investigate the mechanism of inhibition by the reaction products GlcNAc and GalNAc that is observed experimentally (Fig. 2E–H). Blind docking with GlcNAc shows a clear preference for the 'secondary' binding site (Fig. 7C), with an AUTODOCK score of −23.9 kJ·mol^{-1}, whereas the value for docking into the active site was only −14.2 kJ·mol^{-1}, which is significantly lower than with *p*NP-GlcNAc or *p*NP-GalNAc as substrate. On the other hand, the results for GalNAc suggest that docking is only favourable at the active site (−22.0 kJ·mol^{-1}). This active site interaction score is even slightly higher than with *p*NP-GalNAc, indicating a

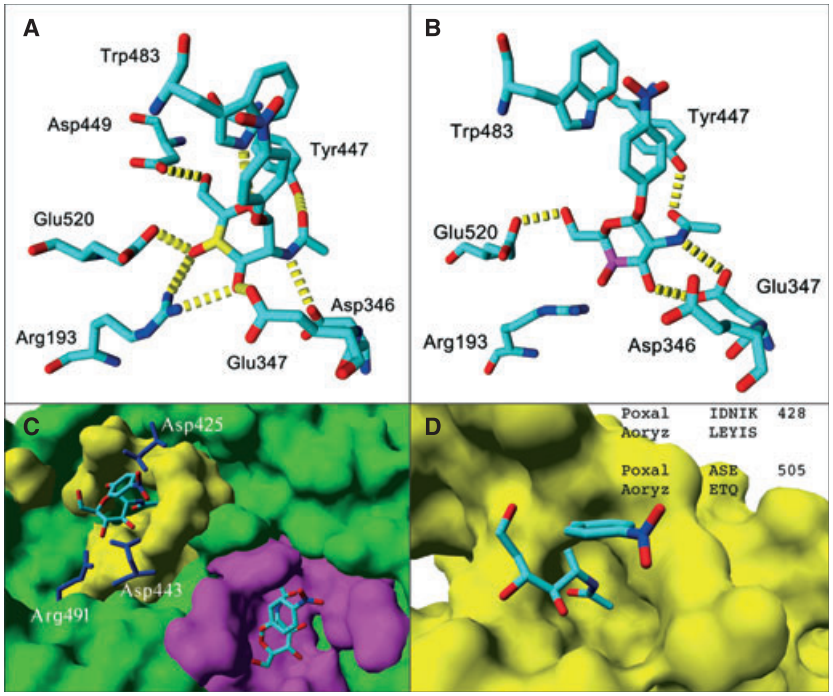

Fig. 7. Docking of *N*-acetylhexosamine substrates into the active site of the PoHex. (A) Active site with docked *p*NP-GlcNAc. The C4 atom is shown in yellow, and hydrogen bonds are shown by yellow dotted lines. (B) Active site with docked *p*NP-GalNAc. Hydrogen bonds are again yellow and the C4 atom is magenta. (C) Molecular surface representation of the protein with 'secondary' binding site (yellow) and active site of the enzyme (magenta) with bound *p*NP-GlcNAc. The position of amino acids responsible for the creation of hydrogen bonds (magenta lines) with the substrate at the secondary site are schematically depicted by blue sticks. (D) Secondary binding pocket of the PoHex with docked *p*NP-GlcNAc overlaid with the Hex from *Aspergillus oryzae* (yellow surface). In the upper right corner is a list of the portions of the sequence alignment that differ between the two enzymes in the vicinity of the secondary binding site.

clear competition between *p*NP-GalNAc and GalNAc for the active site of the enzyme.

The amino acids of the loop that come into close proximity to *p*NP-GlcNAc and GlcNAc at the 'secondary' binding site in the *P. oxalicum* enzyme differ from those in the *A. oryzae* Hex. This difference in models, determined by substitution of residues 503–505 and 424–428 from PoHex in the sequence of the hexosaminidase from *A. oryzae*, leads to a decrease in the size of this region in *A. oryzae* compared with the *P. oxalicum* enzyme (Fig. 7D). A shift in the position of the loops to accommodate the secondary substrate/product binding site may explain the differences in kinetics observed between the two enzymes (Table 1).

Hex substrates modified at their N-acyl residues fall into two different categories. The trifluoroacetyl derivative of *p*NP-GlcNAc is not hydrolysed by the enzyme from either tested species, whereas three other substrates with N-acyl modifications are much better hydrolysed by the *P. oxalicum* enzyme than the *A. oryzae* enzyme. Thus, we performed additional docking experiments in which we docked all four modified substrates in their standard form into the structures of Hex from both *P. oxalicum* and *A. oryzae*. Substrates bearing smaller N-acyl groups docked into the structure of the enzymes with significantly decreased docking energy. For example, the N-formyl substrate, in which the methyl group has been replaced by a much smaller hydrogen atom, docked with a docking energy of 339 kJ·mol^{-1} compared with the standard N-acetyl substrate, which yielded a docking energy of 345 kJ·mol^{-1} (Fig. 8A). The accommodation of this substrate into the substrate binding site is otherwise unaffected and proceeds in the same way as the standard N-acetyl substrate. However, the substrate is shifted in the active site of the enzyme, making the hydrolysed glycosidic bond more distant from the attacking catalytic residues (Table S4). Moreover, we observed a change in the distance from atom O28 (at the C3 atom of the pyranose ring) of this non-reducing sugar to the catalytic aspartic acid (Asp) responsible for proper orientation of the acetyl group during the formation of the oxazolinium ring. This distance shortened from 4.4 Å in the

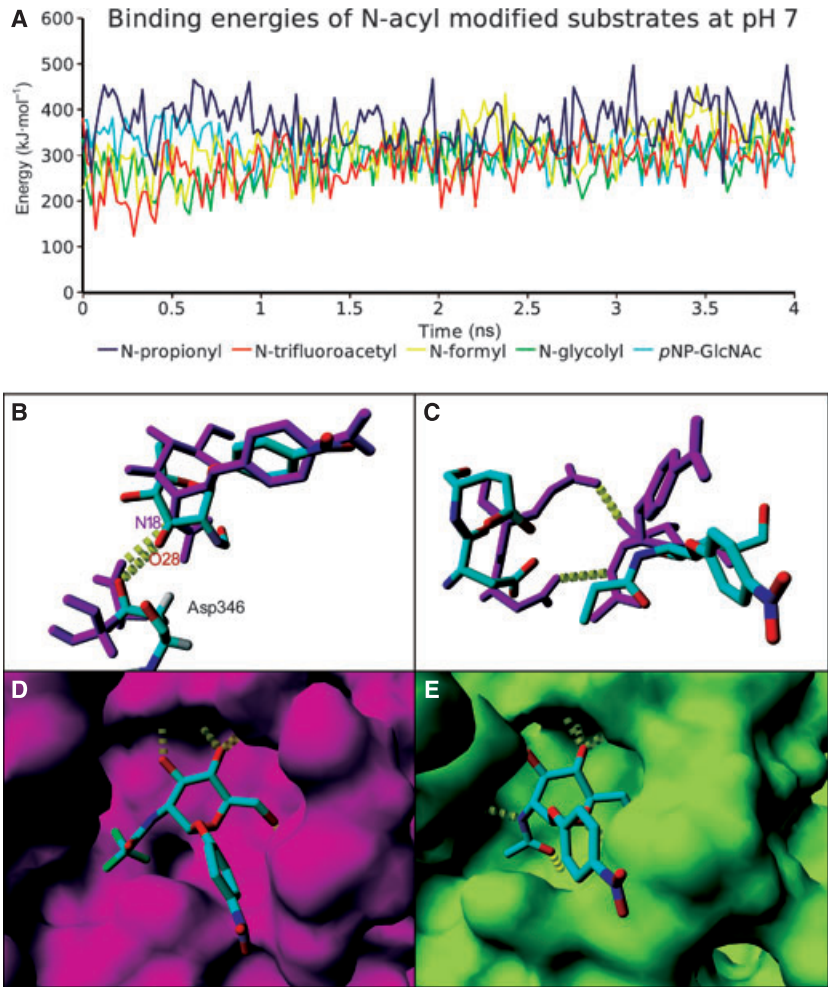

Fig. 8. Docking of modified *N*-acylhexosamine substrates into the active site of PoHex. (A) Binding energies of N-acyl modified substrates with PoHex during 4 ns of MD simulation. (B) Overlay of positions of N-formyl modified (light blue) and standard (magenta) substrates after 4 ns of MD. (C) Overlay of positions of N-propionyl modified (light blue) and standard (magenta) substrates after 4 ns of MD. (D) Molecular surface of the binding pocket of PoHex with bound N-trifluoroacetyl substrate after 4 ns of MD. (E) Molecular surface of the binding site of PoHex with bound *p*NP-GlcNAc after 4 ns of MD.

standard substrate (3.23 Å in *A. oryzae*) to 2.77 Å in the N-formyl derivative (2.56 Å in *A. oryzae*), enabling the formation of a hydrogen bond between the O28 and the O-Asp, which competes with the hydrogen bond formed by the N18(from the acetyl group)-H and the O-Asp (Fig. 8B).

A completely different mechanism applies for substrate analogues bearing longer acyls. These modified substrates generally displayed docking energies comparable with the acetyl-bearing substrate (e.g. the docking energy for N-propionyl substrate was 370.97 kJ·mol^{-1}). Such a high binding energy is the result of a hydrophobic effect that makes burying a long propionyl inside the protein energetically favourable. The accommodation of the larger substrates into the active site of the enzyme leads to a shift in their position, resulting in decreased access of the hydrolysing groups of the active-site amino acids to the glycosidic bond (Fig. 8C). The substrate in the active site is thus more exposed to water (Fig. 8D). This effect was also observed with the N-trifluoroacetyl substrate (Fig. 8D). Despite the fact that the binding energy of the N-trifluoroacetyl substrate was comparable with that observed in the N-glycoloyl substrate, the cleavage of the former substrate is complicated by electrostatic repulsion between the catalytic aspartic acid and the fluorines, which, taking into account the role of the Asp residue in the mechanism [33], can prevent the formation of the oxazolinium ring necessary for creating the intermediate structure.

Discussion

Fungal hexosaminidases have proved useful in biotechnology and chemoenzymatic syntheses of novel oligosaccharide sequences. Unique features of the PoHex among the secreted fungal hexosaminidases (although matched in mammalian HexD and hexosaminidases from *C. elegans* belonging to clade B [28,29]) such as a high ratio of GalNAc-ase/GlcNAc-ase activity and its

ability to cleave substrate analogues with various functional groups found for crude preparations appeared interesting, and deserved detailed molecular investigations using purified enzymes. To obtain standardized preparations, we used two publicly available collection strains, tested several media and cultivation conditions, and optimized a complete purification protocol for PoHex. The highest specific activity (35.6 $U \cdot mg^{-1}$ protein) and total yield (48 $U \cdot L^{-1}$ medium) was obtained using CCF 3438 strain cultivated for 7 days using medium M5. The PoHex isolated from the CCF 1959 strain had a significantly lower specific activity (10.8 $U \cdot mg^{-1}$ protein) despite its identical primary structure. These results may be explained by the low content of the propeptide in the latter preparation. We have previously shown for the Hex from *A. oryzae* that the propeptide represents an important enzyme regulator that associates with the catalytic subunit during its intracellular processing. The amount of propeptide present is related to individual stages of the life-cycle of the producing fungus [10]. Only catalytic subunits that are supplemented with the propeptide possessing proper post-translational modification (proteolysis at the dibasic site, specific glycosylation) can achieve full enzymatic activity and be secreted into the extracellular environment [9,10].

Previously published data describing the results of enzyme kinetics measurements using PoHex [18,19] are mostly in agreement with our results. However, working with the highly purified enzyme under strictly defined chemical conditions, we were able to find a number of hitherto undescribed characteristics that have escaped the attention of previous investigations. First, the inhibition of the enzyme by an excess of *p*NP-GlcNAc (but not *p*NP-GalNAc) is a completely novel finding, as is the fact that the ratio of GalNAcase to GlcNAcase activities is highly dependent on the concentration of the particular substrates. Second, the inhibition of the enzyme by its hydrolytic products (GlcNAc, GalNAc) has been studied in greater detail. Moreover, we now propose a mechanism for both of these phenomena based on the identification of a secondary binding site for *p*NP-GlcNAc (see below). Finally, the ratio of the two enzyme activities has been found to depend on the concentration of certain salts in the reaction buffer, especially ammonium sulfate, a common salt used for enzyme storage and stabilization. Based on these results, 0.4 mM *p*NP-GlcNAc substrate in buffers lacking ammonium sulfate or saturating concentrations of *p*NP-GalNAc in buffers with high ammonium sulfate concentrations appear optimal for practical use.

The purified PoHex displays a number of unique properties in terms of stability and substrate specificity. We have found that, unlike some Hex that are stable under mildly acidic conditions [1,10], PoHex has an optimum stability at pH 7–8. Additional stability maxima were detected for the native enzyme at pH 3 previously [19] and at pH 3 and pH 5 (this work). We have found that the pH stability profile is significantly influenced by enzyme glycosylation: Endo H-treated enzyme has a single stability maximum between pH 5 and 8 and remains completely inactive outside this range. The temperature optimum of the enzyme is somewhat lower than the optimum for the *A. oryzae* Hex (50 and 60 °C, respectively). The *P. oxalicum* enzyme cleaves substrates with N-acyl modifications much better (15% hydrolysis of N-formyl, N-glycolyl and N-propionyl derivatives of the standard N-acetyl substrate) than *A. oryzae* (2–5% hydrolysis).

The cloning and sequencing of PoHex and analysis of the corresponding gene confirmed a significant primary structure similarity to the Hex from *A. oryzae* published previously [10]. However, due to much higher quality of the sequence in the promoter upstream region of the *hex* gene from *P. oxalicum*, we have now been able to identify several regulatory sequences that might be responsible for the induction of enzyme synthesis by GlcNAc and related inducers [23]. Detailed comparisons of the primary structure of the Hex from *P. oxalicum* to *A. oryzae* revealed very small differences between the two enzymes, although the evolutionary pressure, and thus the rate of divergence, differed significantly within the individual segments of the sequence. Signal peptides are known to diverge very rapidly and, indeed, we could not find any similarity between the sequence of the signal peptide of the PoHex and other hexosaminidases. Large propeptides appear to be unique for the subfamily of fungal Hex and also appear to diverge relatively rapidly with 59% identity (70% similarity) to the evolutionarily closest propeptide from *A. terreus* NIH2624. The sequence similarity was particularly high for the catalytic subunit, with 76% identity (88% similarity) to the most closely related Hex from *P. chrysogenum* and 74% identity (87% similarity) to the Hex from *A. oryzae* CCF 1066 studied previously [10,25]. Moreover, important post-translational modifications such as disulfide bond formation [24] and N-glycosylation were also rather similar. The positions of all the cysteine residues, as well as the critical catalytic residues, were identical, and there were only two regions with significant differences in their primary structures.

Molecular modelling techniques that we employed contributed significantly to a better understanding of

the unique catalytic properties of the *P. oxalicum* enzyme that would not have been obvious otherwise considering the minimal differences in primary structure between PoHex and the Hex from *A. oryzae*. Using advanced techniques of contemporary molecular modelling, we were able to construct a three-dimensional model of the complete native Hex in the form of a fully glycosylated dimeric structure. The most dramatic differences between this three-dimensional model and the model of the Hex from *A. oryzae* published previously [25] are in the structure of the lids covering the active sites of the enzymes and the structurally adjacent protein loops. In the *P. oxalicum* enzyme, the corresponding amino acid residues are smaller and more flexible, and the lid loop tends to be in a more open conformation due, in part, to a direct binding to the same monomeric subunit (Fig. 6).

Moreover, techniques of MD simulations and ligand docking allowed us to provide a mechanistic understanding of the details of the kinetics of these complicated enzymes. In particular, the docking of both *p*NP-GlcNAc and *p*NP-GalNAc substrates allowed us to understand important features of Hex found experimentally, such as the values of K_m. In addition, the inhibition by an excess of *p*NP-GlcNAc could be understood by the identification of the 'secondary' binding site close to the active site of the enzyme. The fact that GlcNAc is bound preferentially to this secondary binding site while GalNAc binds only to the active site of the enzyme allows us to understand the molecular nature of the type of inhibition by the products (GlcNAc and GalNAc as non-competitive and competitive inhibitors, respectively). At the same time, we have become aware of certain limits of these techniques that could not fully explain the experimental results obtained with N-acyl modified substrates. In order to understand these experimental data, the catalytic process will have to be considered in its entirety, including not only the binding and hydrolysis of the substrate in the enzyme active site, but also the diffusion and access of the substrate to the binding site of the enzyme, as well as the release of the products. In conclusion, in this paper we have clarified the chemical nature of the unique catalytic properties of the PoHex, which will prove useful for the future use of this enzyme in biotechnology and chemoenzymatic synthesis.

Materials and methods

Microbial strains and growth conditions

The strains of *P. oxalicum* CCF 1959 and CCF 3438 were obtained from the publicly available Culture Collection of Fungi at the Department of Botany, Faculty of Science, Charles University Prague, Czech Republic (http://web.natur.cuni.cz/fccm/collecze.htm#ccf). The following culture media were used: M1 was made up of (per L) 0.3 g KH_2PO_4, 0.5 g $NH_4H_2PO_4$, 0.2 g $(NH_4)_2SO_4$, 0.1 g yeast extract, 0.5 g GlcNAc, 5 g NaCl, 0.05 g $MgSO_4$, pH 6 (adjusted with KOH); M2 was identical to M1 except that the yeast extract was replaced by 0.5 $g{\cdot}L^{-1}$ peptone; M3 was identical to M1 except that the yeast extract was replaced by 0.1 $g{\cdot}L^{-1}$ peptone; M4 was made up of (per L) 0.2 g $NaNO_3$, 0.05 g KCl, 0.001 g $FeSO_4$, 0.1 g KH_2PO_4, 1.0 g GlcNAc, 0.5 g $MgSO_4$, pH 5.5 (adjusted with HCl); M5 was identical to M4 except that the pH was adjusted to 4.5; M6 was identical to M5 except that the pH was set to 6.5. The fungi were cultivated at 28 °C with constant reciprocal shaking (200 r.p.m.) for the times indicated in individual experiments. After the end of the cultivation period, mycelia were removed by filtration, and the remaining medium was used for protein and enzyme activity determinations. The medium used for enzyme purification was collected after 12 and 7 days of cultivation for the CCF 1959 and CCF 3438 strains, respectively.

Purification of PoHex

The crude enzyme obtained by precipitation of the culture medium with ammonium sulfate (80% saturation) was dialysed against 0.6 M ammonium sulfate in 20 mM sodium phosphate buffer (pH 6.8). The enzyme solution was applied to a Phenyl-Sepharose HP column (2.0 × 20 cm; GE Healthcare, Fairfield, CT, USA), equilibrated using the same buffer. The enzyme was eluted with a linear salt gradient decreasing the content of ammonium sulfate (10 $mM{\cdot}min^{-1}$). The enzyme was further purified on a Mono Q column (1.5 × 15 cm; GE Healthcare) to near homogeneity using a linear gradient from 0 to 0.5 M NaCl in 20 mM Bistris buffer (pH 7.0). Final purification was achieved on a Superdex 200 HR column (1 × 30 cm; GE Healthcare), equilibrated and eluted in 10 mM Bistris buffer (pH 7.0) with 0.5 M $(NH_4)_2SO_4$. The enzyme was concentrated to 6 $mg{\cdot}mL^{-1}$ using a Centricon 30 (Millipore, Billerica, MA, USA) and stored at 4 °C.

Enzyme assay

The enzyme activity was monitored using *p*NP-GlcNAc or *p*NP-GalNAc substrates continuously at 348 nm [34] or via the end-point method at 405 nm. The enzyme activity was expressed in units (U, μmol of product formed per minute). The reaction mixture contained 50 mM citrate buffer (pH 3.0) and the corresponding substrate at the concentration corresponding to the saturated reaction rate (0.4 mM *p*NP-GlcNAc, 2 mM *p*NP-GalNAc). After incubation for an appropriate time, 0.2 M sodium carbonate (pH 11.0) was added, and the concentration of liberated *p*-nitrophenol

was determined. Steady-state initial velocity studies were performed at 20 °C in a final volume of 0.2 mL in 50 mM McIlvaine buffer (pH 3.0) containing 0.02–2 mM substrate. These data were fitted to the Michaelis–Menten equation or an equation characterizing substrate inhibition; Michaelis constant, maximal reaction rate, catalytic constant, catalytic efficiency and substrate inhibition constants were calculated. With 4-MU-GlcNAc substrate, the fluorescence of the product (4-MU) was measured (λ_{ex} 380 nm, λ_{em} 520 nm).

pH optimum, pH stability, temperature optimum

pH optimum of the β-hexosaminidase was determined using different buffers for various pH ranges: 0.1 M HCl/KCl buffer (pH 1–2), 0.1 M citrate/phosphate buffer (pH 3–7), 0.1 M Tris/HCl buffer (pH 7–9) and 0.1 M glycine/NaOH buffer (pH 9–11). pH stability of the enzyme was monitored in long-term assays upon incubation in the same series of buffers used for determining the pH optimum. The hexosaminidase was kept in these buffers at 4 °C, and aliquots were screened for enzyme activity at regular intervals. The activity was measured at the pH optimum (3.0). The temperature optimum was measured over the range 25–80 °C in 10 °C increments.

Molecular modelling

To identify homologues for the protein, a BLAST search (http://www.ncbi.nlm.nih.gov) in the databases of non-redundant protein sequences and some protein data banks were used. High-scoring templates were extracted from the Protein Data Bank results (http://www.pdb.org): 1NOW, 1C7S and 1JAK. The identified templates were used for making a structure-based multiple sequence alignment with the TCOFFEE server (http://tcoffee.vital-it.ch/cgi-bin/Tcoffee/tcoffee_cgi/index.cgi) [35] using Expresso mode. This method incorporates a consistency score to evaluate alignment. A structural alignment of the three templates was additionally made with the SHEBA plug-in and visualized with the program YASARA [36]. Structure alignment was further used for checking and correcting the TCOFFEE output according to secondary structure elements and conserved residues in known structures. To gain more information from the alignments, we also used a secondary structure prediction method (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_seccons.html) and two HMM models to gain reliable secondary structure prediction. Additionally, the older alignment used for modelling of the *A. oryzae* enzyme [25] that was based on a careful analysis of the whole protein family was used for comparison. Three-dimensional models were built with the package MODELLER 9.1 [37]. The stereochemical parameters of the models were assessed with the program PROCHECK [38], and energetic parameters were analysed by PROSA [39]. After several iterations of a (re)alignment model building validation process, the best model was selected. Loops were modelled using MODLOOP [27]. The improvement of the quality of models after loop modelling was analysed by PROCHECK and visual control.

The dimeric structure of the *P. oxalicum* enzyme was built by overlying the monomers with the dimeric structure of human β-hexosaminidase 1O7A. Glycosylation was performed online at http://www.glycosciences.de/ with SWEET [40]. Selected glycan antennae were cut down to leave two β-*N*-acetylglucosamine residues. The monomer was further refined with YASARA. The three-dimensional structure was placed in a box with periodic boundary conditions; the cell was filled with TIP3P water and neutralized by placing counter-ions (sodium ions). To remove steric overlaps and correct the covalent geometry, the energy of the complex was minimized with the YAMBER 2 force field and default parameters, followed by a short simulated annealing protocol (atom velocities scaled down by 0.9 every tenth step) until convergence was reached [25,34]. Docking (done by YASARA [36] and AUTODOCK [31,32]) and MD simulations by YASARA are described in supplementary material.

Acknowledgements

The authors acknowledge helpful comments and suggestions by the referees, and technical help by Daniel Kavan, Anna Hodková and Jiří Janovský. This work was supported in part by the Institutional Research Concepts AVOZ50200510 for the Institute of Microbiology and AVOZ60870520 for GCRC, by the Ministry of Education of the Czech Republic (LC06010, MSM 21620808, MSM6007665808 and 1M0505) and by the Grant Agency of the Czech Republic (303/09/0477, 203/09/P024 and 305/09/H008). Additionally, N.K. was supported by the University of South Bohemia, grant GAJU 170/2010/P.

References

1 Slámová K, Bojarová P, Petrásková L & Křen V (2010) β-*N*-Acetylhexosaminidase: what's in a name? *Biotechnol Adv* **28**, 682–693.
2 Proia RL (1988) Gene encoding the human β-hexosaminidase β-chain – extensive homology of intron placement in the α-chain and β-chain genes. *Proc Natl Acad Sci USA* **85**, 1883–1887.
3 Knapp S, Vocadlo D, Gao ZN, Kirk B, Lou JP & Withers SG (1996) NAG-thiazoline, an *N*-acetyl-β-hexosaminidase inhibitor that implicates acetamido participation. *J Am Chem Soc* **118**, 6804–6805.
4 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.
5 Gooday GW, Zhu WY & Odonnell RW (1992) What are the roles of chitinases in the growing fungus? *FEMS Microbiol Lett* **100**, 387–391.
6 Reyes F, Calatayud J, Vazquez C & Martinez MJ (1989) β-*N*-Acetylglucosaminidase from *Aspergillus nidulans* which degrades chitin oligomers during autolysis. *FEMS Microbiol Lett* **65**, 83–87.
7 Křen V & Martínková L (2001) Glycosides in medicine: the role of glycosidic residue in biological activity. *Curr Med Chem* **8**, 1303–1328.
8 Bojarová P & Křen V (2009) Glycosidases: a key to tailored carbohydrates. *Trends Biotechnol* **27**, 199–209.
9 Plíhal O, Sklenář J, Kmoníčková J, Man P, Pompach P, Havlíček V, Křen V & Bezouška K (2004) *N*-Glycosylated catalytic unit meets *O*-glycosylated propeptide: complex protein architecture in a fungal hexosaminidase. *Biochem Soc Trans* **32**, 764–765.
10 Plíhal O, Sklenář J, Hofbauerová K, Novák P, Man P, Pompach P, Kavan D, Ryšlavá H, Weignerová L, Charvátova-Pišvejcová A *et al.* (2007) Large propeptides of fungal β-*N*-acetylhexosaminidases are novel enzyme regulators that must be intracellularly processed to control activity, dimerization, and secretion into the extracellular environment. *Biochemistry* **46**, 2719–2734.
11 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.
12 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.
13 Fialová P, Namdjou DJ, Ettrich R, Přikrylová W, Rauvolfová J, Křenek K, Kuzma M, Elling L, Bezouška K & Křen V (2005) Combined application of galactose oxidase and β-*N*-acetylhexosaminidase in the synthesis of complex immunoactive *N*-acetyl-D-galactos-aminides. *Adv Synth Catal* **347**, 997–1006.
14 Hušáková L, Riva S, Casali M, Nicotra S, Kuzma M, Huňková Z & Křen V (2001) Enzymatic glycosylation using 6-O-acylated sugar donors and acceptors: β-*N*-acetylhexosaminidase-catalysed synthesis of 6-*O,N,N*'-triacetylchitobiose and 6′-*O,N,N*'-triacetylchitobiose. *Carbohydr Res* **331**, 143–148.
15 Slámová K, Gažák R, Bojarová P, Kulik N, Ettrich R, Pelantová H, Sedmera P & Křen V (2010) 4-Deoxy-substrates for β-*N*-acetylhexosaminidases: how to make use of their loose specificity. *Glycobiology* **20**, 1002–1009.
16 Di Giambattista R, Federici F, Petruccioli M & Fenice M (2001) The chitinolytic activity of *Penicillium janthinellum* P9: purification, partial characterization and potential applications. *J Appl Microbiol* **91**, 498–505.
17 Diez B, Rodriguez-Saiz M, de la Fuente JL, Moreno MA & Barredo JL (2005) The nagA gene of *Penicillium chrysogenum* encoding β-*N*-acetylglucosaminidase. *FEMS Microbiol Lett* **242**, 257–264.
18 Rodriguez J, Copapatino JL, Reyes F & Perezleblic MI (1994) A β-*N*-acetylhexosaminidase from *Penicillium oxalicum* implicated in its cell-wall degradation. *Lett Appl Microbiol* **19**, 217–220.
19 Yamamoto K, Lee KM, Kumagai H & Tochikura T (1985) Purification and characterization of β-*N*-acetylhexosaminidase from *Penicillium oxalicum*. *Agric Biol Chem* **49**, 611–619.
20 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.
21 Laemmli UK (1970) Cleavage of structural proteins during assembly of head of bacteriophage-T4. *Nature* **227**, 680–685.
22 Lee DH & Lee CB (2000) Chilling stress-induced changes of antioxidant enzymes in the leaves of cucumber: in gel enzyme activity assays. *Plant Sci* **159**, 75–85.
23 Peterbauer CK, Brunner K, Mach RL & Kubicek CP (2002) Identification of the *N*-acetyl-D-glucosamine-inducible element in the promoter of the *Trichoderma atroviride* nag1 gene encoding *N*-acetyl-glucosaminidase. *Mol Gen Genomics* **267**, 162–170.
24 Pompach P, Man P, Kavan D, Hofbauerová K, Kumar V, Bezouška K, Havlíček V & Novak P (2009) Modified electrophoretic and digestion conditions allow a simplified mass spectrometric evaluation of disulfide bonds. *J Mass Spectrom* **44**, 1571–1578.
25 Ettrich R, Kopecký V, Hofbauerová K, Baumruk V, Novák P, Pompach P, Man P, Plíhal O, Kutý M, Kulik N *et al.* (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.
26 Schuck P (2003) On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. *Anal Biochem* **320**, 104–124.
27 Fisher A, Do RK & Sali A (2000) Modeling of loops in protein structures. *Protein Sci* **9**, 1753–1773.
28 Intra J, Pavesi G & Horner DS (2008) Phylogenetic analyses suggest multiple changes of substrate specificity within the glycosyl hydrolase 20 family. *BMC Evol Biol* **8**, 214.
29 Gutternigg M, Rendič D, Voglauer R, Iskratsch TI & Wilson IBN (2009) Mammalian cells contain a second nucleocytoplasmic hexosaminidase. *Biochem J* **419**, 83–90.
30 Prag G, Papanikolau Y, Tavlas G, Vorgias CE, Petratos K & Oppenheim AB (2000) Structures of chitobiase

mutants complexed with the substrate di-*N*-acetyl-D-glucosamine: the catalytic role of the conserved acidic pair, aspartate 539 and glutamate 540. *J Mol Biol* **300**, 611–617.
31 Goodsell DS & Olson AJ (1990) Automated docking of substrates to proteins by simulated annealing. *Proteins* **8**, 195–202.
32 Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK & Olson AJ (1998) Automated docking using a Lamarckian genetic algorithm and empirical binding free energy function. *J Comput Chem* **19**, 1639–1662.
33 Mark BL, Mahuran DJ, Cherney MM, Zhao DL, Knapp S & James MNG (2003) Crystal structure of human β-hexosaminidase B: understanding the molecular basis of Sandhoff and Tay–Sachs disease. *J Mol Biol* **327**, 1093–1109.
34 Brumer H, Sims PFG & Sinnott ML (1999) Lignocellulose degradation by *Phanerochaete chrysosporium*: purification and characterization of the main α-galactosidase. *Biochem J* **339**, 43–53.
35 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.
36 Krieger E, Darden T, Nabuurs SB, Finkelstein A & Vriend G (2004) Making optimal use of empirical energy functions: force-field parameterization in crystal space. *Proteins* **57**, 678–683.
37 Sali A & Blundell TL (1993) Comparative protein modeling by satisfaction of spatial restraints. *J Mol Biol* **234**, 779–815.
38 Laskowski RA, Macarthur MW, Moss DS & Thornton JM (1993) Procheck – a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* **26**, 283–291.
39 Sippl MJ (1993) Recognition of errors in three-dimensional structures of proteins. *Proteins* **17**, 355–362.
40 Bohne-Lang A & von der Lieth CW (2005) GlyProt: *in silico* glycosylation of proteins. *Nucleic Acids Res* **33**, W214–W219.

Supporting information

The following supplementary material is available:
Doc. S1. Supporting methods.
Fig. S1. Results of the validation of the refined model by PROCHECK.
Fig. S2. Relative mobility of surface residues shown as a function of pH.
Fig. S3. RMSD of C-alpha atoms of the model of PoHex during MD with different substrates.
Fig. S4. Secondary structure comparison with HMM predictions.
Table S1. Purification of the hexosaminidase from *Penicillium oxalicum*.
Table S2. Occurrence of contaminating glycosidases in the course of purification.
Table S3. Occupancy of individual sites of N-glycosylation in PoHex by high mannose glycans identified by mass spectrometry.
Table S4. Trp and position of the catalytic residues for docking of the modified substrates.

This supplementary material can be found in the online version of this article.

Please note: As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be reorganized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.