Expression, purification, crystallization and preliminary crystallographic analysis of a GH20 β -N-acetylglucosaminidase from the marine bacterium *Vibrio harveyi*

Piyanat Meekrathok,^a Marco Bürger,^b Arthur T. Porfetye,^b Ingrid R. Vetter^{b*} and Wipa Suginta^{a*}

Received 11 January 2015

Accepted 27 February 2015

Edited by T. Bergfors, Uppsala University, Sweden

Keywords: β -N-acetylglucosaminidase; *Vibrio harveyi*; GH20 glycoside hydrolase family.

Supporting information: this article has supporting information at journals.iucr.org/f

^aBiochemistry–Electrochemistry Research Unit, School of Biochemistry, Institute of Science, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand, and ^bMax Planck Institute of Molecular Physiology, 44227 Dortmund, Germany. *Correspondence e-mail: ingrid.vetter@mpi-dortmund.mpg.de, wipa@sut.ac.th

Vibrio harveyi β -N-acetylglucosaminidase (*VhGlcNAcase*) is a new member of the GH20 glycoside hydrolase family responsible for the complete degradation of chitin fragments, with *N*-acetylglucosamine (GlcNAc) monomers as the final products. In this study, the crystallization and preliminary crystallographic data of wild-type *VhGlcNAcase* and its catalytically inactive mutant D437A in the absence and the presence of substrate are reported. Crystals of wild-type *VhGlcNAcase* were grown in 0.1 *M* sodium acetate pH 4.6, 1.4 *M* sodium malonate, while crystals of the D437A mutant were obtained in 0.1 *M* bis-tris pH 7.5, 0.1 *M* sodium acetate, 20% PEG 3350. X-ray data from the wild-type and the mutant crystals were collected at a synchrotron-radiation light source and were complete to a resolution of 2.5 Å. All crystals were composed of the same type of dimer, with the substrate *N,N*-diacetylglucosamine (GlcNAc₂ or diNAG) used for soaking was cleaved by the active enzyme, leaving only a single GlcNAc molecule bound to the protein.

1. Introduction

Chitin is a β -1,4-linked *N*-acetylglucosamine (GlcNAc) polymer and is one of the most abundant biopolymers in nature. Chitin serves as a major component of fungal cell walls and also of the exoskeletons of crustaceans and insects and the microfilaria sheaths of parasitic nematodes, molluscs, protozoa and other marine organisms (Armand *et al.*, 1994; Yuli *et al.*, 2004). In marine ecosystems, chitin biomaterials are initially hydrolyzed to short-chain oligosaccharides by chitinases, hydrolytic enzymes that are secreted mainly by marine bacteria such as *Vibrio* sp. (Ohishi *et al.*, 1996; Suginta *et al.*, 2000; Kadokura *et al.*, 2007). *V. harveyi* is a Gram-negative marine bacterium that has been reported to express high levels of chitinases and GlcNAcases in order to efficiently utilize chitin biomaterials as its sole source of energy (Suginta *et al.*, 2000). The bacterium is both a primary and an opportunistic pathogen of marine animals, triggering a lethal disease called luminous vibriosis (Austin & Zhang, 2006), which affects marine fish and prawn farming operations worldwide (Liuxy *et al.*, 1996; Actis *et al.*, 2011). In the chitin-degradation pathway of *V. harveyi*, chitinases initially degrade chitin to small chitooligosaccharide fragments, which are further transported across the bacterial cell wall through the chito-oligosaccharide-uptake channel, which is known as chitoporin (Keyhani *et al.*, 2000; Li & Roseman, 2004; Suginta, Chumjan,

Mahendran, Janning *et al.*, 2013; Suginta, Chumjan, Mahendran, Schulte *et al.*, 2013). In the periplasm, GlcNAcases (EC 3.2.1.52) degrade chitin oligosaccharides to GlcNAc monomers, which are transported across the inner membrane by a specific ABC transporter into the cytoplasm and are then further metabolized to metabolic intermediates that can readily be converted to carbon and nitrogen sources for cells. In bacteria, GlcNAcases act as biocatalysts during the natural recycling of chitin biomaterials. However, the human homologues of bacterial GlcNAcases, known as HexA and HexB, are clinically important since they are critically involved in the degradation of glycosphingolipids, which are deposited in the form of gangliosides GM1, GM2 and GM3 on the plasma membranes of nerve cells. Mutations of these enzymes lead to the accumulation of GM2 gangliosides, which results in fatal diseases such as Tay–Sachs disease and Sandhoff disease (Myerowitz, 1997).

In the Carbohydrate-Active Enzymes (CAZy) database (<http://www.cazy.org/>), GlcNAcases are classified as glycoside hydrolases of family 3 (GH3) or family 20 (GH20), which differ from each other in their sequences and their mechanisms of enzyme action (Henrissat & Bairoch, 1993; Henrissat & Davies, 1997; Lombard *et al.*, 2014). GH3 GlcNAcases catalyze the release of terminal 2-acetamido-2-deoxy- β -D-glycopyranoside from various glycoconjugates (Drouillard *et al.*, 1997) using the covalent α -glycosidic enzyme intermediate mechanism (Vocadlo & Withers, 2005). In contrast, GH20 GlcNAcases typically hydrolyze chitin fragments employing a substrate-assisted mechanism, in which the carbonyl O atom of the C₂ N-acetyl moiety takes part in cleavage of the scissile bond, yielding an oxazolinium ion as the reaction intermediate (Mark *et al.*, 2001, 2003; Lemieux *et al.*, 2006; Langley *et al.*, 2008).

In terms of structural studies, the crystal structure of *Serratia marcescens* chitobiase (SmChiB) was the first GH20 GlcNAcase structure to be solved (in a complex with diNAG; Tews *et al.*, 1996; Prag *et al.*, 2000). To date, a few other GH20 structures have been reported, including those of *Streptomyces plicatus* β -Hex (Williams *et al.*, 2002), human HexB (Mark *et al.*, 2003; Maier *et al.*, 2003), human HexA (Lemieux *et al.*, 2006), *Streptococcus gordonii* GcnA (Langley *et al.*, 2008), *Paenibacillus* sp. TS12 β -Hex (Sumida *et al.*, 2009) and *S. pneumoniae* R6 β -Hex (Jiang *et al.*, 2011). All GH20 hexosaminidases have a structural feature in common: a central catalytic (β/α)₈ TIM-barrel domain with a substrate-binding pocket located within the (β/α)₈ TIM barrel.

In a previous study, we isolated two GH20 GlcNAcase genes encoding two GlcNAcase isoforms, *VhNag1* and *VhNag2*, from the genome of *V. harveyi* strain 650 (Suginta *et al.*, 2010). Both genes were cloned in the pQE60 expression vector and were expressed at high levels in *Escherichia coli* M15 (pREP) host cells. *VhNag1* was inactive and had a molecular mass of about 93 kDa. *VhNag2* had a molecular mass of 75 kDa and was catalytically active, with a pH optimum of 7.0. *VhNag2* is an exolytic enzyme that sequentially hydrolyzes chitin oligosaccharides, releasing GlcNAc as the end product. Kinetic studies of the hydrolysis of the

natural glycosides showed that chitotetraose was the best substrate of *VhNag2* (Suginta *et al.*, 2010). In the present study, we report the expression, purification and crystallization trials of *VhNag2* (hereafter referred to as *VhGlcNAcase*) in the absence and presence of the substrate *N,N'*-diacetylglucosamine (GlcNAc₂ or diNAG). Based on a literature review and multiple sequence alignment, the Asp437-Glu438 residues are equivalent to the catalytic pair Asp539-Glu540 in *SmChiB* (Tews *et al.*, 1996) and Asp313-Glu314 in *S. plicatus* β -Hex (Williams *et al.*, 2002). Therefore, it is presumed that both residues are important in substrate-assisted catalysis and that Asp437 may play the same role as Asp539 in *SmChiB* and Asp313 in β -Hex. Accordingly, we mutated this residue and the corresponding mutant D437A was expressed, purified and crystallized in an attempt to obtain an inactive protein that might still interact with the substrate diNAG, in order to obtain detailed information on the protein–sugar interactions.

2. Materials and methods

2.1. Cloning, expression and purification

The full-length GlcNAcase cDNA (accession No. HM175716) was isolated from the genome of *V. harveyi* by the PCR technique, and the gene encoding *VhGlcNAcase* (amino acids 5–642, lacking the signal peptide) was then cloned into the pQE60 expression vector (Qiagen, Valencia, California, USA), generating a four-amino-acid (MGGS) cloning artefact at the N-terminus as described previously (Suginta *et al.*, 2010). The recombinant protein was shown to be expressed at a high level in *E. coli* M15 (pREP4) cells as a C-terminally His₆-tagged polypeptide (Suginta *et al.*, 2010). The D437A mutant of the GlcNAcase was generated in this study using the QuikChange Site-Directed Mutagenesis Kit (Stratagene, California, USA) according to the manufacturer's protocols, using the full-length GlcNAcase cDNA as template. The mutagenic forward and reverse primers were 5'-GTTCA-CATTGGCGCGCGGAAGTGCCTGTGCCTAACGGC-3' and 5'-GCCGTTAGGCACTTCCGCCGCGCCAATGTGA-AC-3', respectively. The underlined sequences represent the mutated codon. After gene amplification, the mutated DNA was treated with DpnI and then transformed into *E. coli* XL1-Blue competent cells. The mutated plasmids were extracted using QuickClean II Plasmid Miniprep Kits (GenScript, Piscataway, New Jersey, USA) and the mutated sequence was verified by automated sequencing (First Base Laboratories Sdn Bhd, Selangor Darul Ehsan, Malaysia).

In this study, the expression and purification of recombinant wild-type *VhGlcNAcase* and the D437A mutant were optimized in order to improve the yield and the purity of the proteins for crystallization purposes. Cells harbouring the recombinant pQE60/GlcNAcase plasmid were grown at 310 K in Terrific Broth (TB) medium containing 100 μ g ml⁻¹ ampicillin and 25 μ g ml⁻¹ kanamycin until the OD₆₀₀ of the cell culture reached 0.6. GlcNAcase expression was induced by the addition of isopropyl β -D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.4 mM and cell growth continued at

293 K for 18 h. The cell pellet was then collected by centrifugation and resuspended in lysis buffer [20 mM Tris–HCl pH 8.0 containing 150 mM NaCl, 1 mM phenylmethylsulfonyl fluoride (PMSF), 5% (v/v) glycerol and 1 mg ml^{−1} lysozyme]. The resuspended cells were then lysed on ice using a Sonopuls ultrasonic homogenizer with a 6 mm diameter probe (50% duty cycle; amplitude setting 30%; total time 30 s; 6–8 pulses).

For purification, freshly prepared crude lysate obtained after homogenization was centrifuged at 12 000g for 1 h at 277 K and then applied onto a TALON metal-affinity resin (1.0 × 10 cm; Clontech, USA). Affinity chromatography was carried out at 277 K under gravity-dependent flow following the manufacturer's instructions (<http://www.clontech.com>). The column was washed thoroughly with equilibration buffer (20 mM Tris–HCl buffer pH 8.0 containing 150 mM NaCl) followed by 10 mM imidazole in equilibration buffer; GlcNAcase was eluted with 250 mM imidazole in equilibration buffer and three eluted fractions (10 ml each) were collected. After the purity of the proteins had been analyzed by SDS–PAGE, the GlcNAcase-containing fractions were pooled, concentrated using a Vivaspın 20 ultrafiltration membrane concentrator (10 kDa molecular-weight cutoff; Vivascience AG, Hannover, Germany) and then further purified on a HiPrep 16/60 Sephacryl S-200 prepacked column connected to an ÄKTAprime purification system (Amersham Bioscience, Piscataway, New Jersey, USA). For the wild-type protein, the running buffer was low-salt equilibration buffer (20 mM Tris–HCl buffer pH 8.0 containing 30 mM NaCl and 1 mM TCEP). The wild-type protein was prepared in low-salt equilibration buffer to avoid interference by salt in the crystallization trials. However, some of the wild-type protein appeared as an aggregation peak in the gel-filtration profile. Purification of the D437A mutant at a low salt concentration showed complete aggregation of the protein. Therefore, the solubility of the mutated protein was improved by increasing the salt concentration to 150 mM in the same equilibration buffer as used previously (Suginta *et al.*, 2010). The chromatography column was operated at a flow rate of 0.2 ml min^{−1}, and 1.4 ml fractions were collected and assayed for GlcNAcase activity using a colorimetric assay with pNP-GlcNAc as the substrate. The D437A mutant had only 1–2% of the activity of the wild type (data not shown). GlcNAcase-containing fractions were pooled and concentrated using the same type of Vivaspın membrane concentrator. Protein concentrations were initially determined by measuring the absorption at 280 nm (Edelhoc, 1967; Gill & von Hippel, 1989; Pace *et al.*, 1995) and then by the method of Bradford (1976) using a standard calibration curve constructed with BSA. The purity and the molecular mass of the enzyme monomer were determined by SDS–PAGE (Laemmli, 1970). Protein aggregation after purification was evaluated using native PAGE according to a protocol described elsewhere (Arndt *et al.*, 2012).

2.2. Protein crystallization

Initial crystallization screening of wild-type *Vh*GlcNAcase was performed at 293 K using commercially available screens

from Qiagen, Hilden, Germany: the sparse-matrix screens the JCSG Core Suites I, II, III and IV, the Classics and Classics II Suites and the PACT Suite and the grid screens the PEGs Suite and the Anions Suite. The screens were set up in 96-well CrystalQuick plates (Greiner Bio-One, Frickenhausen, Germany) using a Mosquito Crystallization robot (TTP Labtech, Melbourn, England) with the sitting-drop method. The purified enzyme was centrifuged at 13 000g for 5 min and 0.1 µl of the freshly prepared wild-type protein at 10 mg ml^{−1} in equilibration buffer consisting of 20 mM Tris–HCl pH 8.0, 30 mM NaCl was then pipetted into 0.1 µl of each precipitant in the crystallization screens. Small plate-shaped crystals of wild-type GlcNAcase were observed within 1 d in condition F10 from the PACT Suite [20% (w/v) PEG 3350, 0.1 M bis-tris propane pH 6.5, 0.2 M sodium/potassium phosphate] and small crystals also appeared within 3 d in condition C6 of the Anions Suite (0.1 M sodium acetate pH 4.6, 1.2 M sodium malonate). Crystals obtained from both conditions were further optimized by the hanging-drop vapour-diffusion technique using different concentrations of two precipitants: (i) 17–22% (w/v) PEG 3350, 0.1–0.2 M bis-tris propane pH 6.5, 0.1–0.2 M sodium/potassium phosphate and (ii) 0.1–0.2 M sodium acetate pH 4.6, 0.7–1.8 M sodium malonate. In each drop, 1.5 µl GlcNAcase solution was mixed with 1.5 µl of each precipitant and then equilibrated over 1.0 ml of the respective precipitant in a 24-well Linbro tissue-culture plate. The wild-type crystals in 0.1 M sodium acetate pH 4.6, 1.4 M sodium malonate were transferred into drops consisting of 2 µl cryoprotectant solution (0.1 M sodium acetate pH 4.6, 2.9 M sodium malonate). The substrate complex of GlcNAcase was obtained by soaking the wild-type crystals with the natural substrate (diNAG; 10 mM) in the corresponding mother liquor at 293 K for 5, 10, 15 and 30 min. Afterwards, the soaked crystals were immersed for a few seconds into a cryoprotectant consisting of the mother liquor with 2.9 M sodium malonate and 10 mM diNAG before flash-cooling in liquid nitrogen for further storage.

For crystallization of the GlcNAcase D437A mutant, 0.1 µl protein solution (18 mg ml^{−1} in 20 mM Tris–HCl buffer pH 8.0 containing 150 mM NaCl and 1 mM TCEP) was pipetted into 0.1 µl of each precipitant using the same set of crystal screens as described above. Small plate-shaped crystals (form I) were obtained within 3 d in condition G7 from the PACT Suite [20% (w/v) PEG 3350, 0.1 M bis-tris pH 7.5, 0.2 M sodium acetate], condition D9 from the PEGs Suite [25% (w/v) PEG 6000, 0.1 M Tris–HCl pH 8.5] and condition H1 from the PEGs Suite [20% (w/v) PEG 3350, 0.2 M potassium sodium tartrate]. Three-dimensional rod-shaped crystals (form II) were also obtained within 14 d of incubation in condition C12 from the Anions Suite (0.1 M MES pH 6.5, 1.2 M sodium malonate) at 293 K. The D437A crystals from condition G7 of The PACT Suite were further optimized by the hanging-drop vapour-diffusion method in a 24-well Linbro tissue-culture plate as described for the wild-type enzyme. A 1.5 µl droplet of the D437A mutant was mixed with 1.5 µl of precipitants at various concentrations [17–22% (w/v) PEG 3350, 0.1 M bis-tris pH 7.5, 0.1–0.2 M sodium acetate] and then equilibrated over 1.0 ml of

Table 1

Data-collection and processing statistics.

Values in parentheses are for the outer shell.

	Wild type	Wild type, GlcNAc complex	D437A mutant (form I)	D437A mutant (form II)
Space group	$P2_1$	$P2_1$	$P2_1$	$P4_32_12$
Unit-cell parameters				
a, b, c (Å)	90.2, 130.7, 98.5	91.3, 129.6, 100.0	89.4, 129.3, 98.4	165.1, 165.1, 155.2
α, β, γ (°)	90.0, 113.0, 90.0	90.0, 114.4, 90.0	90.0, 112.2, 90.0	90.0, 90.0, 90.0
Resolution range (Å)	48.55–2.43 (2.43–2.37)	46.21–2.50 (2.56–2.50)	48.43–2.60 (2.67–2.60)	49.50–3.00 (3.08–3.00)
V_M (Å ³ Da ^{−1})	3.59	3.62	3.53	3.55
Solvent content (%)	65.73	66.03	65.22	65.39
Subunits per asymmetric unit	2	2	2	2
Total No. of reflections	577490	342683	436477	569579
No. of unique reflections	84783	72950	63617	43294
Wavelength (Å)	0.99980	0.97889	0.97889	0.97779
Data completeness (%)	99.4 (99.1)	99.4 (99.9)	99.8 (99.7)	99.5 (99.5)
Multiplicity	6.6 (7.1)	4.7 (4.5)	6.9 (7.1)	13.2 (13.9)
Mean $\langle I/\sigma(I) \rangle$	9.93 (2.21)	9.71 (3.17)	9.7 (2.18)	12.8 (2.06)
$R_{\text{merge}}^{\dagger}$ (%)	14.6 (108.7)	11.8 (53.3)	12.5 (77.5)	21.0 (167.1)
$CC_{1/2}$	99.4 (84.1)	99.2 (85.1)	99.3 (86.5)	99.6 (79.0)

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity for the i th measurement of an equivalent reflection with indices hkl .

the respective precipitant. After the D437A crystals had been immersed in a cryoprotectant solution consisting of mother liquor supplemented with 25%(v/v) glycerol, they were rapidly transferred to liquid nitrogen and stored.

2.3. Data collection and processing

For in-house crystal testing and the collection of preliminary data sets, crystals were mounted in nylon loops (Hampton Research, Aliso Viejo, California, USA) on a goniometer and data were collected on a Bruker MICROSTAR or Rigaku MicroMax-007 HF rotating-anode generator with a copper anode as the X-ray source. X-ray diffraction data were collected from high-quality crystals using a Pilatus 6M detector on beamline PX-II at the Swiss Light Source (SLS), Paul Scherrer Institute, Villigen, Switzerland. All X-ray images were recorded using 0.25 Å oscillations at 100 K and a wavelength of 1.0 Å with crystal-to-detector distances of 381, 350, 440 and 500 mm for the wild type, the wild type complexed with GlcNAc, the D437A mutant form I and the D437A mutant form II, respectively. Data indexing was carried out using *iMosflm* (Battye *et al.*, 2011) from the *CCP4* suite (Winn *et al.*, 2011). The data were further processed using *XDS* (Kabsch, 2010). The protein content of the asymmetric unit was estimated by calculating the Matthews coefficient (Matthews, 1968), while the solvent content was calculated based on two subunits of protein, each with six attached histidine residues.

3. Results and discussion

The gene encoding *VhGlcNAcase* was cloned in the pQE60 expression vector and the D437A mutant was generated using the full-length *VhGlcNAcase* cDNA as a template as described above. Both wild-type and mutated GlcNAcases were expressed in *E. coli* M15 (pREP4) cells as C-terminally

His₆-tagged polypeptides that could be purified by affinity chromatography. Both GlcNAcases, the wild type and the D437A mutant, were purified to homogeneity by a two-step protocol involving TALON metal resin affinity chromatography followed by HiPrep (16/60) Sephacryl S-200 HR gel-filtration chromatography. The final yield of the purified enzymes was approximately 20 mg protein per litre of bacterial culture. Fig. 1(*a*) shows a chromatographic profile of the elution of the wild-type protein from a HiPrep 16/60 Sephacryl S-200 prepac column. The enzyme was found in two peaks. The first peak, which eluted at the void volume, was shown by native PAGE to be aggregated protein with a molecular mass that was too large to enter the gel (Fig. 1*b*, lane 1). However the second peak, which eluted at 45–70 ml, was non-aggregated protein that migrated with an apparent molecular mass of 75 kDa, corresponding to the molecular mass of the GlcNAcase monomer (Fig. 1*b*, lane 2). The pooled fraction of each protein peak was found to migrate similarly under denaturing conditions on SDS-PAGE (Fig. 1*c*), indicating that they were likely to contain the same protein. When the GlcNAcase activity was assayed with *p*NP-GlcNAc as the substrate, the second peak showed high activity, while no GlcNAcase activity was detected for the void peak. The chromatographic profile of the D437A mutant was similar to that of the wild-type enzyme, but the void peak was not seen, since the higher salt concentration helped to improve the protein solubility (data not shown).

Crystallization trials with the sitting-drop vapour-diffusion method yielded small crystals from various conditions and initial screening of the crystal quality with an in-house X-ray diffractometer showed that most conditions did not give X-ray-quality crystals, with the exception of those obtained from condition C6 of The Anions Suite (0.1 M sodium acetate pH 4.6, 1.2 M sodium malonate), which diffracted to 2.4 Å resolution. Therefore, crystals of wild-type GlcNAcase grown under this condition were further optimized. Small single

crystals of wild-type GlcNAcase were finally observed from 0.1 *M* sodium acetate pH 4.6, 1.3–1.6 *M* sodium malonate. After optimization, large single crystals of wild-type GlcNAcase appeared as thick three-dimensional plates with approximate dimensions of $400 \times 200 \times 20$ μm (Fig. 2*a*). In the case of the crystal complex of wild-type *Vh*GlcNAcase, soaking the wild-type crystal with the diNAG substrate for a short time (5 min) was attempted but without success. We also tried soaking the crystals of the native enzyme with the natural substrate diNAG for various times, but we could not obtain X-ray diffraction data for the GlcNAcase–diNAG complex because the substrate was degraded by the active enzyme. We eventually obtained X-ray diffraction data for the native enzyme complexed with a single GlcNAc molecule. Fig. 3 shows that the degradation product GlcNAc fitted with full

occupancy into the well defined electron density found in the active site of *Vh*GlcNAcase. In the crystallization of the D437A mutant, small plate-shaped crystals (form I) were obtained during optimization of condition G7 from the PACT screen with a reservoir solution consisting of 20% (w/v) PEG 3350, 0.1 *M* bis-tris pH 7.5, 0.2 *M* sodium acetate. After further optimization, the D437A mutant crystals appeared as thick three-dimensional plates in a condition consisting of 20% (w/v) PEG 3350, 0.1 *M* bis-tris pH 7.5, 0.1 *M* sodium acetate. The D437A crystals grew to approximate final dimensions of $500 \times 300 \times 50$ μm (Fig. 2*b*) in a space group similar to that of the wild-type crystals.

The final statistics for X-ray data-collection and processing of the GlcNAcase crystal variants are shown in Table 1. The refined unit-cell parameters of the wild-type crystal are

Figure 1

Purification of *Vh*GlcNAcase expressed from *E. coli* M15 cells using a HiPrep 16/60 Sephacryl S-200 prepacked gel-filtration column. Chromatography was performed under a low salt concentration, as described in the text. (a) A chromatographic elution profile of wild-type *Vh*GlcNAcase obtained from an ÄKTApurifier system. (b) GlcNAcase-containing fractions obtained from the two peaks were pooled separately and then loaded onto native PAGE followed by Coomassie Blue staining. (c) The same protein fractions were analysed by SDS-PAGE. Lane Std, low-molecular-weight protein markers (labelled in kDa); lane 1, pooled fractions from the void peak; lane 2, pooled fractions from the second peak.

Figure 2

(a) A crystal of wild-type *Vh*GlcNAcase, with dimensions of $400 \times 200 \times 20$ μm , obtained from a hanging-drop vapour-diffusion setup using 0.1 *M* sodium acetate pH 4.6 containing 1.4 *M* sodium malonate. (b) A crystal of the D437A mutant (dimensions of $500 \times 300 \times 50$ μm) obtained from a hanging-drop vapour-diffusion setup using 20% (w/v) PEG 3350, 0.1 *M* bis-tris pH 7.5, 0.1 *M* sodium acetate. Both crystals were obtained within 3 d of incubation at 293 K.

$a = 90.2$, $b = 130.7$, $c = 98.5$ Å, $\beta = 113^\circ$. The crystal contains two molecules per asymmetric unit, with an estimated Matthews coefficient (V_M) of 3.59 Å³ Da⁻¹. The indexing statistics were compatible with the monoclinic space groups $P2$ or $P2_1$. Molecular replacement yielded an unambiguous solution in space group $P2_1$. For the wild-type crystal in complex with GlcNAc, the unit-cell parameters are $a = 91.3$, $b = 129.6$, $c = 100.00$ Å, $\beta = 114.4^\circ$, which are very similar to those of the unliganded wild-type crystal in the same space group. For the D437A mutant crystal, the refined unit-cell parameters are again very similar: $a = 89.4$, $b = 129.3$, $c = 98.4$ Å, $\beta = 112.2^\circ$. D437A mutant crystals also grew as rod-shaped crystals (form II) belonging to the tetragonal space group $P4_32_12$ (Fig. 4), with unit-cell parameters $a = 165.1$, $b = 165.1$, $c = 155.2$ Å and again with two molecules per asymmetric unit. Since the preliminary structural model of the D437A mutant from the tetragonal crystal form contained exactly the same dimer as the monoclinic crystal form (r.m.s.d. of 0.82 Å), but with much weaker diffraction of the crystals, this crystal form was not further investigated. Fig. 5 shows

Figure 3

The electron-density map fitted with a GlcNAc molecule after the wild-type *VhGlcNAcase* crystals had been soaked with the substrate diNAG for 30 min. The structure is presented at 2.5 Å resolution and is contoured at the 1σ level.

Figure 4

Crystals of the D437A mutant belonging to the tetragonal space group $P4_32_12$. Crystals grew as thin rod-shaped crystals within 14 d of incubation in condition C12 from The Anions Suite (0.1 M MES pH 6.5, 1.2 M sodium malonate) at 293 K.

representative diffraction images for both wild-type and monoclinic D437A mutant crystals, which diffracted to 2.4 and 2.6 Å, respectively.

For molecular-replacement (MR) calculations, *Phaser* (McCoy *et al.*, 2007) from the *CCP4* package was employed to obtain an initial solution of the structure of *VhGlcNAcase*. The crystal structure of β -hexosaminidase from *Arthrobacter*

(a)

(b)

Figure 5

X-ray diffraction images of (a) wild-type *VhGlcNAcase* in the absence of diNAG and (b) the inactive D437A mutant with resolutions of 2.4 and 2.6 Å, respectively. The X-ray data were collected on beamline PX-II at the Swiss Light Source, Villigen, Switzerland.

aureus (PDB entry 3rcn; 35% identity to GlcNAcase from *V. harveyi*; Midwest Center for Structural Genomics, unpublished work), which lacks the N-terminal carbohydrate-binding domain, was used as a search model. After ten cycles of refinement, the R_{free} values for the wild type, the wild type in complex with GlcNAc, the D437A mutant (form I) and the D437A mutant (form II) were 25.8, 24.2, 25.0 and 23.4%, respectively.

The preliminary crystallographic data suggested the presence of two protein molecules in the asymmetric unit, with a hydrophilic interface area between two molecules contributed mainly by the N-terminal carbohydrate-binding and catalytic domains, with a surface area of about 1800 Å². PISA analysis (Krissinel & Henrick, 2007) predicted no dimer formation in solution through this interface, which is consistent with the molecular mass of the active enzyme (~75 kDa) obtained from size-exclusion chromatography (Fig. 1). Initial structural refinements showed a well defined electron-density map corresponding to a single GlcNAc molecule in the active site of each wild-type enzyme molecule, indicating that the substrate diNAG was hydrolyzed during the soaking process. We failed to obtain the inactive mutant D437A complexed with diNAG either by both soaking or by co-crystallization, probably because substitution of the Asp437 residue with Ala affects not only the catalytic activity but also the affinity of the enzyme for its substrate. Further refinement to obtain final models of the three GlcNAcase structures is in progress.

Acknowledgements

We thank the beamline staff of beamline PXII at the SLS synchrotron at the Paul Scherrer Institute, Villigen, Switzerland and our colleagues Toni Meinhart, Juliane Benz and Ilme Schlichting (Max Planck Institute for Medical Research, Heidelberg) and Eckhard Hofmann and Falk Syberg (Ruhr-Universität Bochum) for help with data collection. This research was financially supported by the Thailand Research Fund (TRF) through the Royal Golden Jubilee PhD Program (Grant No. PHD/0084/2550) and the German Academic Exchange Service (DAAD; Code No. A/10/98020) to PM. WS was also funded by TRF Basic Research Grant (Grant No. BRG578001).

References

- Actis, L. A., Tolmasky, M. E. & Crosa, J. H. (2011). *Fish Diseases and Disorders*, edited by P. T. K. Woo & D. W. Bruno, pp. 570–605. Wallingford: CABI.
- Armand, S., Tomita, H., Heyraud, A., Gey, C., Watanabe, T. & Henrissat, B. (1994). *FEBS Lett.* **343**, 177–180.
- Arndt, C., Koristka, S., Bartsch, H. & Bachmann, M. (2012). *Methods Mol. Biol.* **869**, 49–53.
- Austin, B. & Zhang, X.-H. (2006). *Lett. Appl. Microbiol.* **43**, 119–124.
- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
- Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
- Drouillard, S., Armand, S., Davies, G. J., Vorgias, C. E. & Henrissat, B. (1997). *Biochem. J.* **328**, 945–949.
- Edelhoch, H. (1967). *Biochemistry*, **6**, 1948–1954.
- Gill, S. C. & von Hippel, P. H. (1989). *Anal. Biochem.* **182**, 319–326.
- Henrissat, B. & Bairoch, A. (1993). *Biochem. J.* **293**, 781–788.
- Henrissat, B. & Davies, G. (1997). *Curr. Opin. Struct. Biol.* **7**, 637–644.
- Jiang, Y.-L., Yu, W.-L., Zhang, J.-W., Frolet, C., Di Guilmi, A. M., Zhou, C.-Z., Vernet, T. & Chen, Y. (2011). *J. Biol. Chem.* **286**, 43004–43012.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Kadokura, K., Rokutani, A., Yamamoto, M., Ikegami, T., Sugita, H., Itoi, S., Hakamata, W., Oku, T. & Nishio, T. (2007). *Appl. Microbiol. Biotechnol.* **75**, 357–365.
- Keyhani, N. O., Li, X.-B. & Roseman, S. (2000). *J. Biol. Chem.* **275**, 33068–33076.
- Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
- Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
- Langley, D. B., Harty, D. W., Jacques, N. A., Hunter, N., Guss, J. M. & Collyer, C. A. (2008). *J. Mol. Biol.* **377**, 104–116.
- Lemieux, M. J., Mark, B. L., Cherney, M. M., Withers, S. G., Mahuran, D. J. & James, M. N. G. (2006). *J. Mol. Biol.* **359**, 913–929.
- Li, X. & Roseman, S. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 627–631.
- Liuxy, P.-C., Lee, K.-K. & Chen, S.-N. (1996). *Lett. Appl. Microbiol.* **22**, 413–416.
- Lombard, V., Ramulu, H. G., Drula, E., Coutinho, P. M. & Henrissat, B. (2014). *Nucleic Acids Res.* **42**, D490–D495.
- Maier, T., Strater, N., Schuetz, C., Klingenstein, R., Sandhoff, K. & Saenger, W. (2003). *J. Mol. Biol.* **328**, 669–681.
- Mark, B. L., Mahuran, D. J., Cherney, M. M., Zhao, D., Knapp, S. & James, M. N. G. (2003). *J. Mol. Biol.* **327**, 1093–1109.
- Mark, B. L., Vocadlo, D. J., Knapp, S., Triggs-Raine, B. L., Withers, S. G. & James, M. N. G. (2001). *J. Biol. Chem.* **276**, 10330–10337.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Myerowitz, R. (1997). *Hum. Mutat.* **9**, 195–208.
- Ohishi, K., Yamagishi, M., Ohta, T., Suzuki, M., Izumida, H., Sano, H., Nishijima, M. & Miwa, T. (1996). *J. Ferment. Bioeng.* **82**, 598–600.
- Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. & Gray, T. (1995). *Protein Sci.* **4**, 2411–2423.
- Prag, G., Papanikolaou, Y., Tavlas, G., Vorgias, C. E., Petratos, K. & Oppenheim, A. B. (2000). *J. Mol. Biol.* **300**, 611–617.
- Suginta, W., Chuenark, D., Mizuhara, M. & Fukamizo, T. (2010). *BMC Biochem.* **11**, 40.
- Suginta, W., Chumjan, W., Mahendran, K. R., Janning, P., Schulte, A. & Winterhalter, M. (2013). *PLoS One*, **8**, e55126.
- Suginta, W., Chumjan, W., Mahendran, K. R., Schulte, A. & Winterhalter, M. (2013). *J. Biol. Chem.* **288**, 11038–11046.
- Suginta, W., Robertson, P. A., Austin, B., Fry, S. C. & Fothergill-Gilmore, L. A. (2000). *J. Appl. Microbiol.* **89**, 76–84.
- Sumida, T., Ishii, R., Yanagisawa, T., Yokoyama, S. & Ito, M. (2009). *J. Mol. Biol.* **392**, 87–99.
- Tews, I., Perrakis, A., Oppenheim, A., Dauter, Z., Wilson, K. S. & Vorgias, C. E. (1996). *Nature Struct. Mol. Biol.* **3**, 638–648.
- Vocadlo, D. J. & Withers, S. G. (2005). *Biochemistry*, **44**, 12809–12818.
- Williams, S. J., Mark, B. L., Vocadlo, D. J., James, M. N. & Withers, S. G. (2002). *J. Biol. Chem.* **277**, 40055–40065.
- Winn, M. D. et al. (2011). *Acta Cryst.* **D67**, 235–242.
- Yuli, P. E., Suhartono, M. T., Rukayadi, Y., Hwang, J. K. & Pyun, Y. R. (2004). *Enzyme Microb. Technol.* **35**, 147–153.