The EMBO Journal (2006) 25, 1569–1578 | © 2006 European Molecular Biology Organization | All Rights Reserved 0261-4189/06
www.embojournal.org

THE EMBO JOURNAL

Structural insights into the mechanism and inhibition of eukaryotic *O*-GlcNAc hydrolysis

Francesco V Rao[1], Helge C Dorfmueller[1], Fabrizio Villa[1,2], Matthew Allwood[1], Ian M Eggleston[1] and Daan MF van Aalten[1,*]

[1]Division of Biological Chemistry & Molecular Microbiology, School of Life Sciences, University of Dundee, Dundee, UK and [2]MRC Protein Phosphorylation Unit, School of Life Sciences, University of Dundee, Dundee, UK

***O*-linked *N*-acetylglucosamine (*O*-GlcNAc) modification of specific serines/threonines on intracellular proteins in higher eukaryotes has been shown to directly regulate important processes such as the cell cycle, insulin sensitivity and transcription. The structure, molecular mechanisms of catalysis, protein substrate recognition/specificity of the eukaryotic *O*-GlcNAc transferase and hydrolase are largely unknown. Here we describe the crystal structure, enzymology and *in vitro* activity on human substrates of *Clostridium perfringens* NagJ, a close homologue of human *O*-GlcNAcase (OGA), representing the first family 84 glycoside hydrolase structure. The structure reveals a deep active site pocket highly conserved with the human enzyme, compatible with binding of *O*-GlcNAcylated peptides. Together with mutagenesis data, the structure supports a variant of the substrate-assisted catalytic mechanism, involving two aspartic acids and an unusually positioned tyrosine. Insights into recognition of substrate come from a complex with the transition state mimic *O*-(2-acetamido-2-deoxy-D-glucopyranosylidene)amino-*N*-phenylcarbamate ($K_i = 5.4$ nM). Strikingly, the enzyme is inhibited by the pseudosubstrate peptide Ala-Cys(-*S*-GlcNAc)-Ala, and has OGA activity against *O*-GlcNAcylated human proteins, suggesting that the enzyme is a suitable model for further studies into the function of human OGA.**

The EMBO Journal (2006) **25,** 1569–1578. doi:10.1038/sj.emboj.7601026; Published online 16 March 2006
Subject Categories: structural biology
Keywords: glycosylation; *O*-GlcNAc; phosphorylation; protein structure; PUGNAc

Introduction

Since its discovery over two decades ago (Torres and Hart, 1984), it has become clear that modification of serines/threonines by an *O*-linked *N*-acetylglucosamine (*O*-GlcNAc) is an essential, abundant and dynamic post-translational process (recently reviewed in Zachara and Hart, 2004; Love and Hanover, 2005). *O*-GlcNAcylated proteins have been detected in both the nucleus and cytoplasm (Torres and Hart, 1984; Holt *et al*, 1986) and the levels of *O*-GlcNAcylation respond to nutrient levels and cellular stress (Slawson *et al*, 2005). *O*-GlcNAc has now been detected on hundreds of proteins (Zachara and Hart, 2004), many of which play key roles in cellular processes. For instance, precise levels of *O*-GlcNAcylation of specific sites on insulin receptor substrate 1 (IRS-1) (Park *et al*, 2005), protein kinase Bβ (Park *et al*, 2005), glycogen synthase kinase 3β (Parker *et al*, 2003) and glycogen synthase (Parker *et al*, 2003) are required for proper insulin sensitivity and response. *O*-GlcNAcylation of transcription factors such as c-Myc (Chou *et al*, 1995) and mSin3A (Yang *et al*, 2002) directly affects their activity, and it is therefore not surprising that proper *O*-GlcNAc levels are required for a normal cell cycle (Slawson *et al*, 2005).

Dynamic protein *O*-GlcNAcylation is achieved by interplay of two essential enzymes—*O*-GlcNAc transferase (OGT) and *O*-GlcNAcase (OGA). Both these enzymes are required for life of the metazoan cell, and are highly conserved from *Caenorhabditis elegans* to man (Zachara and Hart, 2004). Human OGT was recently cloned and characterised in terms of its enzyme activity (Lubas and Hanover, 2000) and domain structure (Lazarus *et al*, 2005). The enzyme is thought to possess up to 13 N-terminal tetratricopeptide repeats (TPR), which have been shown to be essential for recognition of large protein substrates (Lubas and Hanover, 2000), although the precise mechanisms of recognition and specificity are not known. The C-terminal domain contains the OGT activity and is part of the family 41 glycosyltransferases (CAZY GT 41, http://afmb.cnrs-mrs.fr/cazy), for which no structural information is currently available.

OGA is a 92 kDa protein, originally purified from rat spleen, followed by cloning (Dong and Hart, 1994; Gao *et al*, 2001) of human OGA (hOGA), also identified as an antigen expressed by meningiomas (MGEA5) (Heckel *et al*, 1998). Subsequent characterisation showed that this enzyme most likely represents the human hexosaminidase C (HexC) activity, which was identified in addition to the lysozomal HexA/HexB activities (Wells *et al*, 2002). The enzyme directly hydrolyses *O*-GlcNAcylated peptides/proteins and the pseudosubstrate *p*-nitrophenyl-GlcNAc (Wells *et al*, 2002). Bioinformatic (Schultz and Pils, 2002), genetic (Heckel *et al*, 1998; Gao *et al*, 2001) and biochemical (Wells *et al*, 2002; Toleman *et al*, 2004) data have shown that the enzyme contains two catalytic activities, the OGA activity and an additional (histone) acetyl transferase (HAT) activity, both of which have been predicted to reside in the C-terminal half of the protein (Schultz and Pils, 2002). Strikingly, the OGA and HAT activities are thought to act synergistically, opening up the chromatin structure and directly activating transcription factors (Toleman *et al*, 2004).

*Corresponding author. Division of Biological Chemistry & Molecular Microbiology, Wellcome Trust Biocentre, School of Life Sciences, University of Dundee, Dundee DD1 5EH, UK.
Tel.: +44 1382 344 979; Fax: +44 1382 345 764;
E-mail: dava@davapc1.bioch.dundee.ac.uk

Received: 15 December 2005; accepted: 8 February 2006; published online: 16 March 2006

In several studies, the pharmacological agents streptozotocin (STZ) and *O*-(2-acetamido-2-deoxy-D-glucopyranosylidene)amino-*N*-phenylcarbamate (PUGNAc) were used to inhibit OGA activity, thus raising levels of *O*-GlcNAcylated proteins in the cell (reviewed in Zachara and Hart, 2004). PUGNAc is a synthetic competitive inhibitor of hexosaminidases that inhibits all three known human hexosaminidases (HexA/B and OGA) (Horsch *et al*, 1991; Haltiwanger *et al*, 1998). In adipocytes, this compound dramatically reduces insulin sensitivity (Park *et al*, 2005). STZ is selectively toxic to pancreatic β-cells and has been used for decades to generate mouse models of diabetes (Konrad *et al*, 2001).

There is currently no detailed information on the structural mechanism of *O*-GlcNAc hydrolysis, and how OGA recognises its protein substrates and is inhibited by the known inhibitors. Indeed, the precise location of the domain responsible for the OGA activity and the residues involved in catalysis remain unknown. Unfortunately, hOGA, which is a family 84 glycoside hydrolase (GH 84) as defined by the CAZY database, has so far resisted overexpression and purification to quantities and purity required for structural studies.

Here we describe the first crystal structure of a GH 84 (from *Clostridium perfringens*) with significant sequence homology to the hOGA N-terminus, showing distant structural homology to the family 20 GHs, and also report the structure in complex with PUGNAc. The active site contains a tight pocket for a single GlcNAc residue, in agreement with the *N*-acetyl-β-D-glucosaminidase activity and the mutagenesis data we report here. Similar to hOGA, PUGNAc competitively inhibits this enzyme in the low nanomolar range and the complex and mutagenesis data are compatible with a substrate-assisted reaction mechanism as proposed recently (Macauley *et al*, 2005b), involving two conserved aspartic acids. Surprisingly, residues in and surrounding the active site are conserved with human OGA, explaining the inhibition of the enzyme by the pseudosubstrate peptide Ala-Cys(-*S*-GlcNAc)-Ala, and its activity on mammalian protein substrates.

Results and discussion

Definition of the family 84 GH fold

Despite attempts by several groups, it has so far been impossible to generate significant quantities of the human OGA (hOGA) with crystallography-grade purity. However, the analysis of the CAZY database reveals that hOGA is part of the GH 84, along with several bacterial enzymes. To gain insight into the structure and function of hOGA, we selected one of the bacterial enzymes with the highest sequence homology, that of *C. perfringens* NagJ (*Cp*NagJ, 34% sequence identity, 51% sequence similarity with hOGA), for structural and biochemical studies (Figure 1C).

Residues 31–624 of *Cp*NagJ (corresponding to the GH 84 domain plus an additional C-terminal 170 residues; Figure 1C) were cloned, overexpressed in *Escherichia coli*, purified using glutathione affinity chromatography and crystallised from PEG solutions. A preliminary crystallisation study of a similar fragment has been reported recently elsewhere (Ficko-Blean and Boraston, 2005). As no significant sequence homology to known protein structures could be detected, the structure was determined with experimental phases from a zinc single-wavelength anomalous dispersion experiment to 2.25 Å resolution (Table I and Figure 1A), and refined to an *R*-factor of 0.177 ($R_{free} = 0.221$) with good stereochemistry (Table I).

The structure consists of three domains (Figure 1B). The N-terminal domain shows an α/β fold, consisting of a seven-stranded mixed β-sheet, with α-helices on both sides. The middle domain is a classic $(\beta/\alpha)_8$ barrel, with the seventh helix replaced by a loop. The C-terminal domain reveals a five-helical bundle. The DALI server (Holm and Sander, 1993) was used to detect structural homology with known protein structures. Strikingly, despite the absence of any significant sequence conservation, the N-terminal and middle domains of *Cp*NagJ adopt a fold similar to that seen in the family 20 GHs (RMSD = 3.3 Å on 365 equivalenced Cα atoms for the *Streptomyces plicatus N*-acetyl-β-hexosaminidase (*Sp*HeX)). The structures of *Cp*NagJ and *Sp*HeX are compared in Figure 1B, showing that not only most of the secondary structure elements in the N-terminal and middle domains are conserved, but in addition two *Cp*NagJ carboxylate side chains (Asp297/Asp298) occupy equivalent positions to those of the catalytic residues Asp313/Glu314 of *Sp*HeX (Williams *et al*, 2002). Notably, DALI also detected structural similarities with families GH 18 and GH 56, which, like *Sp*HeX, are thought to use the unusual 'substrate-assisted' reaction mechanism (Tews *et al*, 1997; Markovic-Housley *et al*, 2000; van Aalten *et al*, 2001; Williams *et al*, 2002).

A DALI search for a structural homologue of the C-terminal domain, which does not fall within the GH 84 signature, gave no significant hits. Similarly, sequence searches in the NCBI protein database with *Cp*NagJ residues 455–624 revealed no significant hits. GHs frequently contain carbohydrate-binding modules that assist in binding of polymeric substrates (Boraston *et al*, 2004), and it is possible that the C-terminal domain could fulfill such a role, although further experiments are required to verify this.

CpNagJ possesses N-acetyl-β-hexosaminidase activity and is inhibited by PUGNAc and STZ

Although *Cp*NagJ falls within the GH 84 family, its kinetics and substrate specificity are not known. Initial steady-state kinetics experiments with *p*-nitrophenyl-GlcNAc (pNPGlcNAc) (not shown) gave a $K_m = 121\,\mu M/k_{cat} = 0.2\,s^{-1}$, similar to the values reported for hOGA ($K_m = 1.1\,mM/k_{cat} = 1\,s^{-1}$) (Wells *et al*, 2002)). A more robust assay was obtained with 4-methylumbelliferyl-GlcNAc (4MU-GlcNAc), which gave $K_m = 2.9\,\mu M/k_{cat} = 10.5\,s^{-1}$ (Figure 2A and Table II), compared to the reported values for hOGA ($K_m = 880\,\mu M$, (Macauley *et al*, 2005b)) and typical GH 20 enzymes (e.g. $K_m = 48\,\mu M/k_{cat} = 222\,s^{-1}$ for *Sp*HeX (Williams *et al*, 2002)).

The *N*-acetyl-β-hexosaminidase inhibitor PUGNAc (Figure 2B)) (Horsch *et al*, 1991) is a potent hOGA inhibitor that has been extensively used to raise levels of *O*-GlcNAc in cells, by selectively inhibiting hOGA, but not hOGT (Haltiwanger *et al*, 1998; Zachara and Hart, 2004). The compound shows diabetogenic properties, inducing insulin resistance by modulating *O*-GlcNAc/*O*-phosphate levels on various proteins in the insulin signalling pathway (Vosseller *et al*, 2002; Zachara and Hart, 2004). It has been shown to competitively inhibit hOGA with a K_i of 70 nM (Macauley

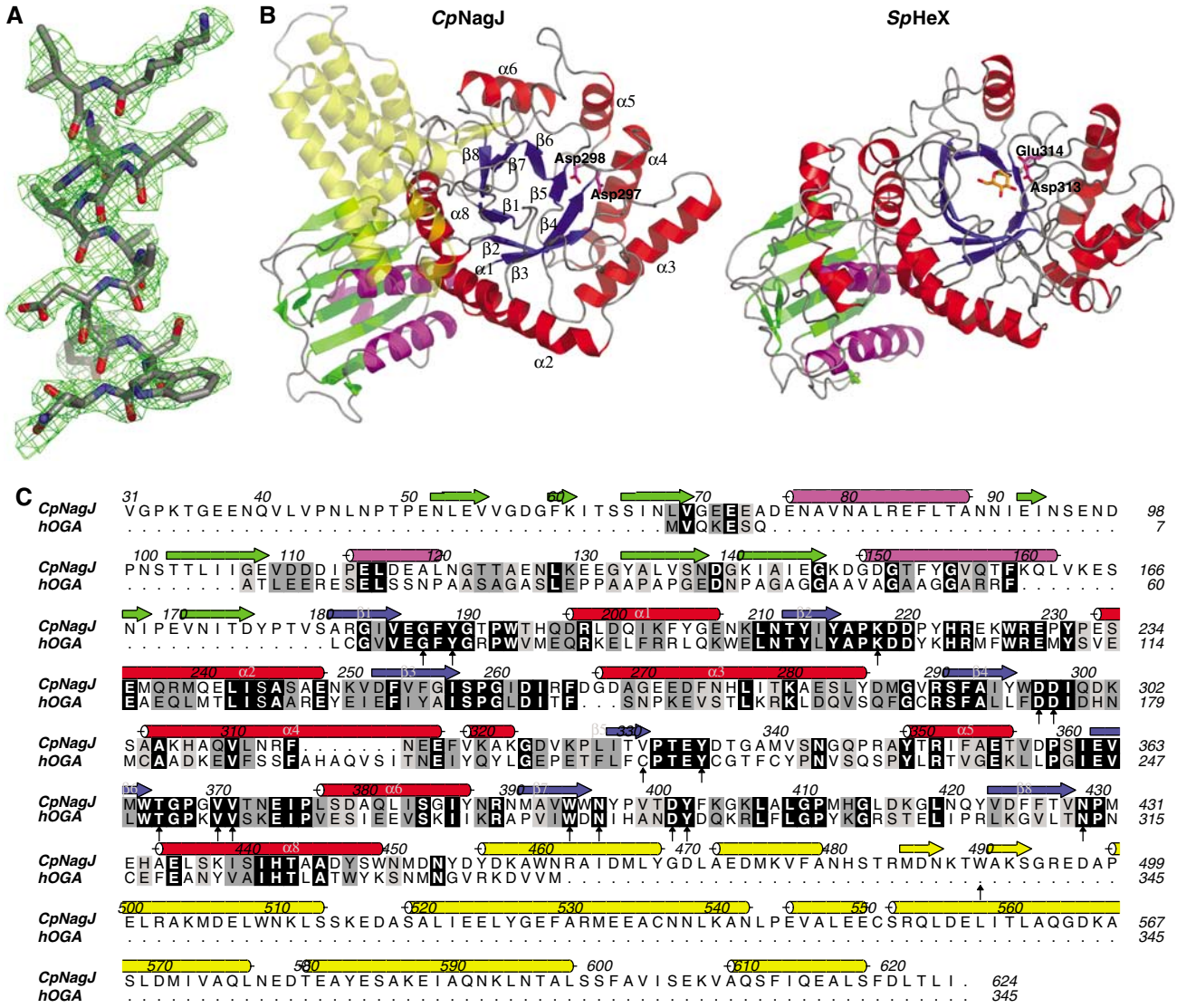

Figure 1 The GH 84 fold. (**A**) Representative section of the 2.25 Å experimental electron density map (green mesh, contoured at 1σ), with the final model for residues 438–450 shown as sticks. (**B**) Comparison of the *Cp*NagJ structure with a representative member (*Sp*HeX, PDB entry 1H15 (Mark *et al*, 2001)) of the family 20 GHs. The structures are shown in a ribbon representation, with red helices and blue strands for the TIM barrel (labelled with α1–8, β1–8), magenta helices and green strands of the N-terminal domain, and transparent yellow helices/strands for the *Cp*NagJ C-terminal domain. The catalytic residues Asp313/Glu314 are shown for the *Sp*HeX structures, the equivalent Asp297/Asp298 are shown for *Cp*NagJ. The thiazoline reaction intermediate mimic, in complex with *Sp*HeX, is shown as a sticks model with orange carbons. (**C**) Sequence alignment of *Cp*NagJ with the hOGA, made with ALINE (CS Bond, personal communication). The *Cp*NagJ secondary structure is indicated, following the colour scheme and labelling as in panel B. Residues in the active site are indicated by arrows.

et al, 2005b). PUGNAc also competitively inhibits *Cp*NagJ with a K_i of 5.4 nM (Figure 2B and Table II), suggesting that the active sites of *Cp*NagJ and hOGA are similar.

STZ (Figure 2C) is a widely used diabetogenic compound, selectively killing β-cells in the pancreas, generating an insulin-deficient mouse model (Konrad *et al*, 2001). There is some evidence that this compounds acts as a suicide inhibitor of OGA (Konrad *et al*, 2001), although this has been controversial (Gao *et al*, 2001; Macauley *et al*, 2005b). Although STZ does inhibit *Cp*NagJ ($IC_{50} = 64$ μM; Figure 2C), we were unable to detect a time-dependent change in inhibition or a covalent enzyme–inhibitor complex by mass spectrometry (data not shown). This is similar to what has been observed in a recent enzymological study of hOGA (Macauley *et al*, 2005b).

PUGNAc inhibits CpNagJ/hOGA by mimicking the transition state

Although the synthesis of PUGNAc (Horsch *et al*, 1991) and its activity against *N*-acetyl-β-hexosaminidases (e.g. Horsch *et al*, 1991; Haltiwanger *et al*, 1998; Macauley *et al*, 2005b) has been known for more than a decade, its structural mode of inhibition has not been defined. To gain insight into the mechanism of PUGNAc inhibition and to study the details of the *Cp*NagJ/hOGA active site, a *Cp*NagJ–PUGNAc complex was determined by soaking the inhibitor into native crystals, followed by refinement against 2.35 Å X-ray diffraction data (R, $R_{free} = 0.177$, 0.232; Table I). Well-defined $|F_o|-|F_c|$, ϕ_{calc} electron density allowed building and refinement of the complete PUGNAc molecule. The inhibitor binds with the GlcNAc sugar deep in a pocket on the enzyme surface,

Table I Details of data collection and structure refinement

	*Cp*NagJ–Zn	*Cp*NagJ–PUGNAc
Wavelength (Å)	1.28202	1.5418
Unit cell (Å)	$a=119.939$	$a=129.613$
	$b=147.380$	$b=145.745$
	$c=157.687$	$c=152.800$
Resolution range (Å)	20.00–2.25 (2.33–2.25)	20.00–2.35 (2.43–2.35)
No. of observed reflections	1 672 845 (120 883)	221 445 (20 790)
No. of unique reflections	65 956 (6502)	56 462 (5727)
Redundancy	25.4 (18.6)	3.9 (3.6)
$I/\sigma I$	7.7 (7.6)	15.8 (2.6)
Completeness (%)	100.0 (100.0)	93.7 (96.0)
R_{merge}	0.114 (0.477)	0.075 (0.609)
No. of protein residues	1170	1170
No. of water molecules	472	587
R, R_{free}	0.177, 0.221	0.177, 0.232
RMSD from ideal geometry		
Bonds (Å)	0.017	0.013
Angles (deg)	1.5	1.3
B-factor RMSD (Å^2) (backbone bonds)	1.33	1.43
$\langle B \rangle$ (Å^2)		
Protein	23.2	49.4
Inhibitor	—	54.0
Water	23.2	49.7

Values between brackets are for the highest resolution shell. All measured data were included in structure refinement. The space group for both crystals was $I2_12_12_1$.

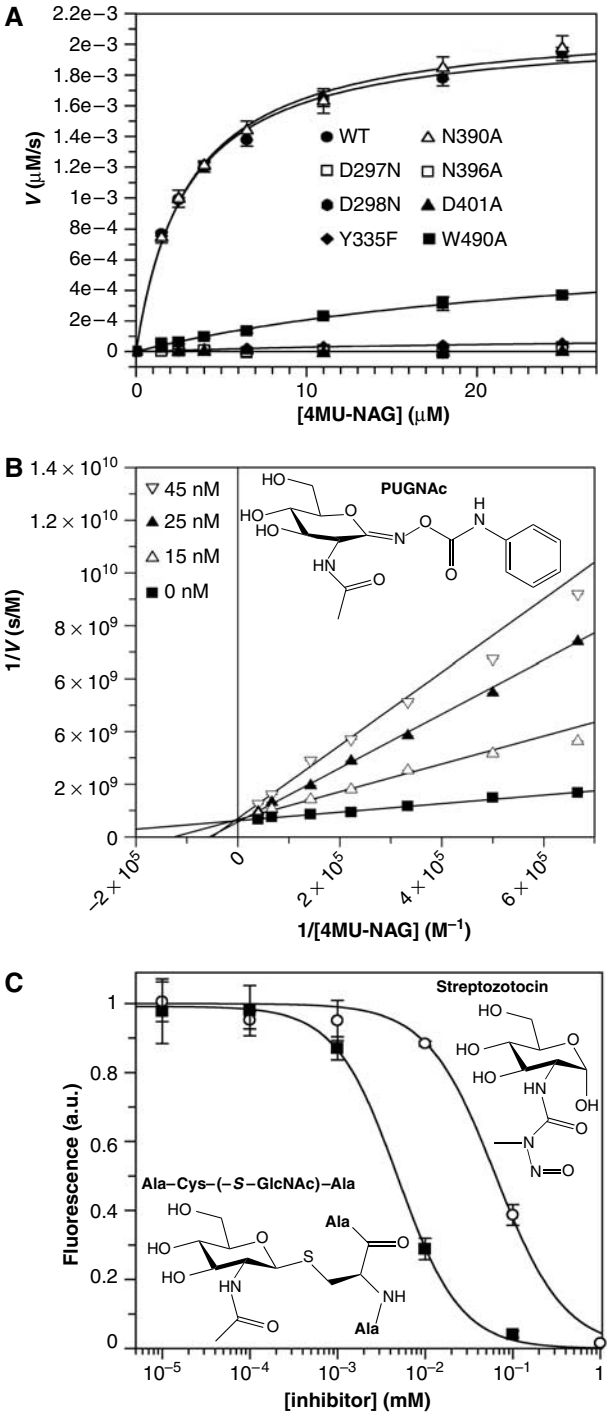

Figure 2 Enzymology. (**A**) Steady-state kinetics for wild-type and mutant *Cp*NagJ, with Michaelis Menten parameters shown in Table II. (**B**). Chemical structure of PUGNAc and Lineweaver–Burk analysis showing competitive inhibition, with $K_i=5.4$ nM ($K_m=2.6$ μM, $k_{cat}=8.0\,s^{-1}$). (**C**) Chemical structures of streptozotocin and the inhibitory peptide together with dose–response curves, giving IC_{50}'s of 64 ± 10 and 4.9 ± 0.8 μM, respectively.

whereas the phenylcarbamate moiety projects towards the solvent. The oxime is in the *Z* configuration, consistent with a recent study showing that inhibition of hOGA with PUGNAc is dependent on the *Z* oxime stereochemistry (Perreira *et al*, 2006). The *E* oxime stereoisomer would not fit the shape of the *Cp*NagJ active site. The sp^2 hybridisation of the C1 carbon helps the pyranose ring assume a ^4E envelope conformation. Strikingly, the oxygen of the 2-acetamido group approaches the C1 carbon to within 3.0 Å. This has also been observed in the GH 18, 20 and 56 families, which have been shown to employ substrate-assisted catalysis, where the oxygen of the 2-acetamido group, rather than a protein side chain, is the catalytic nucleophile, leading to formation of an oxazolinium intermediate (Terwisscha van Scheltinga *et al*, 1994; Tews *et al*, 1997; Markovic-Housley *et al*, 2000; Mark *et al*, 2001; van Aalten *et al*, 2001). Indeed, an elegant mechanistic enzymology study of hOGA has recently shown that it uses the same substrate-assisted double-displacement mechanism with overall retention of stereochemistry (Macauley *et al*, 2005b), compatible with the structural data described here. The GlcNAc moiety of PUGNAc appears to be tightly tethered in the active site, through 10 hydrogen bonds to nine residues, which are all conserved in the hOGA sequence (Figures 1C and 3). This tight hydrogen-bonding network explains why, unlike the human GH 20 hexosaminidases, hOGA is not inhibited by GalNAc (Gao *et al*, 2001). Asp298 and Tyr335 are within hydrogen bonding distance of the oxime nitrogen (approximately equivalent to the position of the substrate glycosidic oxygen), while Asp297 and Asn396 influence the conformation of the acetamido group. The 3, 4 and 6 sugar hydroxyls point deep into the pocket, where they are hydrogen bonded to Lys218, Gly187, Asn429 and, in particular, Asp401. Notably, there are only two residues that contact the inhibitor and that are not conserved between *Cp*NagJ and hOGA. Trp490 projects from the *Cp*NagJ C-terminal domain towards the catalytic domain, and stacks with the PUGNAc phenyl ring, similar to what is observed for the interaction

Table II Steady-state kinetics and PUGNAc inhibition of wild-type and mutant *Cp*NagJ

	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\mu M^{-1} s^{-1}$)	PUGNAc IC_{50} (nM)
WT	2.9±0.2	10.5±0.2	3.6	8.6±0.8
D297N[a]	3.0±0.7	0.0198±0.0006	0.001	ND
D298N[a]	4.0±1.4	0.0013±0.0001	0.0003	ND
Y335F[a]	41±4	0.0028±0.0001	0.0001	ND
N390A	3.1±0.2	11.0±0.2	3.6	10.0±1.4
N396A[a]	13±8	0.0025±0.0005	0.0002	ND
D401A[a]	15±9	0.0044±0.0007	0.0003	ND
W490A	28±5	4.0±0.4	0.1	24.0±6.5

[a]Approximate K_m and k_{cat} values for these rather inactive mutants were determined from a kinetics run over a 9-h period with 5–250 μM substrate concentrations.
Steady-state kinetics data (see also Figure 2A) were fitted to the standard Michaelis Menten equation, with the resulting K_m and k_{cat} values shown below. IC_{50} values for inhibition with PUGNAc are also shown. ND = not determined.

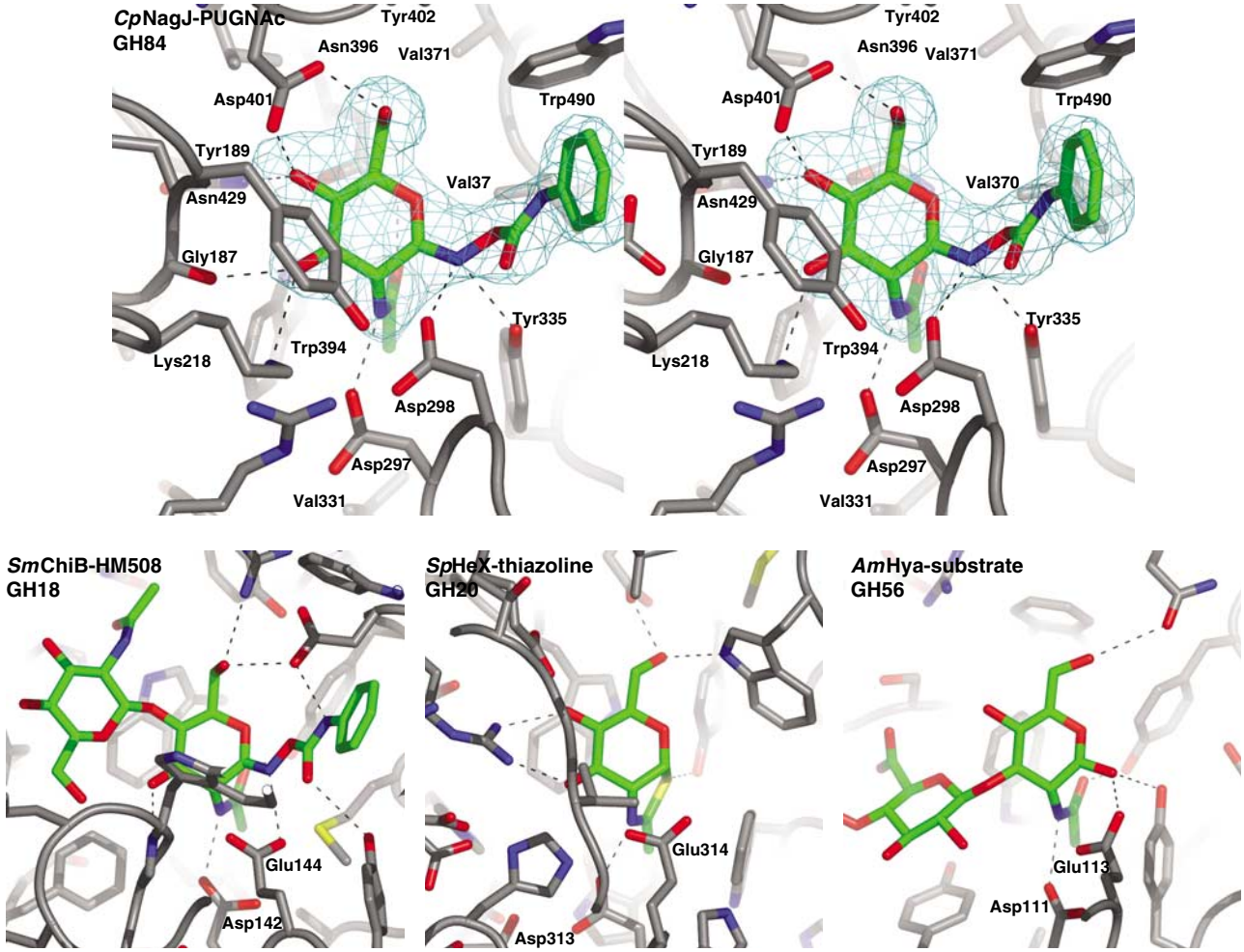

Figure 3 Details of the active site. The *Cp*NagJ–PUGNAc complex is compared to the *Sm*ChiB–HM508 complex (PDB entry 1UR9 (Vaaje-Kolstad *et al*, 2004)), the *Sp*HeX–thiazoline complex (PDB entry 1H15 (Mark *et al*, 2001)) and the *Am*Hya–substrate complex (PDB entry 1FCV (Markovic-Housley *et al*, 2000)). The structures were superimposed using the C2, C3, C5 and O5 atoms of the ligands in the −1 subsite. The ligands are shown as sticks with green carbons. For PUGNAc, the unbiased 2.35 Å $|F_o|-|F_c|$, ϕ_{calc} electron density map is shown in cyan, contoured at 2.5σ. The *Cp*NagJ–PUGNAc complex is shown in stereo, and residues contacting the inhibitor are labelled (see also Figure 1C). For the other complexes the two key carboxylate side chains are identified. The CAZY GH family numbers are also shown.

of the PUGNAc-related compound, HM508, with a GH 18 chitinase (Vaaje-Kolstad *et al*, 2004) (Figure 3). The other nonconserved residue is Val331, at the very bottom of the active site (Figure 3). Interestingly, this residue is a cysteine in hOGA, compatible with the early observation that the enzyme can be inactivated by a thiol-reactive compound (Dong and Hart, 1994). Notably, a cysteine close to the active site of the protein phosphatase PTP1B has recently been shown to play a key role in redox-dependent regulation of phosphatase activity (Salmeen *et al*, 2003). Further experiments are currently in progress to study a potential similar mechanism in hOGA.

The CpNagJ/hOGA active sites are similar to GH 18, 20 and 56

Comparison of *Cp*NagJ with GH 18, 20 and 56 structures reveals a number of significant differences and similarities in the active sites, that allow the putative identification of residues involved in catalysis. *Cp*NagJ Asp298 occupies a position equivalent to a conserved glutamic acid in the GH 18, 20 and 56, which has been shown to be the catalytic acid, protonating the glycosidic bond in the first step of the reaction (Tews *et al*, 1997; Markovic-Housley *et al*, 2000; Mark *et al*, 2001; van Aalten *et al*, 2001). Unusually, Tyr335, also observed in GH 56, but not in GH 18/20, is also within hydrogen-bonding distance of the glycosidic oxygen (Figure 3). *Cp*NagJ Asp297 appears to be structurally conserved in the GH 18, 20 and 56 enzymes, where it is thought to stabilise the conformation of the acetamido group and the developing positive charge on the oxazolinium intermediate (Tews *et al*, 1997; Markovic-Housley *et al*, 2000; van Aalten *et al*, 2001; Williams *et al*, 2002). *Cp*NagJ Asn396 hydrogen bonds the acetamido oxygen, a role fulfilled by a tyrosine in GH 18, 20 and 56 enzymes (Figure 3).

To understand the contributions of active site residues to substrate/inhibitor binding and catalysis, seven mutants were investigated (Figure 2A and Table II). Asn390, a surface-exposed residue away from the active site, was mutated to alanine as a control, and showed kinetic parameters indistinguishable from the wild-type enzyme. The remaining mutants approximately fall into two distinct groups, mainly affecting either binding of the substrate (effects on K_m) or catalysis (effects on k_{cat}). Residues Asp401 and Trp490 make interactions with the PUGNAc inhibitor on the verges of the active site (Figure 3) and, in agreement with this, mutations of these residues (D401A, W490A) have significant effects on the ability of the enzyme to bind the substrate/inhibitor (5–10-fold increase in K_m) (Figure 2A and Table II).

Mutations of several residues showed significant effects on catalysis. Asp297 and Asn396 are involved in stabilising the conformation of the acetamido group and the oxazolinium intermediate, compatible with their 530- and 4200-fold reduction in k_{cat}, respectively. Similar selective effects on k_{cat} have also been demonstrated for mutation of the equivalent residues (an Asp and a Tyr, respectively, see Figure 3) in the GH 18 (e.g. van Aalten *et al*, 2001; Bokma *et al*, 2002) and GH 20 (e.g. Williams *et al*, 2002) families. This has been taken as evidence for the involvement of these side chains in a substrate-assisted catalysis mechanism.

In addition to the increase in K_m for the D401A mutation noted above, there is also a significant drop in k_{cat} (2400-fold; Figure 2A and Table II). Asp401 interacts with the O4/O6 hydroxyls, well away from the glycosidic bond (Figure 3). It is possible that the tight hydrogen bonds to these hydroxyls aid in formation of the ^{4}E envelope conformation of the pyranose ring in the transition state, providing a possible explanation for the large mutational effect on k_{cat}.

Asp298, proposed to be the catalytic acid, and Tyr335, of hitherto unknown function, are both in a position to interact with the glycosidic oxygen (Figure 3). Compatible with its proposed role as protonator of the glycosidic bond, mutation of Asp298 to Asn reduces k_{cat} 8100-fold (Figures 2 and 3 and Table II). Strikingly, mutation of Tyr335 to Phe dramatically also reduces activity (3800-fold). GH 18 and GH 20 enzymes do not possess an equivalent of this residue, whereas it is observed in the only GH 56 structure available (Markovic-Housley *et al*, 2000) (Figure 3). Apart from the glycosidic oxygen, there are no hydrogen-bonding donors or acceptors nearby the tyrosine hydroxyl (it is 4.8 Å away from the pyranose ring oxygen), suggesting that it is protonated in the PUGNAc/Michaelis complex. Interestingly, a very recent study has suggested that hydrolysis of *O*-glycosides by hOGA proceeds via a late transition state in which cleavage of the glycosidic linkage is well advanced (Macauley *et al*, 2005a). Tyr335 would then be well positioned to stabilise any partial negative charge that develops at the glycosidic oxygen in such a transition state, consistent with the requirement of this residue for catalysis.

CpNagJ possesses OGA activity against human proteins and is inhibited by an S-GlcNAc peptide

As discussed above, the *Cp*NagJ active site is highly conserved with hOGA, and the enzymes show similar kinetic parameters for pseudosubstrates. Strikingly, the residues surrounding the active site are also highly conserved between *Cp*NagJ and hOGA (Figure 4A). Whereas the GlcNAc binds deep in a conserved pocket, there also appears to be sequence conservation in the bottom and the walls of a groove running over the surface of the protein, perhaps representing the peptide-binding site in case of hOGA. This leads to the hypothesis that, although the natural substrate of *Cp*NagJ is probably a GlcNAc-containing carbohydrate polymer (e.g. hyaluronan/chitin/peptidoglycan), the enzyme may also be suitable as a model system to study hydrolysis of *O*-GlcNAcylated peptides/proteins *in vitro*. Indeed, the enzyme is able to hydrolyse *O*-GlcNAcylated proteins in Swiss 3T3 cell lysate (Figure 4B), although some proteins appear not to be a substrate. It is possible that on these proteins the *O*-GlcNAc sites are masked, or that the unconserved Trp490 residue in *Cp*NagJ (Figure 4A) prevents binding of some substrate proteins. The inactive D401A mutant of *Cp*NagJ does not show activity against *O*-GlcNAcylated proteins. In agreement with the competitive inhibition of PUGNAc, this compound inhibits the OGA activity of *Cp*NagJ in the cell lysates completely at 10 μM (Figure 4B). As a final verification of the ability to bind *O*-GlcNAcylated peptides, we synthesised the thioglycosidic peptide Ala-Cys-(*S*-GlcNAc)-Ala. Thioglycosides have previously been successfully used to trap glycosidase–substrate complexes (e.g. Varrot *et al*, 2003), and *S*-linked glycopeptides have attracted attention as glycopeptide mimetics due to their enhanced chemical stability and resistance to glycosidases (e.g. Zhu *et al*, 2004). Strikingly, *Cp*NagJ is potently inhibited by this peptide ($IC_{50} = 4.9$ μM; Figure 2C), providing further evidence for the ability of *Cp*NagJ, like hOGA, to bind to and hydrolyse *O*-GlcNAcylated peptide substrates. Interestingly, data from a recent study suggest that inhibition with this peptide may be transient, as it was shown that hOGA is able to hydrolyse model thioglycoside substrates with aromatic leaving groups (Macauley *et al*, 2005a). However, such (good) leaving groups are significantly different from those in the true peptide substrates and further work will be required to study this.

Concluding remarks

The data presented here give the first detailed structural and mechanistic insights into a bacterial GH 84 with significant

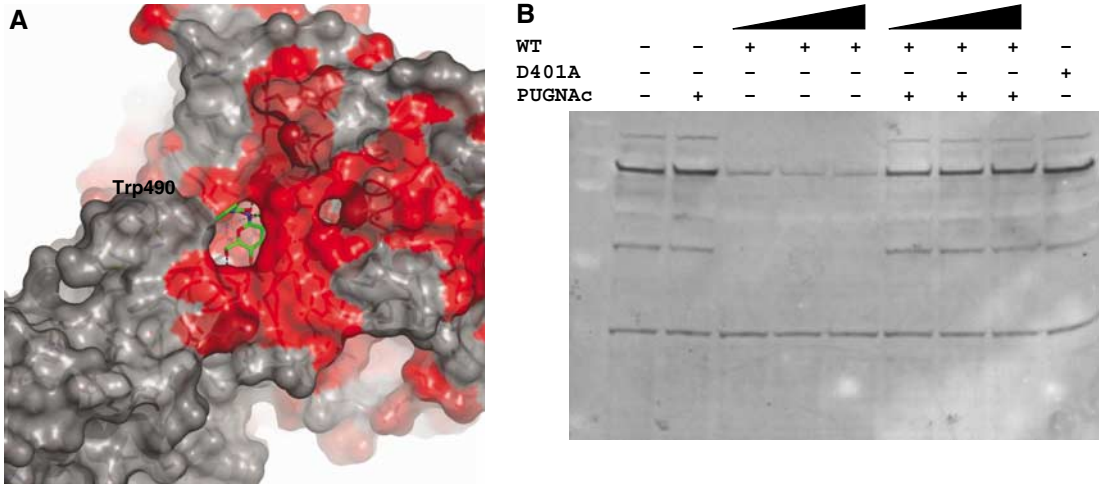

Figure 4 Surface sequence conservation and OGA activity. (**A**) The protein surface of the *Cp*NagJ–PUGNAc complex is shown in three different colours to reflect sequence conservation with the hOGA: grey (not conserved), red (identical) and dark red (similar), following sequence conservation in the alignment shown in Figure 1C. (**B**) Western blot analysis of *O*-GlcNAc levels in Swiss 3T3 cell lysates, after incubation with different amounts of the wild-type enzyme (0.5, 3 and 5 μM), the inactive D401A mutant (5 μM) and the PUGNAc inhibitor (10, 50 and 100 μM). *O*-GlcNAc was detected with a commercially available antibody.

sequence homology to the hOGA. The data show that the *N*-acetyl-β-hexosaminidase activity must be, by sequence homology, located to the N-terminal domain of hOGA, in contrast to what has been proposed previously by a bioinformatics study (Schultz and Pils, 2002). The active sites of the human and bacterial enzyme are, with the exception of two residues, highly conserved, and the sequence conservation extends to the groove around the active site, where peptide binding and recognition take place. The GlcNAc portion of the substrate is tightly bound in a deep pocket, by hydrogen bonds and stacking with conserved aromatic residues. This is somewhat similar to the protein phosphatases, where the *O*-linked phosphate binds deep in a conserved pocket, while the peptide binds in a groove on the surface (e.g. Jia *et al*, 1995; Salmeen *et al*, 2000). Interestingly, hOGA possesses a cysteine in the active site, but away from the catalytic center—similar to the tyrosine phosphatases, where this residue has been shown to be involved in redox-dependent regulation of the enzyme (Salmeen *et al*, 2003).

The structure and enzymology suggest an unusual variant of the substrate-assisted catalysis mechanism previously identified for the GH 18, 20 and 56 enzymes (Tews *et al*, 1997; Markovic-Housley *et al*, 2000; van Aalten *et al*, 2001; Williams *et al*, 2002). In these families the catalytic acid is a glutamic acid residue on the protein, whereas in the GH 84 enzymes it is an aspartic acid (Asp298). Furthermore, a tyrosine appears to be involved in stabilisation of the transition state. An equivalent tyrosine has been observed in the GH 56 bee venom hyaluronidase (Markovic-Housley *et al*, 2000), but its function was hitherto unknown. After more than a decade of use as a modulator of cellular *O*-GlcNAc levels, the PUGNAc complex reveals that this compound inhibits OGA by mimicry of the transition state through the geometry at the C1 carbon. Tight stacking interactions are observed between the PUGNAc acetamido group and a conserved trytophan/tyrosine pair. Interestingly, a recent mechanistic enzymology study has shown that hOGA possesses a cavity near the acetamido methyl group, whereas such cavity is absent in the GH 20 hexosaminidases (Macauley *et al*, 2005b). This is supported by the structure described here, where Val331 forms the bottom of a pocket, whereas a tryptophan tightly packs against the acetamido methyl in the GH 20 enzymes.

In agreement with the sequence homology between the hOGA and the bacterial enzyme, *Cp*NagJ shows OGA activity on several proteins in human cell lysate, and the enzyme is also inhibited by a thioglycosidic inhibitory peptide. This suggests that *Cp*NagJ may be a suitable model for further studies into the mechanisms of recognition/specificity of *O*-GlcNAcylated peptides. Furthermore, the structural data presented here could aid in rational design of more potent/drug-like OGA inhibitors, which could be used to probe or modulate the role of *O*-GlcNAc in diseases such as Alzheimer's, diabetes and cancer.

Materials and methods

Cloning, expression and crystallisation

A fragment corresponding to the N-terminal GH 84 catalytic domain of *Cp*NagJ (GenBank accession no. BAB80940) was PCR amplified (forward primer, 5′-GGATCCGTAGGACCTAAAACTGGG-3′; reverse primer, 5′-CTCGAGTTATCATATTAATGTTAAATCAAAACTTAAAGC-3′) from genomic DNA from *C. perfringens* (Sigma D5139). The PCR product was ligated into pCR 2.1—TOPO (Invitrogen) and subcloned into the pGEX-6P-1 vector (Amersham Biosciences) using the *Bam*HI and the *Xho*I restriction sites. Single amino-acid residue changes were made using the Quick Change Site Directed Mutagenesis Kit (Stratagene), following the manufacturer's protocol. All plasmids were verified by DNA sequencing (The Sequencing Service, School of Life Sciences, University of Dundee, Scotland, UK). Upon sequencing of several independent PCR reactions, it became apparent that there were several nucleotide changes on the template DNA, compared to the GenBank entry. These thus represent strain-specific changes, which result in six conservative amino-acid substitutions that are located in surface loops of the protein.

The *Cp*NagJ–pGEX-6P-1 construct was transformed into *E. coli* BL21 (DE3) pLysS cells. Cells were grown overnight in Luria-Bertani medium (LB) + ampicillin (100 μg/ml). From this culture, 10 ml of cells were used to inoculate 1 l of LB media. The cells were grown to $OD_{600} = 0.5$ before expression was induced by the addition of 250 μM of isopropyl-β-D-thiogalactopyranoside, and then the cells

were cultured overnight at room temperature. The cells were harvested by centrifugation at 2500 *g* for 30 min, flash frozen in liquid nitrogen, thawed at 37°C, and resuspended in 25 ml of buffer A (50 mM HEPES, 250 mM NaCl, pH 7.5) with half a protease inhibitor tablet (Roche). Cells were lysed by the addition of 10 mg of DNAse-1 and the use of a cell disrupter at 30 K psi. The lysate was centrifuged at 18 900 *g* for 30 min at 4°C and passed through a 0.45 μm filter. The filtrate was then incubated at 4°C for 2.5 h on a rotating platform with 2 ml of glutathione-Sepharose beads (Amersham Biosciences), prewashed with buffer A. The N-terminal GST tag was removed from the GST–*Cp*NagJ fusion protein by incubating the beads with PreScission Protease (80 μg) at 4°C overnight. The supernatant of the beads and a subsequent wash were passed over a Bio-Rad 20 ml disposable column to remove the beads. The resulting filtrate was concentrated to 4 ml, and loaded onto a Superdex 75, 26/60 gel filtration column pre-equilibrated in buffer A. The pure fractions were verified by SDS–PAGE, pooled and then dialysed overnight in 50 mM Tris (pH 8.0) at 4°C. The *Cp*NagJ mutants were expressed and purified using an identical protocol, exluding the final gel filtration step.

Pure *Cp*NagJ protein was spin concentrated to approximately 28 mg/ml. Vapour diffusion crystallisation experiments were set up by mixing 1 μl of protein, 1 μl of mother liquor (0.2 M ammonium sulfate, 0.1 M sodium cacodylate, pH 6.5, and 30% PEG 8000) and 0.25 μl of 40% v/v γ-butyrolactone. Rod-shaped crystals appeared after 4 days, growing to a maximum size of approximately $0.3 \times 0.1 \times 0.1$ mm.

Data collection, structure solution and refinement

The crystals were cryoprotected by a 5 s immersion in a solution containing 0.17 M ammonium sulfate, 0.085 M sodium cacodylate, pH 6.5, 25.5% PEG 8000 and 15% v/v glycerol, and then frozen in a nitrogen cryostream. Crystals used for phasing were soaked in 1.5 M $ZnSO_4$ for approximately 5 s and cryoprotected as above. An inhibitor complex of PUGNAc (Toronto Research Chemicals Inc., Canada) was obtained by adding 0.25 μl of 10 mM PUGNAc to the crystals for 10 min, followed by cryoprotection.

A 25-fold redundant 2.25 Å single-wavelength anomalous dispersion data set was collected from the $ZnSO_4$-soaked crystals at beamline BM14 at the European Synchrotron Radiation Facility (ESRF). Initial phases were calculated with SOLVE (Terwilliger and Berendzen, 1999) to 3.25 Å (figure of merit = 0.35), from the 11 Zn sites located. With the help of PROFESS (Collaborative Computational Project, 1994), two-fold noncrystallographic symmetry was detected, compatible with eight of the 11 sites. Solvent flattening, phase extension to 2.25 Å and twofold averaging was then performed with DM (Cowtan, 1994), yielding a good-quality electron density map. This was used as input for warpNtrace (Perrakis *et al*, 1999), which was able to build 976 out of 1186 possible residues, covering two molecules in the asymmetric unit. Iterative model building (with O (Jones *et al*, 1991) and COOT (Emsley and Cowtan, 2004)) and refinement (with CNS (Brunger *et al*, 1998) and REFMAC (Murshudov *et al*, 1997)) then yielded the final model with statistics shown in Table I.

Diffraction data of PUGNAc-soaked crystals were collected on a rotating anode to 2.35 Å (Table I). Refinement was initiated from the native structure, immediately revealing well-defined $|F_o|-|F_c|$, ϕ_{calc} electron density for the inhibitor, which was built in with the help of PRODRG (Schuettelkopf and van Aalten, 2004)-generated inhibitor structure and topology. Further refinement and model building then yielded the final model with statistics shown in Table I.

Enzymology

Steady-state kinetics of wild-type and mutant *Cp*NagJ were determined using the fluorogenic substrate 4-methylumbelliferyl-*N*-acetyl-β-D-glucosaminide (4MU-NAG; Sigma). Standard reaction mixtures (50 μl) contained 0.2 nM enzyme in 50 mM citric acid 125 mM NaH_2PO_4, pH 5.5, 0.1 mg/ml BSA and 1.5–25 μM of substrate in water. The reaction mixture was incubated for 7 min at 20°C (RT). The reaction was stopped by the addition of a twofold excess (100 μl) of 3 M glycine-NaOH, pH 10.3. The fluorescence of the released 4-methylumbelliferone was quantified using an FLX 800 Microplate Fluorescence Reader (Bio-Tek), with excitation and emission wavelengths of 360 and 460 nm, respectively. The production of 4-methylumbelliferone was linear with time for the incubation period used, and less than 10% of the available substrate was hydrolysed. Experiments were performed in triplicate; all of the spectra were corrected for the background emission from the buffer and the protein. Michaelis–Menten parameters were obtained by fitting the fluorescence intensity with GraFit (Leatherbarrow, 2001).

The IC_{50}'s of Ala-Cys(*S*-GlcNAc)Ala, STZ and PUGNAc against wild type and mutant *Cp*NagJ were determined using the same protocol, but with the following changes: the enzyme was incubated for 1 min with the corresponding inhibitor (1 nM–1 mM) prior to starting the reaction by the addition of a constant substrate concentration equivalent to the K_m determined by steady-state kinetics (e.g. 2.9 μM for wild type). All parameters were obtained by fitting the fluorescence intensity data to the standard IC_{50} equation in the software GraFit (Leatherbarrow, 2001).

Determination of the PUGNAc K_i was performed by steady-state kinetics in the presence of different concentrations (0, 15, 25 and 45 nM) of the inhibitor. The mode of inhibition was visually verified by a Lineweaver–Burk plot, and the K_i determined by fitting all fluorescence intensity data to the standard equation for competitive inhibition in GraFit (Leatherbarrow, 2001).

Western blotting

Swiss 3T3 cells were cultured using standard techniques and lysed in a buffer containing 50 mM Tris–HCl, pH 7.5, 1 mM EDTA, 1 mM EGTA, 270 mM sucrose, 0.1% β-mercaptoethanol, 1 mM Na_3VO_4, 1% (v/v) Triton X-100, 50 mM NaF, 5 mM sodium pyrophosphate and proteinase inhibitor cocktail (one tablet/25 ml). For the OGA assay, 100 μg of cell lysate was incubated with various amounts of enzyme and inhibitor for 30 min at 20°C in a final volume of 100 μl. For Western-blot purposes, 50 μl was separated on a 10% SDS–PAGE gel and transferred onto a nitrocellulose membrane. The membrane was blocked in 50 mM Tris–HCl, pH 7.5, 0.15 M NaCl, 0.1% (v/v) Tween (TBS-Tween) and 10% (w/v) BSA for 1 h. The membrane was then incubated with the same buffer for 16 h at 4°C in the presence of 1 μg/ml of *O*-GlcNAc primary antibody (Pierce). Detection was performed using horseradish peroxidase-conjugated secondary antibody and the enhanced chemiluminescence (ECL Amersham Pharmacia Biotech) reagent.

Synthesis of Ala-Cys-(S-GlcNAc)-Ala

N-(*tert*-butoxycarbonyl)-*O*-β-(2-acetamido-2-deoxy-3,4,6-tri-*O*-acetyl-β-D-glucopyranosyl)-L-cysteine benzyl ester (Boc-L-Cys(*S*-GlcNAc(Ac)$_3$-β-D)-OBn) was synthesised from *N*-(tert-butoxycarbonyl)-L-iodoalanine (Aldrich) using the method of Ohnishi *et al* (2000). The sugar amino acid was then elaborated to the fully protected glycopeptide derivative, Z-L-Ala-L-Cys(*S*-GlcNAc(Ac)$_3$-β-D)-L-Ala-OBn, by standard peptide synthesis methodology, followed by removal of the *N*-benzyloxycarbonyl (Z) and benzyl ester (Bn) protection by catalytic transfer hydrogenolysis (Anwer and Spatola, 1980), and *O*-deacetylation by treatment with sodium methoxide in anhydrous methanol (Paulsen and Holck, 1982). The final glycopeptide was purified by reversed-phase HPLC on a Dionex HPLC system equipped with a Phenomenex Gemini 5 μM C-18 (250×10 mm) column (mobile phase A = 0.1% trifluoroacetic acid in H_2O; mobile phase B = 0.1% trifluoroacetic acid in acetonitrile; linear gradient 5–95% B in 10 min; flow rate 2.5 ml/min), and was characterised by ^{1}H NMR, low-resolution and high-resolution electrospray mass spectrometry.

Acknowledgements

We thank the European Synchrotron Radiation Facility, Grenoble, for the time at beamline BM14, the EPSRC National Mass Spectrometry Service (University of Swansea, UK) for accurate mass analysis, Nick Leslie for the Swiss 3T3 cell lysates and Mike Ferguson for *Cp*NagJ-streptozotocin mass spectrometry. DvA is supported by a Wellcome Trust Senior Research Fellowship and the EMBO Young Investigator Programme. FVR is supported by a BBSRC CASE studentship together with Syngenta, and HCD is supported by the University of Dundee Alumni Studentship. MA is supported by a Wellcome Trust studentship. The coordinates and structure factors have been deposited with the PDB (entries 2CBI, 2CBJ).

Note added in proof

While completing the proofs of this manuscript, another report on a bacterial family 84 glycoside hydrolase has appeared (RJ Dennis, EJ Taylor, MS MacAuley, KA Stubbs, JP Turkenburg, SJ Hart, G Black,

DJ Vocadlo, GJ Davies (2006) Structure and mechanism of a bacterial β-glucosaminidase having *O*-GlcNAcase activity, *Nat Struct Mol Biol*, in press). In agreement with the work described here, this report provides solid structural evidence for a substrate-assisted catalytic mechanism and describes the structural similarities with the family 18/20/56 glycoside hydrolases.

References

Anwer MK, Spatola AF (1980) An advantageous method for the rapid removal of hydrogenolysable protecting groups under ambient conditions—synthesis of leucine-enkephalin. *Synthesis Stuttgart* **11:** 929–932

Bokma E, Rozeboom HJ, Sibbald M, Dijkstra BW, Beintema JJ (2002) Expression and characterization of active site mutants of hevamine, a chitinase from the rubber tree *hevea brasiliensis*. *Eur J Biochem* **269:** 893–901

Boraston AB, Bolam DN, Gilbert HJ, Davies GJ (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J* **382:** 769–781

Brunger AT, Adams PD, Clore GM, Gros P, Grosse-Kunstleve RW, Jiang J-S, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography and NMR system: a new software system for macromolecular structure determination. *Acta Crystallogr D* **54:** 905–921

Chou TY, Hart GW, Dang CV (1995) C-myc is glycosylated at threonine-58, a known phosphorylation site and a mutational hot-spot in lymphomas. *J Biol Chem* **270:** 18961–18965

Collaborative Computational Project, N (1994) The ccp4 suite: programs for protein crystallography. *Acta Crystallogr D* **50:** 760–763

Cowtan K (1994) *Jt CCP4 ESF-EACBM Newslett Prot Crystallogr* **31:** 34–38

Dong DLY, Hart GW (1994) Purification and characterization of an *o*-glcnac selective *n*-acetyl-β-D-glucosaminidase from rat spleen cytosol. *J Biol Chem* **269:** 19321–19330

Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D* **60:** 2126–2132

Ficko-Blean E, Boraston AB (2005) Cloning, recombinant production, crystallization and preliminary X-ray diffraction studies of a family 84 glycoside hydrolase from clostridium perfringens. *Acta Crystallogr F* **61:** 834–836

Gao Y, Wells L, Comer FI, Parker GJ, Hart GW (2001) Dynamic *o*-glycosylation of nuclear and cytosolic proteins—cloning and characterization of a neutral, cytosolic beta-*n*-acetylglucosaminidase from human brain. *J Biol Chem* **276:** 9838–9845

Haltiwanger RS, Grove K, Philipsberg GA (1998) Modulation of *o*-linked *n*-acetylglucosamine levels on nuclear and cytoplasmic proteins *in vivo* using the peptide *o*-glcnac-β-*n*-acetylglucosaminidase inhibitor *o*-(2-acetamido-2-deoxy-D-glucopyranosylidene)amino-*n*-phenylcarbamate. *J Biol Chem* **273:** 3611–3617

Heckel D, Comtesse N, Brass N, Blin N, Zang KD, Meese E (1998) Novel immunogenic antigen homologous to hyaluronidase in meningioma. *Hum Mol Genet* **7:** 1859–1872

Holm L, Sander C (1993) Protein structure comparison by alignment of distance matrices. *J Mol Biol* **233:** 123–138

Holt GD, Snow CM, Gerace L, Hart GW (1986) Localization of a novel form of glycosylation to the cytosolic faces of the nuclear–pore complex. *J Cell Biol* **103:** A320

Horsch M, Hoesch L, Vasella A, Rast DM (1991) *N*-acetylglucosaminono-1,5-lactone oxime and the corresponding (phenylcarbamoyl)oxime—novel and potent inhibitors of β-*n*-acetylglucosaminidase. *Eur J Biochem* **197:** 815–818

Jia ZC, Barford D, Flint AJ, Tonks NK (1995) Structural basis for phosphotyrosine peptide recognition by protein-tyrosine-phosphatase 1b. *Science* **268:** 1754–1758

Jones TA, Zou JY, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47:** 110–119

Konrad RJ, Mikolaenko I, Tolar JF, Liu K, Kudlow JE (2001) The potential mechanism of the diabetogenic action of streptozotocin: inhibition of pancreatic beta-cell *o*-glcnac-selective *n*-acetyl-beta-D-glucosaminidase. *Biochem J* **356:** 31–41

Lazarus BD, Roos MD, Hanover JA (2005) Mutational analysis of the catalytic domain of *o*-linked *n*-acetylglucosaminyl transferase. *J Biol Chem* **280:** 35537–35544

Leatherbarrow RJ (2001) *GraFit Version 5*. Horley, UK: Erithacus Software Ltd

Love DC, Hanover JA (2005) The hexosamine signaling pathway: deciphering the *O*-glcnac code. *Sci STKE* **312:** 1–14

Lubas WA, Hanover JA (2000) Functional expression of *o*-linked glcnac transferase—domain structure and substrate specificity. *J Biol Chem* **275:** 10983–10988

Macauley MS, Stubbs KA, Vocadlo DJ (2005a) *O*-GlcNAcase catalyzes cleavage of thioglycosides without general acid catalysis. *J Am Chem Soc* **127:** 17202–17203

Macauley MS, Whitworth GE, Debowski AW, Chin D, Vocadlo DJ (2005b) *O*-glcnacase uses substrate-assisted catalysis—kinetic analysis and development of highly selective mechanism-inspired inhibitors. *J Biol Chem* **280:** 25313–25322

Mark BL, Vocadlo DJ, Zhao DL, Knapp S, Withers SG, James MNG (2001) Biochemical and structural assessment of the 1-*n*-azasugar galnac-isofagomine as a potent family 20 beta-n acetylhexosaminidase inhibitor. *J Biol Chem* **276:** 42131–42137

Markovic-Housley Z, Miglierini G, Soldatova L, Rizkallah PJ, Muller U, Schirmer T (2000) Crystal structure of hyaluronidase, a major allergen of bee venom. *Structure* **8:** 1025–1035

Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* **53:** 240–255

Ohnishi Y, Ichikawa M, Ichikawa Y (2000) Facile synthesis of *n*-fmoc-serine-*s*-glcnac: a potential molecular probe for the functional study of *o*-glcnac. *Bioorg Med Chem Lett* **10:** 1289–1291

Park SY, Ryu JW, Lee W (2005) *O*-glcnac modification on irs-1 and akt2 by pugnac inhibits their phosphorylation and induces insulin resistance in rat primary adipocytes. *Exp Mol Med* **37:** 220–229

Parker GJ, Lund KC, Taylor RP, McClain DA (2003) Insulin resistance of glycogen synthase mediated by *o*-linked *n*-acetylglucosamine. *J Biol Chem* **278:** 10022–10027

Paulsen H, Holck JP (1982) Synthesis of the glycopeptide *o*-β-D-galactopyranosyl-(1-]3)-*o*-(2-acetamido-2-desoxy-α-D-galacto pyranosyl)-(1-]3)-L-serine and L-threonine. *Carbohydr Res* **109:** 89–107

Perrakis A, Morris R, Lamzin VS (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* **6:** 458–463

Perreira M, Kim Ju E, Thomas CJ, Hanover JA (2006) Inhibition of *O*-GlcNAcase by PUGNAc is dependent upon the oxime stereochemistry. *Bioorg Med Chem* **14:** 837–846

Salmeen A, Andersen JN, Myers MP, Tonks NK, Barford D (2000) Molecular basis for the dephosphorylation of the activation segment of the insulin receptor by protein tyrosine phosphatase 1b. *Mol Cell* **6:** 1401–1412

Salmeen A, Andersen JN, Myers MP, Meng TC, Hinks JA, Tonks NK, Barford D (2003) Redox regulation of protein tyrosine phosphatase 1b involves a sulphenyl-amide intermediate. *Nature* **423:** 769–773

Schuettelkopf AW, van Aalten DMF (2004) Prodrg: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D* **60:** 1355–1363

Schultz J, Pils B (2002) Prediction of structure and functional residues for *o*-glcnacase, a divergent homologue of acetyltransferases. *FEBS Lett* **529:** 179–182

Slawson C, Zachara NE, Vosseller K, Cheung WD, Lane MD, Hart GW (2005) Perturbations in *o*-linked beta-*n*-acetylglucosamine protein modification cause severe defects in mitotic progression and cytokinesis. *J Biol Chem* **280:** 32944–32956

Terwilliger TC, Berendzen J (1999) Automated mad and mir structure solution. *Acta Crystallogr D* **55:** 849–861

Terwisscha van Scheltinga AC, Kalk KH, Beintema JJ, Dijkstra BW (1994) Crystal-structures of hevamine, a plant defense protein with chitinase and lysozyme activity, and its complex with an inhibitor. *Structure* **2:** 1181–1189

Tews I, Terwisscha van Scheltinga AC, Perrakis A, Wilson KS, Dijkstra BW (1997) Substrate-assisted catalysis unifies two families of chitinolytic enzymes. *J Am Chem Soc* **119:** 7954–7959

Toleman C, Paterson AJ, Whisenhunt TR, Kudlow JE (2004) Characterization of the histone acetyltransferase (hat) domain of a bifunctional protein with activable *o*-glcnacase and hat activities. *J Biol Chem* **279:** 53665–53673

Torres CR, Hart GW (1984) Topography and polypeptide distribution of terminal *n*-acetylglucosamine residues on the surfaces of intact lymphocytes—evidence for *o*-linked glcnac. *J Biol Chem* **259:** 3308–3317

Vaaje-Kolstad G, Vasella A, Peter MG, Netter C, Houston DR, Westereng B, Synstad B, Eijsink VGH, van Aalten DMF (2004) Interactions of a family 18 chitinase with the designed inhibitor hm508, and its degradation product, chitobiono-δ-lactone. *J Biol Chem* **279:** 3612–3619

van Aalten DMF, Komander D, Synstad B, Gåseidnes S, Peter MG, Eijsink VGH (2001) Structural insights into the catalytic mechanism of a family 18 exo-chitinase. *Proc Natl Acad Sci USA* **98:** 8979–8984

Varrot A, Frandsen TP, von Ossowski I, Boyer V, Cottaz S, Driguez H, Schulein M, Davies GJ (2003) Structural basis for ligand binding and processivity in cellobiohydrolase Cel6A from *Humicola insolens*. *Structure* **11:** 855–864

Vosseller K, Wells L, Lane MD, Hart GW (2002) Elevated nucleocytoplasmic glycosylation by *o*-glcnac results in insulin resistance associated with defects in akt activation in 3t3-l1 adipocytes. *Proc Natl Acad Sci USA* **99:** 5313–5318

Wells L, Gao Y, Mahoney JA, Vosseller K, Chen C, Rosen A, Hart GW (2002) Dynamic *o*-glycosylation of nuclear and cytosolic proteins—further characterization of the nucleocytoplasmic beta-*n*-acetylglucosaminidase, *o*-glcnacase. *J Biol Chem* **277:** 1755–1761

Williams SJ, Mark BL, Vocadlo DJ, James MNG, Withers SG (2002) Aspartate 313 in the *Streptomyces plicatus* hexosaminidase plays a critical role in substrate-assisted catalysis by orienting the 2-acetamido group and stabilizing the transition state. *J Biol Chem* **277:** 40055–40065

Yang XY, Zhang FX, Kudlow JE (2002) Recruitment of *o*-glcnac transferase to promoters by corepressor msin3a: coupling protein *o*-glcnacylation to transcriptional repression. *Cell* **110:** 69–80

Zachara NE, Hart GW (2004) *O*-glcnac a sensor of cellular state: the role of nucleocytoplasmic glycosylation in modulating cellular function in response to nutrition and stress. *Biochim Biophys Acta* **1673:** 13–28

Zhu XM, Haag T, Schmidt RR (2004) Synthesis of an *S*-linked glycopeptide analog derived from human Tamm–Horsfall glycoprotein. *Org Biomol Chem* **2:** 31–33

