Structural and Kinetic Analysis of Bacillus subtilis N-Acetylglucosaminidase Reveals a Unique Asp-His Dyad Mechanism*§

Received for publication, April 6, 2010, and in revised form, September 3, 2010. Published, JBC Papers in Press, September 7, 2010. Published, JBC Papers in Press, September 7, 2010, DOI 10.1074/jbc.M110.131037

Silke Litzinger1,2, Stefanie Fischer5,1, Patrick Polzer5, Kay Diederichs5, Wolfram Welte4, and Christoph Mayer1,3

From the Departments of 1,5 Molecular Microbiology and 4 Biophysics, Fachbereich Biologie, University of Konstanz, 78457 Konstanz, Germany and the 3 Max-Planck-Institute of Quantum Optics, 85748 Garching, Germany

Three-dimensional structures of NagZ of Bacillus subtilis, the first structures of a two-domain β-N-acetylglucosaminidase of family 3 of glycosidases, were determined with and without the transition state mimicking inhibitor PUGNac bound to the active site, at 1.84- and 1.40-Å resolution, respectively. The structures together with kinetic analyses of mutants revealed an Asp-His dyad involved in catalysis: His234 of BsNagZ acts as general acid/base catalyst and is hydrogen bonded by Asp232 for proper function. Replacement of both His234 and Asp232 with glycine reduced the rate of hydrolysis of the fluorogenic substrate 4′-methylumbelliferyl N-acetyl-β-D-glucosaminide 1900- and 4500-fold, respectively, and rendered activity pH-independent in the alkaline range consistent with a role of these residues in acid/base catalysis. N-Acetylglucosaminyl enzyme intermediate accumulated in the H234G mutant and β-azide product was formed in the presence of sodium azide in both mutants. The Asp-His dyad is conserved within β-N-acetylglucosaminidases but otherwise absent in β-glycosidases of family 3, which instead carry a “classical” glutamate acid/base catalyst. The acid/base glutamate of Hordeum vulgare exoglucanase (Exo1) superimposes with His234 of the dyad of BsNagZ and, in contrast to the latter, protrudes from a second domain of the enzyme into the active site. This is the first report of an Asp-His catalytic dyad involved in hydrolysis of glycosides resembling in function the Asp-His-Ser triad of serine proteases. Our findings will facilitate the development of mechanism-based inhibitors that selectively target family 3 β-N-acetylglucosaminidases, which are involved in bacterial cell wall turnover, spore germination, and induction of β-lactamase.

Glycosidases that catalyze the hydrolysis of glycosidic linkages under retention of anomeric configuration of substrate and product (retaining glycosidases) proceed via a two-step double displacement mechanism (1, 2). In most cases, the catalytic machinery of these enzymes involves two carboxylic acids, which are located ~5.5 Å apart in the active site (Fig. 1A) and function as catalytic nucleophile and catalytic acid/base (also general acid/base catalyst). In the first step (glycosylation), a carboxylic acid that hydrogen bonds to the glycosidic oxygen acts as a general acid facilitating the leaving group departure simultaneously with a nucleophilic attack by the second carboxylate that forms a covalent glycosyl enzyme intermediate. In the second step (deglycosylation), the first residue then functions as a general base to activate an incoming water molecule for nucleophilic attack that hydrolyzes the glycosyl enzyme to form a sugar hemiacetal product with overall retention of stereochemistry (1, 2).

This mechanism, for instance, is performed by the well studied lysozyme that hydrolyzes the β-glicosidic linkage between N-acetylmuramic acid (MurNAc) and N-acetylgalactosamine (GlcNAc) of the backbone polysaccharide of the bacterial cell wall component peptidoglycan (murein). It was shown only recently that lysozyme proceeds through a covalent α-glycosyl enzyme (3) and not a long-lived oxocarbenium-ion intermediate as was proposed earlier (4). We are studying a group of bacterial β-N-acetylglucosaminidases, which hydrolyze the other glycosidic linkage in peptidoglycan, between GlcNAc and MurNAc, and are involved in turnover and recycling of the bacterial cell wall (5–10). These enzymes are classified on the basis of amino acid sequence and secondary structure to family 3 of glycosidases (according to the carbohydrate active enzymes (CAZY) data base), which comprises a heterogeneous group of exo-acting, retaining β-glycosidases that besides β-N-acetylglucosaminidases (EC 3.2.1.52), include β-glucosidases (EC 3.2.1.21), xylan 1,4-β-xyllosidases (EC 3.2.1.37), glucan 1,3-1,4-β-glucosidases (EC 3.2.1.58 and 3.2.1.74), α-L-arabinofuranosidases (EC 3.2.1.55), and exo-1,3,1,4-glucanases (EC 3.2.1.-).

Family 3 β-N-acetylglucosaminidase from Vibrio furnissii (VfExoII) as well as related β-glucosidases were shown to proceed through a glycosyl enzyme intermediate. A conserved aspartate residue was identified in all cases as the catalytic

* This work was supported in part by Grant MA2436/4 and a Heisenberg stipend (MA2436/3) from the Deutsche Forschungsgemeinschaft (to C. M.).
§ The on-line version of this article (available at http://www.jbc.org) contains supplemental Table S1 and Figs. S1–S4.
1. Both authors contributed equally to this work.
2. Supported by a fellowship of the federal government of Germany (Landesgraduiertenförderungsgesetz stipend).
3. To whom correspondence should be addressed: Universitätsstr. 10, 78457 Konstanz, Germany. E-mail: ch.mayer@uni-konstanz.de.

This is an Open Access article under the CC BY license.
Asp-His Dyad Glycosidase

A. Acid/base catalyst

B. Catalytic nucleophile

C. Glycosyl-enzyme intermediate

D. Catalytic dyad

E. Catalytic nucleophile

F. Glycosyl-enzyme intermediate
Asp-His Dyad Glycosidase

FIGURE 1. Mechanism and structure of β-exo-glucanase I of H. vulgare (HvExol) and β-N-acetylglucosaminidase of B. subtilis (BsNagZ) of family 3 of glycosidases. A, schematic of the general two-step double displacement mechanism of retaining β-glycosidases as proposed also for HvExol. Two catalytic carboxyl groups, which are located ~5.5 Å apart in the active site, act as the catalytic nucleophile and general acid/base catalyst, respectively. B, ribbon model of HvExol (PDB identifier 1X38). View of the top of the catalytic N-terminal (β/αI)-barrel domain (magenta) to which the inhibitor glucopyranosidazole is bound in the active site (gray sticks; for chemical structures see supplemental Fig. S1), which carries the catalytic nucleophile (orange stick, right). The amino acid sequence of Exol shows 22% identity with BsNagZ. D, schematic of the modified double displacement mechanism as proposed for BsNagZ. Similar to HvExol, an asparagine residue acts as the catalytic nucleophile (Asp<sup>318</sup>), however, instead of an acid/base glutamate residue an Asp-His dyad catalytic dyad was now identified as the acid/base catalyst of β-N-acetylglucosaminidases of family 3. The catalytic groups are located ~6.3 Å apart in the active site. E, ribbon model of BsNagZ (PDB identifier 3NVD). View on the top of the catalytic N-terminal (β/αI)-barrel domain (magenta) to which the inhibitor PUGNAc is bound (gray sticks; for chemical structures see supplemental Fig. S1) in the active site that carries the catalytic nucleophile (orange stick, left). In contrast to HvExol, the C-terminal domain of BsNagZ, shown in green, is further apart from the TIM-barrel and does not contribute to the active site. F, in BsNagZ an Asp-His dyad mediates the acid/base function that superimposes with the acid/base glutamate of HvExol (cf. C). Residues Asp<sup>332</sup> and His<sup>334</sup> (shown in orange) are in H-bond distance to each other and His<sup>334</sup>, acid/base catalyst, as well as Asp<sup>316</sup>, the catalytic nucleophile (shown in orange), are H-bonding the PUGNAC inhibitor. 

1 domain β-N-acetylglucosaminidases

P75949 NAG3_ECOLI 161 TKHGFCGAVTABSKHE
Q85Q06 NAG5_SALTY 161 TKHGFCGAVTABSKHE
P96157 NAG5_VIBFU 158 TKHGFCGAVIASSHLE
Q9KU37 NAG8_VIBCH 158 TKHGFCGAVIASSHLE
P44955 NAG5_HAIN 163 TKHGFCGAVLASSHLE
Q9HZK0 NAG3_PSEA 159 TKHGFCGAEASSHVA
Secondary structure ->5->

2 domain β-N-acetylglucosaminidases

P40406 NAG3_BACSU 219 ALKFGCHGTVDSHLYG 509 Y1ITGSYVKNPVPVNDGV
Q68823 HEXA_PSEO7 206 ALKFGCHGTVSHGY 481 MI2AASPPQSAAEIGGM
Q82840 NAG3_STRTL 261 TKHGFCGAVTAATGC 530 AVWATNYTASAQTLTV
Q9R099 NAG4_STRCC 265 TKHGFCGAVTATGC 514 AVWATNYTASAQTLTV
Q9K81U NAG5_STRCC 241 TKHGFCGATQSHGY 510 AVWATNYTASAQTLTV
Q8XNR6 HEXA_CLOPE 224 TKHGFCGTVSSHGY 509 YIVASLSNSNLKGFAW
Q97ML4 NAG3_CLOB 178 VAHGFGCTVDTHLNL 440 SIVIGIYNAHIEGKL
Secondary structure ->5-> ->IV-> => =>

β-glucosidases

Q9E532 EXO1_HORVD 229 CAKEVLCPTVTDGIVEN 480 AIVAYEHPYTPKGDNLN
Q42835 EXO2_HORVD 229 CAKEVLCPTVTDGIVEN 490 AVVYVEPPTFTMGDPNLN
P33363 BGLX_ECOLI 206 SWKHPAYVAGKKEYN 505 WWVVYRAQGMHRASSRT
P56078 BGLX_SALTY 206 SWKHPAYVAGKKEYN 505 WWVVVSEQGMHRASSRT
Q87013 BGLX_FLAME 166 CWKALYLVAFGDRYN 460 VWLAFPTLBSGGSARA
Q8A79 GLUC_PAESP 141 SLIZFEVNNQHEHRMTT 398 AIVLFLDPRYESGGDTRT
P27034 BGLX_AGRTU 171 TKFEVAN-ESEEIRQT 546 VLLVRRGEGWDTHLDP
Secondary structure ->5-> / / / / / /

Secondary structure

FIGURE 2. Partial multiple sequence alignment of selected glycosidases of family 3, including one domain and two domain β-N-acetylglucosaminidases as well as two domain β-glucosidases. The subfamilies can be distinguished by differences in the sequence pattern next to the conserved KH(F/Y) domain (boxed in gray), which carries the catalytic nucleophile (Asp<sup>318</sup>) on the N-terminal domain adjacent to the β-N-acetylglucosaminidases subfamily can be distinguished by differences in the sequence pattern next to the conserved KH(F/Y) domain (boxed in black). Figure 2 shows the alignment of a two-domain β-N-acetylglucosaminidase of V. cholerae (NagZ), the first structure of a two-domain β-N-acetylglucosaminidase, along with kinetic analyses, which provide evidence for participation of the side chains of the conserved Asp and His residues during catalysis. Our results indicate that the histidine, instead of a glutamate, acts as acid/base catalyst, which undergoes hydrogen bonding with the aspartate residue, thereby forming a catalytic dyad that protonates the glycosidic oxygen in the first (glycosylation) step and deprotonates and activates water for nucleophilic attack of the glycosyl enzyme in the second (deglycosylation) step of the reaction (Fig. 1D). The function of this unique Asp-His dyad in glycoside hydrolysis resembles that of the Asp-His-Ser triad of serine proteases (19) as well as the Asp-His dyad of ribonucleases (20).
Asp-His Dyad Glycosidase

EXPERIMENTAL PROCEDURES

Bacterial Strains and Plasmids—B. subtilis 168 was obtained from the Bacillus Genetic Stock Center. Escherichia coli BL21(DE3) and E. coli expression vectors pET16b (P7, AmpR, ori<sub>pBR322</sub>, lacI, N-terminal His<sub>6</sub> tag) and pET29b (P7, Kan<sub>4</sub>, ori<sub>pBR322</sub>, lacI, C-terminal His<sub>6</sub> tag) were from Novagen. The reagents 4’-methylumbelliferyl N-acetyl-β-d-glucosaminide (4-Mu-β-GlcNAc) and 4’-nitrophenyl N-acetyl-β-d-glucosaminide (pNP-β-GlcNAc) were obtained from Glycosynth. N-Acetyl-β-d-glucosaminyl azide (β-GlcNAc azide) was from Carbosynth. All other reagents were from Sigma unless otherwise stated.

Cloning and Site-directed Mutagenesis—B<sub>s</sub> nagZ was cloned as an cytoplasmic construct in which the N-terminal signal sequence was removed and exchanged by a His tag. DNA preparation, restriction enzyme digestion, ligation, and transformation were performed according to standard techniques using 5 units of PWO DNA polymerase (Genaxxon Biosciences, Biberach, Germany) and 30 ng of chromosomal DNA. B<sub>s</sub> nagZ from B. subtilis 168 was cloned into pET16b without signal sequence (pET16b-<i>Bs</i>nagZ) using the following primers: <i>Bs</i>nagZ/FP, 5’-AAA ACC ATG GGC CAT ATG TTT TTC GAG TTA-3’ and <i>Bs</i>nagZ/RP, 5’-T TT TCT CTC GAG TTA AAG CGG TCT TCC CGT TTT G-3’ (underlined are the recognition sites for the endonucleases NdeI and XhoI, respectively). Thirty cycles (15 s at 94 °C, 30 s at 55 °C, and 120 s at 72 °C) were performed in a thermal cycler and revealed a single 1.9-kb fragment (underlined are the recognition sites for the endonucleases NdeI and XhoI, respectively). Thirty cycles (15 s at 94 °C, 30 s at 55 °C, and 120 s at 72 °C) were performed in a thermal cycler and revealed a single 1.9-kb fragment (<i>Bs</i>nagZ). <i>Bs</i>nagZ’ missing the N-terminal signal sequence and the C-terminal domain was amplified using primers with recognition sites for restriction endonucleases NdeI and XhoI, respectively (underlined): <i>Bs</i>nagZ’/FP, 5’-AAA ACC ATG GGC CAT ATG TTT TTC GGG GCC AGA CAG AC-3’; and <i>Bs</i>nagZ’/RP, 5’-T TT TCT CTC GAG TTA AAG CGG TCT TCC GTT TTT G-3’ (underlined are the recognition sites for the endonucleases NdeI and XhoI, respectively). Thirty cycles (15 s at 94 °C, 30 s at 55 °C, and 120 s at 72 °C) were performed in a thermal cycler and revealed a single 1.9-kb fragment (Bs<nagZ’>). The vector pET16b-B<sub>s</sub>nagZ’ was used as template for site-directed mutagenesis of H234G. The following degenerated primers (Thermo Fisher) were used: H234G/FP as well as D323G/FP, 5’-GG AGA TAT ACC ATG GGC CAT C-3’; H234G/RP, 5’-CC TTG GCC ATG GGA AAC GAG CCG CAG TCC AT AYM AYM GCT GCT AAC GTC GTG TCT CTC-3’; and D323G/RP, 5’-C TTG GCC ATG GGA AAC GAG CCG CAG TCC AT AYM GCT GCT AAC GTC GTG TCT ATG TC-3’ (the mutated codon is shown in bold with Y = C or T, M = A or C, and B = C, G, or T; the underlined residue is the restriction site for Ncol). Thirty cycles (30 s at 94 °C, 30 s at 50 °C, and 60 s at 72 °C) were performed. The amplified 785-bp fragments carrying the mutation were cleaved with Ncol and exchanged with the wild type fragment of pET16b-<i>Bs</i>nagZ. Two mutants were further processed that give rise to H234G (AYM = ACC) and D323G (CBC = CCC) protein variants.

Overproduction and Purification of Proteins—E. coli strain BL21(DE3) (F<sup>ompT</sup> hisD<sub>λ</sub> gal dcm) harboring pET16b-<i>Bs</i>nagZ, pET16b-<i>H</i>234G, pET16b-D323G, or pET29b-<i>Bs</i>nagZ’ were grown at 37 °C in Luria-Bertani (LB) medium supplemented with ampicillin (100 µg/ml) or kanamycin (50 µg/ml), respectively, under vigorous shaking to logarithmic phase (<i>A<sub>600</sub></i> 0.5–0.6). Production of the enzyme was induced by the addition of isopropyl β-d-thiogalactopyranoside to a final concentration of 1 mM. Incubation was continued for a further 3 h and the cells were harvested by centrifugation (4000 × g for 30 min at 4 °C), resuspended in ice-cold phosphate buffer (20 mM Na<sub>2</sub>HPO<sub>4</sub> × 2H<sub>2</sub>O, 500 mM NaCl, pH 7.5), and lysed by passing the cells three times through a French pressure cell. The lysates were clarified by centrifugation at 100,000 × g for 1 h at 4 °C. The His-tagged proteins in the supernatant were purified by Ni<sup>2+</sup>-affinity chromatography on a 5-ml HiTrap column (Amersham Biosciences) pre-equilibrated with phosphate buffer and eluted from the column with imidazole (phosphate buffer supplemented with 500 mM imidazole, pH 7.5). To avoid cross-contamination of the proteins, new columns were used for each protein. The purity of the enzymes was assessed by SDS-PAGE. Fractions containing homogeneous protein were pooled, concentrated, and desalted by dialysis against 10 mM sodium acetate, pH 4.5, for crystallization or against 20 mM phosphate buffer, pH 7.5, for enzymatic assay studies at 4 °C. The protein concentrations were measured according to the method of Bradford with bovine serum albumin as a standard. 3 to 5 mg of pure protein was obtained from a 1-liter culture.

Crystallization and Data Collection—Purified B<sub>s</sub>nagZ was concentrated to 14 mg/ml using an Amicon Ultra Centrificon (Millipore) with a 10-kDa cut off. Complexes of B<sub>s</sub>nagZ with the inhibitor for co-crystallization were prepared by mixing PUGNAc with B<sub>s</sub>nagZ solution at a 10:1 molar ratio and incubating for 30 min at room temperature. Crystals of unliganded B<sub>s</sub>nagZ and its co-crystals with inhibitor PUGNAc were grown in 0.1 M sodium acetate, pH 4.9, using 24-well hanging drop plates. Presaturated protein drops were prepared by mixing protein and reservoir solution at ratios 1:1, 1:2, and 2:1, yielding final drop volumes of 2–3 µl. The reservoir contained 0.1 M sodium acetate, pH 4.9, and varying concentrations of polyethylene glycol 1000. Crystals were shock-frozen in liquid nitrogen. Data sets were collected at the Swiss Light Source beamline X06SA of the Paul Scherrer Institut, Villigen, Switzerland. Data processing was done with XDS (21) (see Table 1).

Structure Determination—The structure of B<sub>s</sub>nagZ without ligand was solved first. An ensemble of PDB files (1EXI, 1IEX, 1J8V, 1LQ2, and 1X38 (14–16)) was defined in Phaser (22) as the starting point for molecular replacement. Overall amino acid sequence identities between the target enzyme and the related search models was low (23% and less) (23). Buccaneer (24) was used to build parts of two molecules of B<sub>s</sub>nagZ (996 out of 1284 residues as poly-Ala model). ARP/wARP (25) and Resolve (26) were used for completing the model. By using Resolve, 2-fold non-crystallographic symmetry (NCS) could be determined and used for phase improvement, leading finally to an almost complete model with 1139 amino acids of the B<sub>s</sub>nagZ sequence. The resulting electron density map was used for further manual model building in COOT (27) and refinement with Refmac5 and Phenix (28). After modeling water molecules and adding hydrogen atoms, the last anisotropic refinement resulted in a final R-factor of 12.7% (R<sub>free</sub> 16.6%). The structure of B<sub>s</sub>nagZ in complex with its inhibitor PUGNAc was
determined by using the first structure of BsNagZ as the search model in Molrep (29). The pseudo-merohedral twinning law was identified with the help of the CCP4 (30) program SFCHECK (31). Model refinement was done in COOT. After adding waters and modeling the inhibitor PUGNAc, the last isotropic refinement resulted in a final R-factor of 18.7% (R_{free}, 24.6%). Refinement statistics are presented in Table 1. Atomic coordinates and structure factors for unliganded and liganded N-acetylglucosaminidases of *B. subtilis* have been deposited in the Protein Data Bank under accession code 3BMX and 3NVD, respectively.

Kinetic Studies—All Michaelis-Menten kinetics were carried out at least in triplicates using a discontinuous assay measuring the release of 4-methylumbelliferone from 4-Mu β-GlcNAc. The fluorescence of the released 4-methylumbelliferone was measured using a 96-well plate (Greiner bio-one) in a SpectraMax M2 Microplate Reader (Molecular Devices) (excitation, 362 nm; emission, 448 nm) at 37 °C. Enzyme activity was measured using various concentrations of 4-Mu β-GlcNAc in Clark and Lubs solution (0.1 M KH₂PO₄, 0.1 M NaOH) at pH 5.8. Reactions (300 μl) were initiated by addition of enzyme (0.0158 mg of His, 0.01465 mg of Asp, and 1.5 × 10⁻⁵ mg of BsNagZ) and incubated over a period of 5 to 30 min at 37 °C.

The pH activity profiles of wild type *BsNagZ* and the mutants were determined in 0.1 M citric acid, 0.2 M disodium phosphate buffer (McIlvaine) ranging from 4.0 to 8.0, in 0.2 M sodium acetate/acetic acid buffer ranging from 4.0 to 5.6 and in Clark and Lubs solution in the range of 5.8–8.0. Reactions (300 μl) were initiated by addition of 1.5 mM 4-Mu β-GlcNAc and incubated over a period of 5–30 min at 37 °C. Protein amounts used were 4.5 × 10⁻⁵, 0.0114, and 0.01465 mg for *BsNagZ*, H234G, and D232G, respectively. All reactions were stopped by addition of 270 μl of 0.2 M Na₂CO₃, pH 10.0, to 30 μl of the reaction mixture. Kinetic parameters were obtained by direct fit of the rate versus substrate concentration data to the Michaelis-Menten equation using Prism 4 software (Graph Pad Software, Inc., La Jolla, CA). The extinction coefficient was determined by calibration using 4'-methylumbelliferone. One unit is defined as the amount of enzyme that hydrolyzes 1 μmol of substrate per min at pH 5.8 at 37 °C.

Mass Spectrometry—H234G was incubated with pNP β-GlcNAc azide over a period of 10 min until the enzyme/substrate solution exhibited a light yellow color indicating the release of pNP. The mass spectrometry analysis of the trapped intermediate (glycosyl-enzyme) was performed by electrospray ionization time of flight mass spectrometry (ESI-TOF-MS) (at the ZMBH, University of Heidelberg).

HPLC Analysis—High performance liquid chromatography (HPLC) was used for analysis of the β-GlcNAc azide generated by reaction of *BsNagZ*-H234G and *BsNagZ*-D232G with pNP β-GlcNAc (6 mM) in 100 mM sodium phosphate, pH 7.5, and 500 mM sodium azide, pH 7.5. In brief a reversed-phase column (Gemini RP-C18, 150 × 4.6 mm, 5 μm particle size, Phenomenex) was used at a flow rate of 0.5 ml/min, isocratic elution with 0.05% trifluoroacetic acid for 5 min, followed by a gradient from 0.05% trifluoroacetic acid to 10% acetonitrile containing 0.035% trifluoroacetic acid over a period of 50 min according to Ref. 32.

RESULTS

Purification, Crystallization, and Structure Determination—*BsNagZ* was overproduced in *E. coli* cytoplasm and purified by Ni²⁺-affinity chromatography to apparent homogeneity (23). The protein retained activity for several months at 4 °C and in the pH range between pH 4.0 and 8.0. *BsNagZ* tends to precipitate at a pH above 6 and low ionic strength but can be kept soluble at 12–13 mg/ml and high ionic strength (e.g. 20 mM sodium phosphate, pH 7.5, 500 mM sodium chloride). At acidic pH solubility behavior is reversed but the protein shows only low enzymatic activity; *BsNagZ* is soluble at low ionic strength (e.g. 20 mM sodium acetate, pH 4.5) but precipitates upon addition of 100 mM sodium chloride. Crystals of unliganded *BsNagZ* and its complex with the transition state-like inhibitor PUGNAc, O-(2-acetamido-2-deoxy-D-glucopyranosylidene)amino-N-phenylcarbamate (chemical structures of inhibitors and substrates of *BsNagZ* are shown in supplemental Fig. S1) were grown at pH 4.9 and 18 °C by vapor-phase equilibration as described under “Experimental Procedures.” The best results were obtained at polyethylene glycol 1000 reservoir concentrations between 27 and 33% (w/v). Within 2–3 weeks crystals reached a size of approximately 50 μm² and exhibited triangular, rhombic, or cuboid habits. Data were collected at the Swiss Light Source synchrotron to resolutions of 1.4-Å and 1.84-Å for the native (unliganded) and inhibitor complex (liganded) protein crystals, respectively. All crystals belonged to space group P1 and possessed very similar unit cells with two *BsNagZ* molecules per cell (cf. Table 1). The inhibitor complex crystals displayed pseudo-merohedral twinning with a twinning fraction of 13.7%. The structures of native *BsNagZ* and its inhibitor complex were solved by molecular replacement as described under “Experimental Procedures.”

*BsNagZ* Has a Unique Two-domain Structure—The final atomic model comprises residues 26–642 (native enzyme numbering, corresponding to residues 31–647 of the engineered *BsNagZ*, lacking the N-terminal His tag) from both monomers of the asymmetric unit (structural parameters and refinement statistics, see Table 1). *BsNagZ* structure reveals two separate domains (Fig. 1E). The N-terminal domain (residues 26–420) adopts a (β/α)₁₀-barrel fold (TIM-barrel), which is typical for catalytic domains of glycosidases and contain the conserved aspartate, the catalytic nucleophile of family 3 glycosidases. The C-terminal domain (residues 421–642) displays an αβα-sandwich fold. The surface buried between the domains (1407.5 Å²) indicates that both domains are intimately associated (33), however, no part of the C-terminal domain comes into close contact with the substrate/inhibitor binding site on top of the (β/α)₁₀-barrel domain (Fig. 1, E and F). This distinguishes *BsNagZ* from *HvExoI* although both display weak sequence similarity (22% overall amino acid sequence identity) (15). As mentioned above, in *HvExoI* a short helix of the C-terminal domain, which carries the general acid/base catalyst (Glu⁴⁹⁴), protrudes into the active site (Fig. 1, B and C). This helix is missing in *BsNagZ* (Fig. 1, E and F).
**Asp-His Dyad Glycosidase**

### TABLE 1
Crystallographic data collection and refinement statistics

<table>
<thead>
<tr>
<th>Parameters</th>
<th>BsNagZ</th>
<th>BsNagZ-PUGNAc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P1</td>
<td>P1</td>
</tr>
<tr>
<td>Solvent content (%)</td>
<td>48.1</td>
<td>47.2</td>
</tr>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unit cell dimensions (Å)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>85.8</td>
<td>85.8</td>
</tr>
<tr>
<td>b</td>
<td>73.1</td>
<td>73.2</td>
</tr>
<tr>
<td>c</td>
<td>83.6</td>
<td>83.8</td>
</tr>
<tr>
<td>Unit cell dimensions (Å)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>79.8</td>
<td>69.6</td>
</tr>
<tr>
<td>b</td>
<td>69.6</td>
<td>69.4</td>
</tr>
<tr>
<td>c</td>
<td>83.6</td>
<td>83.8</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>83.8</td>
<td>83.2</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>48.18–1.40</td>
<td>48.18–1.40</td>
</tr>
<tr>
<td>Total observations</td>
<td>1,601,297 (66995)</td>
<td>207,544 (32774)</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>228,046 (25215)</td>
<td>104,555 (16575)</td>
</tr>
<tr>
<td>Flack parameter</td>
<td>–0.099</td>
<td>0.0979</td>
</tr>
<tr>
<td>Flack R factor</td>
<td>0.261</td>
<td>0.262</td>
</tr>
<tr>
<td>Flack R value</td>
<td>0.260</td>
<td>0.261</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>90.8</td>
<td>93.5</td>
</tr>
<tr>
<td>R_onset (%)</td>
<td>9.0</td>
<td>10.9</td>
</tr>
<tr>
<td>R_misfit (%)</td>
<td>8.0</td>
<td>7.6</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>42.84–1.40 (1.42–1.40)</td>
<td>48.18–1.40 (1.47–1.48)</td>
</tr>
<tr>
<td>R_work (%)</td>
<td>12.7</td>
<td>18.7</td>
</tr>
<tr>
<td>R_res (%)</td>
<td>16.6</td>
<td>24.6</td>
</tr>
<tr>
<td>Model composition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein residues</td>
<td>1,234</td>
<td>1,234</td>
</tr>
<tr>
<td>Ligand atoms</td>
<td>103</td>
<td>103</td>
</tr>
<tr>
<td>Water molecules</td>
<td>1,288</td>
<td>1,527</td>
</tr>
<tr>
<td>Average B (Å²)</td>
<td>20.9</td>
<td>23.5</td>
</tr>
<tr>
<td>Root mean square deviations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond length (Å)</td>
<td>0.007</td>
<td>0.017</td>
</tr>
<tr>
<td>Bond angles (Å)</td>
<td>0.967</td>
<td>1.611</td>
</tr>
<tr>
<td>Ramachandran plot</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Favored regions (%)</td>
<td>97.8</td>
<td>96.8</td>
</tr>
<tr>
<td>Allowed regions (%)</td>
<td>99.8</td>
<td>99.8</td>
</tr>
<tr>
<td>Twin law (Å)</td>
<td>0.173</td>
<td>0.173</td>
</tr>
<tr>
<td>Twin fraction (Å)</td>
<td>0.137</td>
<td>0.137</td>
</tr>
</tbody>
</table>

*Values in parentheses refer to the high resolution shell.

**An Asp-His Dyad in the Inhibitor Binding Site**—The overall structures of BsNagZ with and without the inhibitor (liganded and unliganded structures) are basically identical, including the inhibitor binding site arrangement. An acetate molecule in the unliganded structures (a fairly good substrate; pKₐ of methylumbellifereone = 7.79 compared with the natural muropeptide substrates with an estimated pKₐ of about 14). The apparent second-order rate constant (k_cat/K_m) for substrate hydrolysis of 4-Mu β-GlcNAc by H234G thus was much less affected than the apparent first-order rate constant (k_cat), indicating less impairment of the glycosylation step (reflected in k_cat/K_m) than the deglycosylation step (reflected in k_cat), indicating a particularly important role of His²⁵⁴ in base catalysis. This leads to a significant accumulation of the covalent glycosyl-enzyme intermediate (cf. Fig. 4), which is reflected in a 24-fold reduced Michaelis-Menten constant (K_m) for H234G compared with the wild type enzyme.

The rate of 4-Mu β-GlcNAc hydrolysis by the D232G mutant was 4500-fold reduced compared with wild type, which is an even larger reduction compared with the effect of the H234G mutation (Table 2). However, the K_m was only a little affected by the D232G mutation (Table 2). This indicates a shift in the rate-determining step in the D232G mutant compared with H234G from deglycosylation in the latter to glycosylation in the former. This might be explained by a larger impairment of the protonation of the leaving group of the substrate (glycosylation step) in D232G compared with H234G, presumably because Asp²³² mostly is required for His²⁵⁴ to function as a proton donor. Protonation of the glycosidic oxygen in catalysis by the H234G mutant might be substituted by small organic acids of the buffer (phosphate, acetate), which, however, cannot substitute general base catalysis (deglycosylation step). Congruently, the activity of the H234G mutant of BsNagZ in the Tris buffer was lower than in phosphate or acetate buffer but to low to determine kinetic parameters.

**pH Dependence of Catalysis**—The pH activity profile of wild type BsNagZ was compared with H234G and D232G mutants. The plots of log k_cat as a function of pH are shown in Fig. 5 and supplemental Fig. S3. The pH activity profile for the BsNagZ-catalyzed hydrolysis of 4-Mu β-GlcNAc resembles a bell-shaped curve as expected due to involvement of two ionizable  β-GlcNAc and 4-Mu β-GlcNAc, respectively, which are convenient substrates for continuous assays and kinetic studies. Kinetics were determined with 4-Mu β-GlcNAc, because this substrate generates a fluorogenic product that can be measured with much higher sensitivity than chromogenic products. Hydrolysis of 4-Mu β-GlcNAc by wild type BsNagZ and mutants obey Michaelis-Menten kinetics and the kinetic constants are given in Table 2. To investigate the role of the Asp-His dyad in family 3 β-N-acetylglucosaminidases, His²³⁴ and Asp²³² of BsNagZ were exchanged by a glycine. Both mutants showed severely reduced activity. As shown in previous studies with β-glycosidases, glycosylation (the first irreversible step that is reflected through k_cat/K_m) requires major assistance in protonation of the glycosidic oxygen by the general acid/base catalyst for cleavage of the substrates that have poor leaving groups. By contrast, the second step (deglycosylation, reflected through k_cat) depends on the general acid/base catalyst functioning as base at this stage independent of the leaving group of the substrate (34). Exchanging His²³⁴ with glycine resulted in a 1900-fold reduction in k_cat but only a 80-fold reduction in k_cat/K_m compared with wild type BsNagZ and with 4-Mu β-GlcNAc as substrate (a fairly good substrate; pKₐ of methylumbellifereone = 7.79 compared with the natural muropeptide substrates with an estimated pKₐ of about 14). The apparent second-order rate constant (k_cat/K_m) for substrate hydrolysis of 4-Mu β-GlcNAc by H234G thus was much less affected than the apparent first-order rate constant (k_cat), indicating less impairment of the glycosylation step (reflected in k_cat/K_m) than the deglycosylation step (reflected in k_cat), indicating a particularly important role of His²⁵⁴ in base catalysis. This leads to a significant accumulation of the covalent glycosyl-enzyme intermediate (cf. Fig. 4), which is reflected in a 24-fold reduced Michaelis-Menten constant (K_m) for H234G compared with the wild type enzyme.
groups in catalysis. The maximal catalytic activity ranges from 5.8 to 6.2 in Clark and Lubs solution (pH 5.8 – 8.0) as well as in McIlvaine buffer (pH 4.0 – 8.0). The pH activity profile of BsNagZ indicates acid ionization constants of the nucleophile and the general acid/base catalytic residue (pK_a1 and pK_a2) around 5.0 and 7.0, respectively. Hydrolysis of 4-Mu β-GlcNAc by H234G and D232G mutants was extremely slow and the pH profiles for both protein variants retained activity at alkaline pH, suggesting the elimination of a catalytic acid/base residue around 5.0 and 7.0 in the protein variants.

Accumulation of the Glycosyl-Enzyme Intermediate—Removing the general acid/base catalysts in family 3 glycosidases leads to accumulation of the glycosyl-enzyme intermediate. This was also the case in the H234G mutant as reflected by a small $K_{cat}$ value. Furthermore, direct evidence that His^{234} functions as the acid/base catalyst was performed by ESI-MS measurement. After reaction of the H234G mutant with the substrate pNP β-GlcNAc two protein species were observed (Fig. 4). The mass of 71.133 kDa corresponds to the unmodified protein, whereas the mass of 71.336 kDa corresponds to the glycosyl-enzyme intermediate. The mass difference of 203 Da is in accordance with the mass of bound N-acetylglucosaminyl.

Chemical Rescue—As shown above, deglycosylation is the rate-limiting step in hydrolysis of 4-Mu β-GlcNAc by BsNagZ. H234G, which leads to accumulation of the glycosyl-enzyme intermediate. It has been shown for many glycosidases that the addition of external small anionic nucleophiles like azide can compensate for the loss of the base residue, resulting in rescue of activity (13, 18). However, activity of the H234G mutant of BsNagZ could not be restored by the addition of azide, although replacement of His^{234} by the small glycine should allow accommodation of small compounds in the active site (see supplemental Table S1). Possibly, binding of the small anionic azide to the active site is disfavored by its negative charge. However, β-azide product formation was identified upon pNP β-GlcNAc cleavage by D232G (and to some extent also by H234G) in the presence of azide (supplemental Fig. S4). With D232G a 2-fold rate enhancement was observed in the presence of azide (supplemental Table S1), possibly due to accelerating the rate-determining glycosylation step in these mutant.

DISCUSSION

N-Acetylgalactosaminidases of family 3 of glycosidases are involved in peptidoglycan turnover and cell wall recycling of bacteria (23, 32). In this process they liberate, for instance, inducers of chromosomal β-lactamases in Gram-negative bacteria (35, 36) as well as spore germinants in the Gram-positive bacterium B. subtilis (37). A detailed understanding of the mechanism of these enzymes is a prerequisite for the design of potent selective inhibitors that may serve as novel therapeutic agents. The identification of evolutionary and structural rela-
Asp-His Dyad Glycosidase

A study now showed that family 3 β-N-acetylgalactosaminidases, instead of a glutamate general acid/base, involve an Asp-His catalytic dyad, which is unique for glycosidases. The crystal structure of BsNagZ, the first structure of a two-domain β-N-acetylgalactosaminidase of family 3 glycosidases, revealed that the Asp-His dyad superimposes with a glutamate residue that had been identified as the general acid/base catalyst in other members of family 3 glycosidases. Most intriguing evidence for the Asp

### TABLE 2
Kinetic parameters for substrate hydrolysis by BsNagZ and mutants

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Substrate</th>
<th>$K_m \mu M$</th>
<th>$k_{cat}$</th>
<th>$k_{cat}/K_m$</th>
<th>$K_m (wt)/K_m$</th>
<th>$k_{cat} (wt)/k_{cat}$</th>
<th>$(k_{cat}/K_m (wt))/(k_{cat}/K_m)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>4-Mu β-GlcNAc</td>
<td>109.6 ± 4.3</td>
<td>6.42 ± 0.07</td>
<td>58.58</td>
<td>1</td>
<td>1</td>
<td>79.2</td>
</tr>
<tr>
<td>H234G</td>
<td>4-Mu β-GlcNAc</td>
<td>4.57 ± 0.39</td>
<td>3.37 × 10⁻³</td>
<td>0.74</td>
<td>24</td>
<td>1905</td>
<td>92.0</td>
</tr>
<tr>
<td>D232G</td>
<td>4-Mu β-GlcNAc</td>
<td>56.24 ± 3.56</td>
<td>1.40 × 10⁻²</td>
<td>0.025</td>
<td>1.95</td>
<td>4586</td>
<td>2343</td>
</tr>
</tbody>
</table>

FIGURE 4. Transform of the electrospray mass spectrum of BsNagZ (left) and H234G (right) after incubation with pH 11001-11002 in triplicate. The peak with a molecular mass of 71,133 amu is the H234G mutant (the calculated value for [H234G]⁻ with His₅₉ tag and lacking methionine is 71,133 amu), the peak with a molecular mass of 71,133 amu is the H234G mutant (the calculated value for [H234G]⁻ with His₅₉ tag and lacking methionine is 71,133 amu), the peak with a molecular mass of 71,336 amu is the H234G mutant (the calculated value for [H234G]⁻ with His₅₉ tag and lacking methionine is 71,336 amu). The mass shift of 203 Da corresponds to the GlcNAc residue covalently bound to H234G.

FIGURE 5. pH activity profiles of BsNagZ and the mutants. $k_{cat}$ values at different pH were determined for wild type BsNagZ (○) and mutants H234G (□) and D232G (■). The buffers were: 0.1 M NaAc, pH 4.0–5.6, solid symbols; and 0.1 M KHPO₄, 0.1 M NaOH, pH 5.8–8.0, open symbols. Data were from a representative experiment measured in triplicate (cf. supplemental Fig. S3). Solid lines are fits of the data supposing that $k_{cat}$ values depend on the enzyme being in an acid form (H234G, D232G) or depending on two residues being in an acidic and basis form (BsnagZ).


dence that His$^{234}$ of the Asp-His dyad of family 3 $\beta$-N-acetylglucosaminidases acts as the general acid/base that is assisted by Asp$^{232}$. This contrasts with a study on Clostridium perfringens M-21 $\beta$-N-acetylglucosaminidase (Nag3A). A conserved aspartate residue (Asp$^{177}$) on the N-terminal domain was proposed as the acid/base catalyst (39) and replacement of Asp$^{232}$ with Ala abolished the activity of Nag3A. However, clear kinetic evidence for a role as acid/base catalyst of the above mentioned residue is lacking.

The amino acid histidine is perfectly qualified as a general acid/base because it has a $pK_a$ near neutrality. Nevertheless, involvement of a histidine residue, respectively, in an Asp-His dyad, is unique in glycosidas. It is, however, commonly found in enzymes cleaving phosphodiester bonds, e.g. ribonucleases (20, 40). The major role of the aspartate of the dyad in ribonucleases may be to orient the proper tautomer of the histidine for catalysis. Notably, in bovine ribonucleases N81 rather that Ne2 faces the phosphodiester (20). The His-Asp-Ser catalytic triad is renowned for serine proteases (19) and lipases (41) and the Asp-His hydrogen bond in the catalytic triad is known to contribute greatly to catalysis, potentially via forming of a strong, low-barrier hydrogen bond (42). There is no evidence for the formation of a low-barrier hydrogen bond between His$^{234}$ and Asp$^{232}$ for the family 3 $\beta$-N-acetylglucosaminidases. Possibly the major role of Asp$^{232}$ in BsNagZ is its influence on proton dissociation of Ne2 of His$^{234}$ for catalysis. Frank and Wen (43) suggested a cooperative behavior in chains of hydrogen-bonded molecules that may sharpen the acid/base behavior of the His for catalysis (reviewed in Ref. 44). The catalytic triad His$^{57}$-Asp$^{102}$-Ser$^{195}$ of chymotrypsin (bovine chymotrypsin numbering) operates in this way: Ser$^{195}$ functions as a nucleophilic catalyst, assisted by Ne2 of His$^{57}$ that serves as acid/base catalyst, and residue Asp$^{102}$ assists in acid/base catalysis by hydrogen bonding to N81 of His$^{57}$ increasing the $pK_a$ of His$^{57}$. In the native BsNagZ structure (Fig. 3A), Asp$^{232}$ is in short H-bond distance (2.6 Å) to His$^{234}$ and electron density that can be attributed to a water molecule in the H-bond distance to Ne2 of His$^{234}$. Although, at 1.4-Å resolution this has to be taken with caution, the situation clearly resembles that of chymotrypsin, in which the serine of the triad forms a short H-bond with the His$^{57}$ (19). In the liganded structure His$^{234}$-H-bonds to the glycosidic oxygen-mimicking nitrogen of PUGNAc and the Asp$^{232}$--His$^{234}$ distance is significantly larger (2.9 Å). According to the suggested mechanism (Fig. 1D), in BsNagZ the catalytic His$^{234}$ forms an ion pair with Asp$^{232}$ thereby allowing protonation of the extracyclic oxygen of the glycosidic bond and facilitating removal of the leaving group upon nucleophilic attack of Asp$^{318}$. In a second step, the His removes a proton of an incoming water molecule thereby hydrolyzing the glycosyl-enzyme intermediate.

The catalytic nucleophile and the Asp-His dyad are located ~6.3 Å apart, which is in the range of catalytic residues involved in bond cleavage via a two-step double displacement mechanism in glycosidases (1). They reside on the N-terminal domain of BsNagZ and are positioned at the carboxy-terminal ends of $\beta$-strands 5 and 7, respectively. The Asp-His dyad lays on an extended loop that occupies a position after strand 4 in the BsNagZ molecule due to the shortened strand 4 loop, thereby resembling the situation of the 4/7-superfamily of glycoside hydrolases (clan GH-A glycoside hydrolases) in which the nucleophile is positioned at the C terminus of $\beta$-strand 7 as in family 3 glycosidases but the acid/base catalyst lays in a loop extending $\beta$-strand 4 (45, 46).

Recently, the crystal structure of a single domain $\beta$-N-acetylglucosaminidases of family 3 of glycosidases from Vibrio cholerae, VcNagZ, in the presence of the competitive transition state-like inhibitor PUGNAc has been reported (36). This structure, however, provided no information regarding the identity of a putative acid/base catalyst. Apparently in this structure, a flexible loop carrying the Asp-His dyad is flipped outward. Moreover, the aspartate nucleophile within this structure is distorted, indicating a non-physiological conformation of the active site in the crystal or an unproductive binding of the inhibitor (cf. Fig. 1E). We recently solved a further BsNagZ structure, in which a short loop of the protein that carries the Asp-His dyad is moved outwards away from the active site. It can be speculated that the proper orientation of the Asp-His dyad in this enzymes is induced upon substrate binding, providing a high degree of substrate specificity.

The obvious question is why the sub-family 3 $\beta$-N-acetylglucosaminidases apparently are the only glycosidases that act by a catalytic mechanism involving an Asp-His dyad. One rational for the replacement of an acid/base glutamate for a dyad might be the negative charge of the natural substrates of these enzymes, which are MurNAc or 1,6-anhydroMurNAc containing cell wall fragments as mentioned above. The negative charge of the carboxylic acid of these molecules might interfere with the use of a negative charged acid/base catalyst in the active site. A similar situation holds for sialidases, which have been shown to utilize tyrosine as a catalytic nucleophile rather than a carboxylate nucleophile (47, 48). It was argued that the anomic center of the sialic acid sugars bears an anionic carboxylate residue and the nucleophile attack by an anionic nucleophile is therefore disfavored. Sialytransferases of family 42 glycosyltransferases have been reported to utilize a histidine as the base catalyst that abstracts the proton from the nucleophilic hydroxyl group of the sugar acceptor, thereby facilitating attack on the CMP-Neu5Ac donor nucleotide (49, 50). Further precedence for the role of histidine residues as base catalysts are reported (51, 52).

The residues His$^{234}$ and Asp$^{232}$ are completely conserved in the subfamily of $\beta$-N-acetylglucosaminidases of family 3 glycosidases located in the conserved sequence pattern KH(F/L)PG(H/L)GX(4)D(S/T)H, which is used as an identifier for members of the subfamily. The Asp-His dyad is suitably positioned to act as the acid/base catalyst (cf. supplemental Fig. S2) and may substitute the “normal” carboxylic acid acid/base catalyst to act on substrates bearing a negative charged residue. Our findings will facilitate the development of mechanism-based inhibitors that selectively target family 3 $\beta$-N-acetylglucosaminidases, which are involved in cell wall turnover, release of spore germinants, and induction of $\beta$-lactamase in bacteria.

---

Asp-His Dyad Glycosidase

Acknowledgments—We appreciate the help of the Swiss Light Source beamline staff of the Paul Scherrer Institute, Villigen, Switzerland, with data collection. We thank David Vocadlo, Simon-Fraser University, Burnaby, Canada, for providing PUGNAc inhibitor and Jocelyne Fiaux, ZMBH, University of Heidelberg, for measuring ESI-MS.

REFERENCES