Mechanism of Action and Identification of Asp242 as the Catalytic Nucleophile of *Vibrio furnisii N*-Acetyl-β-D-glucosaminidase Using 2-Acetamido-2-deoxy-5-fluoro-α-L-idopyranosyl Fluoride†

David J. Vocadlo, Christoph Mayer, Shouming He, and Stephen G. Withers*

Protein Engineering Network of Centres of Excellence of Canada and Department of Chemistry, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1

Received August 20, 1999; Revised Manuscript Received October 18, 1999

ABSTRACT: The novel mechanism-based reagent 2-acetamido-2-deoxy-5-fluoro-α-L-idopyranosyl fluoride has been synthesized, and the kinetic parameters $K_m = 0.23$ mM and $k_{cat} = 0.55$ min^{-1} for its hydrolysis by *Vibrio furnisii* β-*N*-acetylglucosaminidase (ExoII) have been determined. Investigation of mixtures of enzyme with this slow substrate by electrospray mass spectrometry revealed a high steady-state population of the 2-acetamido-2-deoxy-5-fluoro-β-L-idopyranosyl-enzyme, indicating that the hydrolytic mechanism of ExoII involves the formation and rate-determining hydrolysis of a glycosyl-enzyme intermediate. Analysis of a peptic digest of the glycosyl-enzyme by HPLC/ESMS/MS in the neutral-loss mode permitted identification of a peptide bearing the 5-fluoro-sugar moiety. Tandem MS sequencing of the labeled peptide, in conjunction with multiple sequence alignments of family 3 members, allowed the identification of Asp242 as the catalytic nucleophile within the sequence IVFS**D**DLSM.

β-*N*-Acetylglucosaminidases are glycosyl hydrolases that catalyze the hydrolysis of *N*-acetylglucosamine from the nonreducing end of oligosaccharides and glyconjugates. Considerable interest in these enzymes has developed in the past few years as a consequence of their importance in the pathology of certain lysosomal storage diseases (*1*) and their potential uses in industrial biotechnology (*2*), and as targets for antifungal agents (*3*). Within the classification system of glycosyl hydrolases devised by Henrissat, *N*-acetyl-β-glucosaminidases have been found in family 20 and, very recently, family 3 of the glycosyl hydrolases (*4*, *5*). However, *Vibrio furnisii* β-*N*-acetylglucosaminidase (ExoII) is one of only three cloned β-*N*-acetylglucosaminidases (*6*) within this family, the rest of the family comprising primarily β-glucosidases, β-xylosidases, cellodextrinases, and *exo*-β-1,3-1,4-glucanases. All members of this family are defined by a section of high sequence similarity in a region encompassing the proposed nucleophile.

The mechanism of family 3 β-glucosidases (*7*) and β-glucan exohydrolases (*8*) has been the subject of some investigation. Stereochemical outcome studies with both *Hordeum vulgare* β-glucan exohydrolase (*9*) and *Aspergillus wentii* β-glucosidase (*10*) have shown that this family operates via a retaining mechanism. Thus, owing to their sequence similarity with other family 3 enzymes, it may be expected that these β-*N*-acetylglucosaminidases catalyze hydrolysis of their substrates using a double displacement mechanism in which two active site carboxyl groups assist in the formation and breakdown of a glycosyl-enzyme intermediate. The first step involves the attack by one of the carboxylates at the anomeric center, along with general acid catalysis from the other carboxyl group, expediting departure of the leaving group to yield a glycosyl-enzyme intermediate with inverted configuration at the anomeric center. In a second step, the intermediate is hydrolyzed by general base-catalyzed attack of water at the anomeric center, resulting in cleavage of the glycosidic bond with overall net retention of the anomeric configuration (Figure 1) (*11*).

The first X-ray structure of a family 3 enzyme was that of the *H. vulgare* β-glucan exohydrolase recently published by Varghese et al. (*12*). On the basis of this structure, they tentatively proposed Asp285 as the nucleophile and Glu491 as the acid/base catalyst. The authors, however, indicated some uncertainty in their assignments of these catalytic carboxyl groups. Much earlier, Legler, using the active site directed inactivator conduritol B epoxide (*13*) and the slow substrate D-glucal with the *Aspergillus wentii* β-glucosidase (*10*), proposed an aspartic acid within a 63 residue peptide as the nucleophile. However, in a number of cases, conduritol epoxides have been found to label residues other than the catalytic nucleophile. For example, in the case of *E. coli* β-galactosidase, the D-galacto derivative labeled the acid/base residue (*14*) while in bitter almond β-glucosidase a carboxylate in close proximity to the active site was tagged by conduritol B epoxide (*15*). Working with the *A. wentii* β-glucosidase, Legler also conducted a series of kinetic experiments including kinetic isotope effects, Hammett correlations, and inhibition studies (*7*). Relative to the glucosidases in this family, the detailed mechanism of family

† Financial assistance from the Protein Engineering Network of Centres of Excellence of Canada, the Natural Sciences and Engineering Research Council of Canada, and HelixBiopharma Inc. is gratefully acknowledged.

* To whom correspondence should be addressed. Tel: (604) 822-3402; Fax: (604) 822-2847; E-mail: withers@chem.ubc.ca.

10.1021/bi991958d CCC: $19.00
Published on Web 12/09/1999

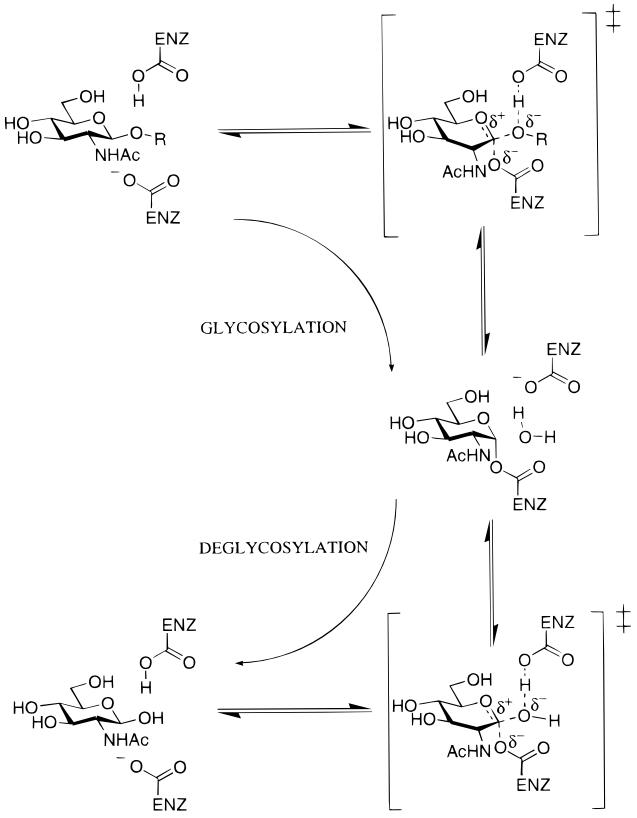


FIGURE 1: Proposed mechanism for family 3 *β*-*N*-acetylglucosaminidases.

3 *β*-*N*-acetylglucosaminidases remains essentially uncharacterized, and as will become apparent below, there remained some question as to the actual mechanism followed.

The functionally related family 20 *β*-*N*-acetylglucosaminidases have recently received considerable attention; X-ray structural data (*16*) and kinetic studies with inhibitors (*17*−*19*) have provided strong evidence for catalysis involving neighboring group participation from the 2-acetamido group. These enzymes lack an apparent nucleophile, and the 2-acetamido group of the substrate acts in its place to generate an enzyme-bound oxazolinium ion intermediate (Figure 2) (*20*). This intermediate is then hydrolyzed by general base-catalyzed attack of water at the anomeric center in a manner analogous to the double displacement mechanism described above.

The mechanism of action of another interesting, functionally related enzyme, chicken egg white lysozyme (ChEWL),[1] remains an issue of contention despite its textbook familiarity. There are three realistic mechanistic alternatives for this enzyme, which differ primarily in the nature of the enzyme intermediate: oxacarbenium ion, glycosyl-enzyme, and oxazoline. These mechanisms are expected to differ in their requirement for the apparent catalytic carboxylate Asp52. Indeed, some support for a mechanism involving anchimeric assistance has come from studies in which the nucleophile has been removed and apparently shown to not play an essential role in catalysis (*21*, *22*). In view of the residual catalytic power of these ChEWL mutants, anchimeric assistance may play a significant role in the mechanism of wild-type ChEWL or this may simply be a case of efficient chemical rescue by the substrate acetamido group, which presumably must have a high effective molarity.

The significant role that substrate assistance appears to play in the mechanisms of the family 20 hydrolases and possibly in wild-type and/or nucleophile mutants of ChEWL brings to question the mechanism of action of these family 3 *N*-acetyl-*β*-D-glucosaminidases. Whether they operate with an anchimeric assistance mechanism in which the conserved aspartate is no longer of critical importance for catalysis, or whether this residue functions as a critically important nucleophile as in most retaining glycosidases is unclear.

2-Deoxy-2-fluoro-*β*-D-glycosyl fluorides have proved to be highly successful reagents for identifying the active site nucleophiles in a range of *β*-glycosidases (*23*−*25*). 5-Fluoroglycosyl fluorides have also been found to be of great utility in the labeling of both α- and *β*-glycosidases, and, interestingly, in most cases the fluorinated version of the C-5 epimer of the natural substrate (e.g., 5-F-α-L-idosyl fluoride for a *β*-glucosidase) functioned at least as well as the 'correct' 5-fluoro-sugar (*26*−*28*). The mode of action of these mechanism-based inactivators involves the stabilization and consequent trapping of the glycosyl-enzyme intermediate. The good leaving group at C-1 accelerates the first step of the double displacement mechanism (glycosylation) whereas the C-2 or C-5 fluorine slows down both steps via inductive destabilization of the carbocationic transition states and results in the trapping of a 5-fluoro- or 2-deoxy-2-fluoroglycosyl-enzyme intermediate. Subsequent peptic digestion of the inactivated enzyme commonly yields short peptides, one of which bears the fluoro-sugar label. The sole ester in the complex, the glycosyl-enzyme linkage, is uniquely fragile to fragmentation in the collision cell of a mass spectrometer. Thus, neutral-loss tandem MS can be used to localize the labeled peptide (*29*). Purification of the tagged peptide and further tandem MS analysis permit sequencing of the peptide and identification of the catalytic nucleophile.

Our interest in the mechanisms of *β*-*N*-acetylglucosaminidases prompted us to investigate whether these enzymes employ an enzymic nucleophile. Additionally, should the enzyme utilize an enzymic nucleophile, we hoped to identify the residue using a mechanism-based inhibitor, thereby removing any ambiguity regarding the identity of this residue. Since the 2-acetamido group is important for catalysis in ExoII (*6*), it was necessary to prepare an inhibitor bearing this functional group. However, compounds in which the 2-acetamido and 2-fluoro groups are both present are unstable. Consequently, we have chosen to prepare the novel substrate analogue 2-acetamido-2-deoxy-5-fluoro-α-L-idopyranosyl fluoride. In this paper, we report the synthesis of 2-acetamido-2-deoxy-5-fluoro-α-L-idopyranosyl fluoride and its action on *V. furnisii β*-*N*-acetylglucosaminidase. Additionally, we describe the use of this reagent in conjunction with ESMS for the identification of the catalytic nucleophile.

EXPERIMENTAL PROCEDURES

Reagents, Enzymes, and Bacterial Strains. Growth media components were obtained from Difco. Plasmid-containing

[1] Abbreviations: 5FIdNAcF, 2-acetamido-2-deoxy-5-fluoro-α-L-idopyranosyl fluoride; pNPGlcNAc, *p*-nitrophenyl 2-acetamido-2-deoxy-*β*-D-glucopyranoside; 5FIdNAc-enzyme, 2-acetamido-2-deoxy-5-fluoro-*β*-L-idopyranosyl-enzyme intermediate; ISV, ion-source voltage; OR, orifice energy; LC/MS, tandem liquid chromatography/MS; MS/MS, tandem MS; ESMS, electrospray MS; TFA, trifluoroacetic acid; TIC, total ion chromatogram; ChEWL, chicken egg white lysozyme.

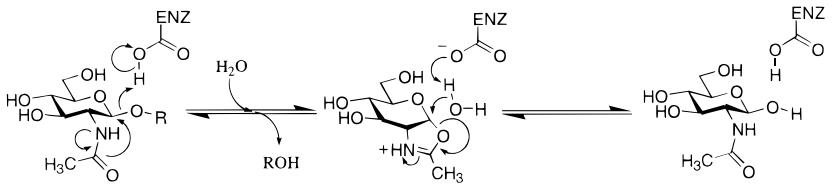


FIGURE 2: Proposed mechanism for family 20 β-*N*-acetylglucosaminidases.

strains were grown in Luria broth containing 50 μg/mL kanamycin (LB_{kan}) or in TYP (16 g/L tryptone, 16 g/L yeast extract, 5 g/L NaCl, 2.5 g/L K_2HPO_4) containing 50 μg/mL kanamycin (TYP_{kan}). Pwo DNA polymerase and deoxynucleoside triphosphates were from Boehringer Mannheim. Restriction endonucleases and T4 DNA ligase were from New England BioLabs. *Escherichia coli* Topp10 cells and the pZeroBlunt cloning kit were from Invitrogen. The pET-29b(+) expression vector, *E. coli* BL21(DE3) cells, and His-Bind metal chelation resin were from Novagen. PCR DNA fragment purification and plasmid purification kits were from Qiagen and Promega. Preparation of oligonucleotide primers and DNA sequencing were performed at the Nucleic Acids and Peptide Service facility (NAPS), University of British Columbia.

Amplification and Subcloning of ExoII. The β-*N*-acetylglucosaminidase gene from *Vibrio furnissii* (*exoII*) was amplified via polymerase chain reaction (PCR). The PCR mixture contained 10 μM oligonucleotide primers (shown below), 1 mM concentrations of the four deoxynucleoside triphosphates in 100 μL of DNA polymerase buffer, and 50 ng of plasmid pXE18, a generous gift from Dr. Saul Roseman (Department of Biology and the McCollum-Pratt Institute, The John Hopkins University, Baltimore, MD). Plasmid pXE18 contains a 1.8 kb *Sph*I/*Sal*I fragment of *V. furnissii* DNA, carrying the entire *exoII* gene. After heating the reaction mixture to 95 °C, the PCR reaction was started by adding 5 units of Pwo DNA polymerase (Boehringer Mannheim). Thirty PCR cycles (45 s at 94 °C, 45 s at 56 °C, and 70 s at 72 °C) were performed in a thermal cycler (Perkin-Elmer, GeneAmp PCR System2400). Agarose gel electrophoresis of the PCR product revealed a single DNA fragment of 980 bases.

The forward primer was as follows: 5′-AA CAT ATG GGA CCG TTA TGG CTA GAC-3′ (*Nde*I underlined). The reverse primer was as follows: 5′-AAT GTC CTC GAG GCT GTG CGC GTC AAT CAA ACG-3′ (*Xho*I underlined). After purification of the PCR product using the Qiaquick PCR purification kit according to the manufacturer's protocol (Qiagen), a blunt-end ligation into plasmid pZero2.0 was performed according to the manufacturer's protocol (Invitrogen). Electrocompetent Topp10 cells (Invitrogen) were subsequently transformed with the ligation mixture using a BioRad GenePulser II. Single colonies were selected and grown overnight in LB_{kan}, and DNA was isolated via the minipreparation technique (Promega Wizard *Plus* kit). Restriction endonuclease mapping revealed positive clones, which were subsequently sequenced to verify the published sequence of *exoII*.

The cloned *exoII* in pZero2.0 was cut out using the unique sites engineered into the oligonucleotide primers (see above). The 6-histidine fusion protein expression vector, pET-29b-(+) (Novagen), was also digested with *Nde*I and *Xho*I. A ligation reaction was performed at a ratio of 10:1 (insert to vector) using gel-purified DNA fragments and T4 DNA ligase (1 unit/10 ng of DNA) at 25 °C. The cloned product, called pET29exoII, was subsequently transformed into electrocompetent Topp10 cells, selected by the kanamycin resistance conferred by pET-29b(+). Single colonies were selected and grown overnight in LB_{kan}, and DNA was isolated via the minipreparation technique. Restriction endonuclease mapping revealed positive clones. Topp10 transformed cells were used for preparation of large amounts of plasmid pET29exoII (Qiagen Plasmid Maxi kit) and long-term storage of the vector as glycerol stock.

Overexpression and Purification of 6-Histidine-Tagged ExoII. The constructed pET29exoII expression vector was used to transform electrocompetent BL21(DE3) cells. The *E. coli* transformants were selected on LB_{kan} (50 μg/mL) agar plates. A single colony was picked to grow overnight in 3 mL of LB_{kan}, and this culture was subsequently used to inoculate 1 L of TYP_{kan}. After the culture grew to an A_{600} of 2−3 at 30 °C, 0.4 mM isopropyl-β-D-thiogalactoside (IPTG) was added to induce protein expression from the lac promoter and grown for an additional 4 h at 25 °C. Overexpression of the enzyme was monitored by sampling both induced and noninduced cells using SDS−PAGE. Induced cells were then harvested and suspended in 30 mL of binding buffer (5 mM imidazole, 500 mM NaCl, 20 mM Tris-HCl, pH 7.9). The cell suspension was passed 2 times through a French press at 0−5 °C and centrifuged at 10000*g* for 30 min at 4 °C to yield soluble cell extract.

A 60 mL slurry of His-Bind resin (nitriloacetic acid−agarose, Novagen) was placed into a 50 mL column, yielding a bed volume of 30 mL. The column was washed with 10 bed volumes of sterile deionized water and then charged with nickel by adding 5 bed volumes of 50 mM $NiSO_4$. Unbound Ni^{2+} was washed away with 5 bed volumes of binding buffer (see above). The soluble cell extract from a 1 L cell culture was applied to the column. The column was washed with 5 bed volumes of binding buffer. ExoII protein was eluted from the Ni^{2+}-column with a linear imidazole gradient from 30 to 200 mM in a buffer containing 500 mM NaCl and 20 mM Tris-HCl (pH 7.9). Fractions (3 mL) were collected (1.0 mL/min flow rate) and assayed for hexosaminidase activity using 4-methylumbelliferyl β-*N*-acetyl-D-glucosaminide. Active fractions were further analyzed via SDS−polyacrylamide gel electrophoresis. Fractions containing pure enzyme were pooled, mixed with glycerol to give a 20% solution, and stored at 4 °C.

General Procedures and Synthesis. All buffer chemicals and other reagents were obtained from the Sigma/Aldrich Chemical Co. unless otherwise noted. Synthetic reactions were monitored by TLC using Merck Kieselgel 60 F_{254} aluminum-backed sheets. Compounds were detected by charring with 10% ammonium molybdate in 2 M H_2SO_4 and heating. Flash chromatography under a positive pressure was performed with Merck Kieselgel 60 (230−400 mesh) using

the specified eluants. ^{1}H NMR spectra were recorded on a Bruker WH-400 spectrometer at 400 MHz (chemical shifts quoted relative to $CDCl_3$ or to DSS when taken in D_2O). ^{19}F NMR spectra were recorded on a Bruker AC-200 at 188 MHz and are proton-coupled with CF_3CO_2H as a reference.

*3,4,6-Tri-O-acetyl-2-acetamido-2-deoxy-5-fluoro-*α*-*L*-idosyl Fluoride (**3**)*. 3,4,6-Tri-*O*-acetyl-2-acetamido-2-deoxy-*β*-D-glucopyranosyl fluoride (**1**) (*30*) (0.922 g, 4.1 mmol) was suspended in carbon tetrachloride (220 mL), and *N*-bromosuccinimide (2.36 g, 13.3 mmol) was added. This suspension was heated to reflux between two 200 W light bulbs. After 2 h, the reaction mixture was allowed to cool to room temperature, and dichloromethane (160 mL) was added. The solution was washed with saturated sodium hydrogen carbonate (150 mL) and water (2 × 150 mL). The organic layer was dried ($MgSO_4$) and filtered, and the solvent was removed in vacuo. The resulting yellow gum was immediately purified on a short column of silica gel (ethyl acetate/hexanes/triethylamine, 2:1:0.015) to yield the crude intermediate 3,4,6-tri-*O*-acetyl-2-acetamido-5-bromo-2-deoxy-*β*-D-glucosyl fluoride (**2**): ^{19}F NMR (188 MHz, $CDCl_3$, CF_3CO_2H reference) δ −44.7 (1 F, dd, $J_{F1,H1}$ = 57.0 Hz). This golden syrup (223 mg) was dried under vacuum for 3 h, and under an atmosphere of nitrogen silver fluoride (0.94 g, 7.4 mmol), and activated 4 Å molecular sieves (50 mg) were added. Dry acetonitrile (50 mL) was added to the flask, and the reaction mixture was stirred under nitrogen for 18 h. The reaction mixture was then filtered through Celite/silica to remove silver salts and the solvent removed in vacuo to yield a translucent pale yellow syrup. This residue was purified by careful gradient flash chromatography (ethyl acetate/hexanes, 1:1 to 2:1) to yield the product (**3**) as a clear gum (120 mg, 0.33 mmol, 8%): ^{19}F NMR (188 MHz, $CDCl_3$, CF_3CO_2H reference) δ −44.3 (1 F, dd, $J_{F1,H1}$ = 52.4 Hz, $J_{F1,H2}$ = 10.1 Hz, F-1), −30.8 (1 F, dddd, $J_{F5,H6}$ = 17.0 Hz, $J_{F5,H6'}$ = 15.3 Hz, $J_{F5,H4}$ = 8.5 Hz, $J_{F5,H1}$ = 1.5 Hz, F-5); ^{1}H NMR (400 MHz, $CDCl_3$) δ 5.99 (1 H, d, $J_{NH,H2}$ = 8.6 Hz, NH), 5.63 (1 H, ddd, $J_{H1,H2}$ = 3.9 Hz, H-1), 5.47 (1 H, dd, $J_{H4,H3}$ = 5.6 Hz, H-4), 5.08 (1 H, dd, $J_{H3,H2}$ = 6.0 Hz, H-3), 4.47 (1 H, ddd, H-2), 4.31 (1 H, d, H-6), 4.31 (1 H, d, H-6′), 2.1 (9 H, s, 3 × OAc), 2.0 (3 H, s, NAc); +CIMS (NH_3): *m*/*z* 385 $(M + NH_4)^+$ (22%), 368 $(M + H)^+$ (100%), 348 $(M - F)^+$ (20%); high-resolution +CIMS $(M + NH_4)^+$: calculated, 368.1158; found, 368.1150.

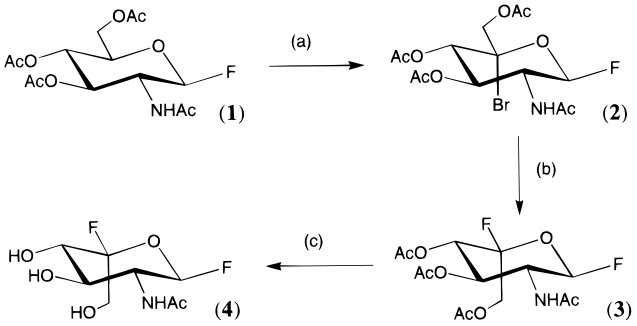
*Synthesis of 2-Acetamido-2-deoxy-5-fluoro-*α*-*L*-idosyl Fluoride (**4**)*. The protected 5-fluoro-sugar (23.5 mg, 0.064 mmol) (**3**) was dissolved in dry methanol (3 mL). After being cooled to 4 °C, the stirred solution was saturated with ammonia. This mixture was warmed to room temperature and stirred for 25 min, after which time the solvent was evaporated in vacuo to give a clear gum. The desired product was then purified by flash chromatography on silica gel (ethyl acetate/methanol/water, 17:2:1) to yield the desired product (**4**) (13.6 mg, 0.056 mmol, 88%) as a transparent gum: ^{19}F NMR (188 MHz, CD_3OD, CF_3CO_2H reference) δ −44.1 (1 F, dd, $J_{F1,H1}$ = 54.7 Hz, $J_{F1,H2}$ = 13.6 Hz, F-1), −33.2 (1 F, dddd, $J_{F5,H6}$ = 14.6 Hz, $J_{F5,H6'}$ = 12.2 Hz, $J_{F5,H4}$ = 12.2 Hz, $J_{F5,H1}$ = 1.7 Hz, F-5); ^{1}H NMR (400 MHz, CD_3OD) δ 5.49 (1 H, dd, $J_{H1,H2}$ = 3.9 Hz, H-1), 4.19 (1 H, ddd, $J_{H2,H3}$ = 8.4 Hz, H-2), 3.98 (1 H, dd, $J_{H4,H3}$ = 7.1 Hz, H-4), 3.80 (1 H, dd, $J_{H6,H6'}$ = 12.2 Hz, H-6), 3.72 (1 H, dd, H-6′), 3.71 (1 H, dd, H-3), 2.01 (3 H, s, NAc); +CIMS (NH_3): *m*/*z* 242 $(M + H)^+$ (29%), 222 $(M - F)^+$ (100%); high-resolution +CIMS $(M + H)^+$: calculated, 242.0840; found, 242.0838.

Enzyme Kinetics. Kinetic studies were all performed in 50 mM sodium phosphate, 200 mM sodium chloride, pH 7.00, containing 0.1% bovine serum albumin. A continuous spectrophotometric assay based on the rate of release of *p*-nitrophenolate upon hydrolysis of pNPGlcNAc was used. The resulting absorbance change was measured at 400 nm ($\Delta\epsilon_{400}$ = 7280 mM^{-1} cm^{-1}) using a Pye-Unicam UV/VIS spectrophotometer equipped with a circulating water bath at 23.5 °C. Michaelis−Menten parameters for the hydrolysis of pNPGlcNAc by ExoII were determined by directly fitting the intial rate data to the Michaelis−Menten equation using GraFit version 3.0 (*31*). K_m and k_{cat} values for the hydrolysis of 5FIdNAcF by ExoII were determined by monitoring the release of fluoride using a 9606BN Ionplus Orion fluoride ion electrode interfaced to a pH/ion selective meter (Fischer Scientific) at 25 °C. Stock enzyme (10 μL) was added to glass cells containing various concentrations of 5FIdNAcF in the above-mentioned buffer to give a final volume of 250 μL. Reaction cells were preequilibrated in a water bath to 25 °C. It is known that 5-fluoroglycosyl fluorides, after being hydrolyzed to give the hemiacetal, rapidly lose another fluoride ion, resulting in a net release of two fluoride ions for each catalytic event. Initial rates, which were measured over 5 min, were therefore corrected by dividing by a factor of 2. Calculation of kinetic parameters for this substrate was carried out by directly fitting the intial rate data to the Michaelis−Menten equation using GraFit version 3.0 (*31*).

Labeling and Proteolysis. Labeling of *V. furnisii β-N*-acetylglucosaminidase (8.2 mg/mL) was accomplished by incubating the enzyme with 5FIdNAcF (8.2 mM) for 5 min in 50 mM sodium phosphate, 200 mM sodium chloride, pH 7.00. This sample was then analyzed immediately by injecting the mixture onto a reverse-phase column (PLRP-S, 1 × 50 mm) equilibrated with solvent A [solvent A: 0.05% trifluoroacetic acid (TFA)/2% acetonitrile in water] on an Ultrafast Microprotein Analyzer (Michrom BioResources Inc., Pleasanton, CA). Elution of the enzyme was accomplished using solvent A at a flow rate of 50 μL/min. Proteolytic digestion of the enzyme was performed by mixing the labeled enzyme (25 μL of 8.2 mg/mL) with 20 μL of 2.1 M sodium phosphate, pH 1.7, 135 μL of H_2O, and 20 μL of 1 mg/mL pepsin in 200 mM sodium phosphate, pH 2.0. This sample and a control in which the enzyme was not exposed to the inhibitor were incubated at 25 °C for 2 h. At 0.25, 0.5, 1.0, and 2.0 h, aliquots (50 μL) of the sample were frozen over CO_2/acetone. Analysis of these samples by ESMS revealed that the 2 h digest ensured complete digestion of the enzyme and generated peptides of a size suitable for sequencing by MS/MS.

ESMS Analysis of the Proteolytic Digest. Mass spectra were recorded on a PE-Sciex API 300 triple-quadrupole mass spectrometer (Sciex, Thornhill, Ontario, Canada) equipped with an Ionspray ion source. Peptides were separated by reverse-phase HPLC on an Ultrafast Microprotein Analyzer (Michrom BioResources Inc.) directly interfaced with the mass spectrometer. In each of the MS experiments, the proteolytic digest was loaded onto a C18 column (Reliasil, 1 × 150 mm) equilibrated with solvent A [solvent A: 0.05% trifluoroacetic acid (TFA)/2% acetonitrile in water]. Elution of the peptides was accomplished using a gradient (0−60%)

Scheme 1[a]



[a] (a) NBS, CCl_4, *hν*. (b) AgF, MeCN, 8% over two steps. (c) NH_3, MeOH, 88%.

of solvent B over 60 min followed by 100% solvent B over 2 min (solvent B: 0.045% TFA/80% acetonitrile in water). Solvents were pumped at a constant flow rate of 50 μL/min. Spectra were obtained in the single-quadrupole scan mode (LC/MS), the tandem MS neutral-loss scan mode, or the tandem MS product-ion scan mode (MS/MS). In the single-quadrupole mode (LC/MS), the quadrupole mass analyzer was scanned over a mass-to-charge ratio (*m*/*z*) range of 300−2200 Da with a step size of 0.5 Da and a dwell time of 1.5 ms per step. The ion source voltage (ISV) was set at 5.5 kV, and the orifice energy (OR) was 45 V. In the neutral-loss scanning mode, MS/MS spectra were obtained by searching for the mass loss of $m/z = 222$, corresponding to the loss of the 5FIdNAc label from a peptide in the singly charged state. In the tandem MS daughter-ion scan mode, the spectrum was obtained by selectively introducing the parent ion ($m/z = 1248$) from the first quadrupole (Q1) into the collision cell (Q2) and observing the product ions in the third quadrupole (Q3). Thus, Q1 was locked on *m*/*z* 1248; the Q3 scan range was 50−1260; the step size was 0.5; the dwell time was 1 ms; ISV was 5 kV; OR was 45 V; Q0 = −10; IQ2 = −48.

Aminolysis of the Labeled Enzyme Digest. Concentrated ammonium hydroxide (5 μL) was added to an aliquot of the 2 h digest mixture (10 μL, 1.0 mg/mL). This sample was incubated for 15 min at 50 °C, acidified with 50% TFA, and analyzed by LC/MS as described above.

RESULTS AND DISCUSSION

Synthesis. 5FIdNAcF (**4**) (Scheme 1) was synthesized from 3,4,6-tri-*O*-acetyl-2-acetamido-2-deoxy-β-D-glucosyl fluoride (**1**) by a three-step procedure as follows (Scheme 1). Radical bromination (*32*) of **1** in carbon tetrachloride generated the 5-bromo derivative (**2**) although attempts to isolate this material were unsuccessful. Monitoring of a crude sample of 3,4,6-tri-*O*-acetyl-2-acetamido-5-bromo-2-deoxy-β-D-glucosyl fluoride (**2**) in $CDCl_3$ by ^{19}F NMR suggested that the product underwent significant decomposition overnight. The bromo compound (**2**) was therefore used immediately after minimal purification on a plug of silica. Displacement of the bromine at C-5 using silver fluoride in acetonitrile yielded the product with the inverted stereochemistry at C-5, namely, 3,4,6-tri-*O*-acetyl-2-acetamido-2-deoxy-5-fluoro-α-L-idosyl fluoride (**3**), in a modest but acceptable yield. Deprotection of the acetylated 5-fluoro compound (**3**) was accomplished by treatment with ammonia in methanol to yield, after purification, the target compound 5FIdNAcF (**4**). We

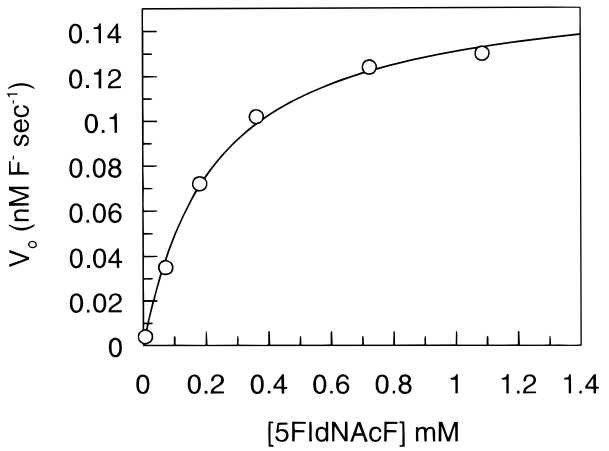
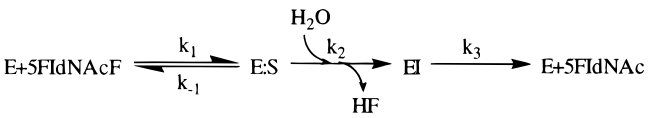


FIGURE 3: Michaelis-Menten plot of initial rates of hydrolysis of 5FIdNAcF by *Vibrio furnisii* ExoII.

opted to conduct the reaction series on the desired 2-acetamido compound and accept the loss of yield arising from partial N-halogenation of the alkyl amide rather than employ a transient protecting group on a 2-amino-sugar. This decision was reached on the basis of avoiding potential difficulties in deprotection and reacetylation during subsequent steps, since compounds with two anomeric fluorine atoms are relatively reactive.

Kinetic Studies. Comparison of the Michaelis−Menten parameters for the hydrolysis of pNPGlcNAc by the His_6-tagged ExoII (23.5 °C, $K_m = 0.73$ mM, $V_{max} = 2.37$ μmol mg^{-1} min^{-1}) used in this study and the wild-type ExoII as reported by Roseman (22 °C, $K_m = 0.44$ mM, $V_{max} = 1.1$ μmol mg^{-1} min^{-1}) indicates that modification of the wild-type ExoII by addition of the hexa-histidine tag does not alter its catalytic function significantly. The greater specific activity of the His_6-tagged ExoII may arise from the expedient purification of the His_6-tagged ExoII as compared to the wild-type ExoII that is known to be thermally unstable.

Incubation of *V. furnisii β-N*-acetylglucosaminidase with 5FIdNAcF did not result in time-dependent inactivation of the enzyme relative to a control. The absence of any observable inactivation of the enzyme by 5FIdNAcF therefore prompted us to investigate whether the enzyme was capable of hydrolyzing this analogue. Rates of release of fluoride upon incubation of *V. furnisii β-N*-acetylglucosaminidase at a number of different inhibitor concentrations [(0.15−5) × K_m] as measured by a fluoride ion selective electrode indicated that the enzyme hydrolyzed 5FIdNAcF as a Michaelian substrate ($K_m = 0.23 \pm 0.02$ mM, $k_{cat} = 0.0091 \pm 0.0004$ s^{-1}, Figure 3). The hydrolysis of 5FIdNAcF by the enzyme follows the scheme shown below (eq 1) where



E = enzyme, F = fluoride, 5FIdNAc = 2-acetamido-2-deoxy-5-fluoro-L-idose, E:S = Michaelis complex, and EI = enzyme intermediate. As mentioned above, the 5FIdNAc product rapidly decomposes with the release of a second equivalent of fluoride; thus, all kinetic parameters have been calculated on the basis of 2 equiv of fluoride released per molecule of inhibitor.

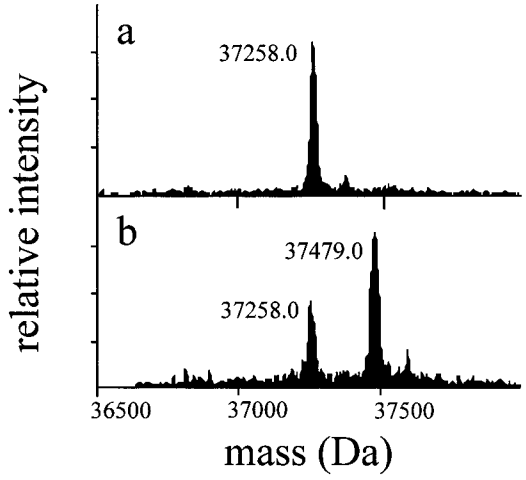


FIGURE 4: Transform of the electrospray mass spectrum of (a) native ExoII and (b) ExoII incubated with 8.2 mM 5FIdNAcF.

The observed K_m is somewhat lower than that observed for pNPGlcNAc ($K_m = 0.44$ mM) despite the absence of an aromatic aglycon, which has generally been found to enhance binding. This relatively low K_m is even more surprising given the structurally significant epimerization of the 5-position from the configuration found in GlcNAc to that found in IdNAc. This modification would be expected to result in decreased inherent binding affinity of the enzyme for the inhibitor. It therefore seems likely that this relatively low K_m value has its origin in a kinetic phenomenon rather than reflecting enhanced binding interactions. The K_m value for a two-step enzyme having a ping-pong-type mechanism in which one of the substrates is water is expressed as

$$K_m = \left(\frac{k_{-1} + k_2}{k_1}\right)\left(\frac{k_3}{k_2 + k_3}\right) \qquad (2)$$

As can be seen from the expression, a low K_m can result from an increase in k_1 relative to k_{-1} or k_2, or a decrease in k_3 relative to k_2, which in either case results in an increased steady-state concentration of the enzyme intermediate (EI). However, in this case the low observed k_{cat} suggests that k_3 has decreased significantly relative to k_2 such that the breakdown of the enzyme intermediate is now rate-determining. Such kinetic behavior is consistent with that observed for 5-fluoroglycosyl fluorides studied previously, as discussed in the introduction (*26*−*28*), and opens the possibility of investigating the nature of what must be a fairly well-populated enzyme intermediate.

Nature of the Intermediate. The mass of native *V. furnisii* *β-N*-acetylglucosaminidase was found by ESMS to be 37 258 Da (Figure 4a), which is in agreement, within error, with the theoretical mass of the cloned enzyme (37 247 Da). After incubation with 5FIdNAcF, two species are observed: the native, unlabeled enzyme; and another species with a mass of 37 479 Da (Figure 4b). The mass difference observed between the native and inhibited enzymes is 221 Da, a value which is consistent, within error, with the addition of a single 5FIdNAc label (222 Da). The partial labeling of the enzyme is consistent with the expectation of a high steady-state population of the enzyme intermediate rather than a stoichiometrically inactivated species.

The observation of a covalent glycosyl intermediate provides strong evidence for a mechanism involving an enzymic nucleophile rather than the mechanism involving substrate assistance from the 2-acetamido group followed by the family 20 *β-N*-acetylglucosaminidases. This latter mechanism involves the formation of an oxazoline or oxazolinium ion intermediate, noncovalently bound to the enzyme, and it is very unlikely that such a complex would survive the LC/ESMS conditions. Indeed, subsequent peptide analysis (vide infra) confirms the covalent nature of the intermediate. This enzyme therefore operates by a double displacement mechanism involving an enzymic nucleophile as do other family 3 members. The relatively large steady-state 5FIdNAc-enzyme population observed by ESMS should permit the determination of the site of attachment of the label.

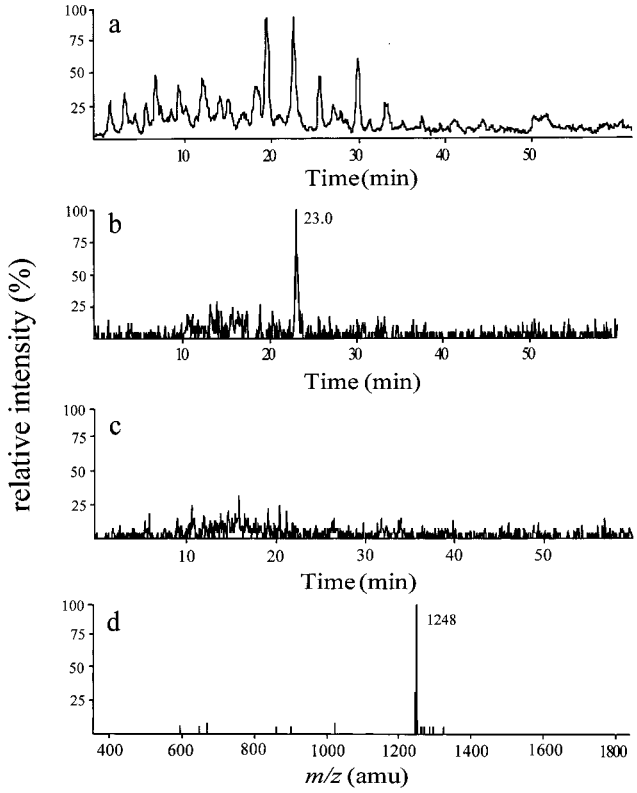


FIGURE 5: ESMS experiments on peptic digests of *V. furnisii* ExoII. (a) Enzyme incubated with 5FIdNAcF, TIC in normal MS mode. (b) Enzyme incubated with 5FIdNAcF, TIC in the neutral-loss mode. (c) Unlabeled enzyme, TIC in the neutral-loss mode. (d) Mass spectrum of the peptide eluting at 23.0 min.

Identification of the Labeled Active Site Peptide. Peptic hydrolysis of the 5FIdNAc-enzyme resulted in a mixture of peptides which were separated by reverse-phase HPLC using the ESMS as detector. When scanned in the normal LC/MS mode, the total ion chromatogram (TIC) showed a large number of peaks, each corresponding to one or more peptides in the digest mixture (Figure 5a). The 5FIdNAc-peptide was then located in a second experiment using a tandem mass spectrometer set up in the neutral-loss mode. This technique involves the limited fragmentation of the ions by collision with an inert gas (N_2) in a collision chamber (Q2) between two quadrupoles (Q1 and Q3). Since the ester linkage between label and enzyme is one of the most labile linkages present, homolytic cleavage of this bond is expected to occur in the collision chamber. The resulting neutral loss of this label leaves the peptide with an unchanged charge state but a mass decrease of 222 Da. The two quadrupoles were therefore scanned in a linked mode to permit only passage

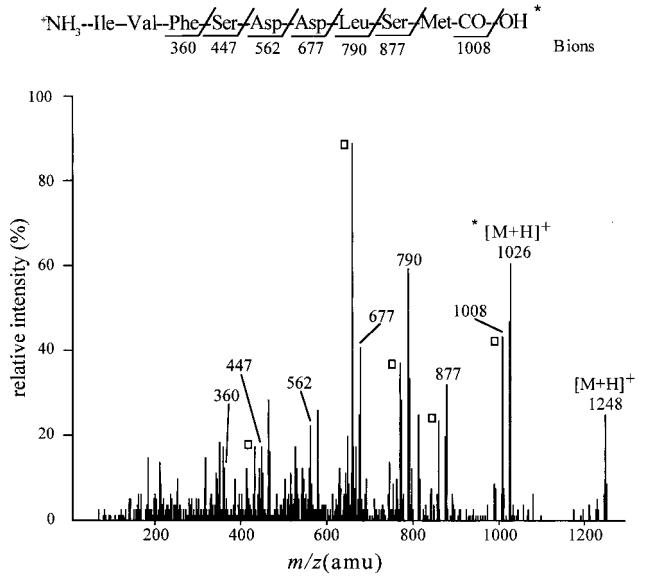


FIGURE 6: ESMS/MS daughter-ion spectrum of the 5FIdNAc-peptide (*m*/*z* 1248, in the singly charged state). Observed B series fragments are shown below the peptide sequence. An asterisk indicates the unlabeled peptide arising from neutral loss of the label. B ions arising from dehydration of serine are indicated by (□).

of those ions which lose a mass of 222 Da in the collision cell. A peak at 23.0 min was observed (Figure 5b), and no such peak was observed in a control in the neutral-loss spectrum of the unlabeled *β-N*-acetylglucosaminidase (Figure 5c). These results indicate that a singly charged peptide bearing the 5FIdNAc label is preferentially detected with a mass of 1248 Da (Figure 5d); thus, the uncharged peptide has a mass of 1247 Da (1248 − 1 H). Since the mass of the label is 222 Da, the unlabeled, protonated peptide must have a mass of 1026 Da (1247 − 222 + 1 H). Aminolysis of the labeled peptide resulted in disappearance of the peak in the ion chromatogram of *m*/*z* 1247 (±1) (data not shown) and an increase in the intensity of the peak at 1027. The mass loss of 221 Da (1248 − 1027) resulting from aminolysis of the labeled peptide to yield a lower molecular weight species (*m*/*z* = 1027) is consistent, within error, with the expected mass loss from cleavage of the ester-linked 5FIdNAc label. The susceptibility of the label to aminolysis provides evidence for an ester linkage between the sugar moiety and the enzyme and is entirely consistent with the formation of a glycosyl-enzyme intermediate on an enzymic carboxylate. Candidate peptides of this mass (1026 Da) were identified by inspection of the enzyme amino acid sequence and searching for all possible peptides of this mass. Three possible peptides with a mass of 1026 ± 1 Da were identified: $_{135}$DVQTVLTYS$_{143}$, $_{238}$IVFSDDLSM$_{246}$, and $_{287}$PISVVPQAQS$_{296}$ The peptide bearing the label was then identified unambiguously by peptide sequencing using MS/MS.

Peptide Sequencing. Information on the sequence was obtained by additional fragmentation of the peptide of interest (*m*/*z* 1248) in the daughter ion scan mode (Figure 6). Both the labeled parent ion and the unlabeled peptide (*m*/*z* = 1026) arising from loss of the 5FIdNAc label appear as singly charged species. Peaks resulting from B ions correspond to IVF (*m*/*z* 360), IVFS (*m*/*z* 447), IVFSD (*m*/*z* 562), IVFSDD (*m*/*z* 677), IVFSDDL (*m*/*z* 790), IVFSDDLS (*m*/*z* 877), and IVFSDDLSM (*m*/*z* 1008). Several peaks with *m*/*z* values 18 units lower than the B ions are observed. These ions are believed to arise from dehydration of a serine residue and are marked (□) on the spectrum. This sequence corresponds exactly to that of $_{238}$IVFSDDLSM$_{246}$ as deduced from the gene sequence. Within this peptide, there are four amino acids that could potentially function as nucleophiles: two serines and two aspartic acids. Precedent alone would make the two serine residues unlikely candidates, and both the susceptibility of the linkage to aminolysis (vide supra) and the lack of sequence conservation (vide infra) confirm this. This leaves two aspartic acid residues that could function as the catalytic nucleophile, and, unfortunately, the MS/MS sequencing does not allow the assignment of which one was labeled. However, sequence conservation and precedent from other enzymes provides a clear answer, as described below.

Multiple Sequence Alignments. Multiple sequence alignments of the members of family 39 glycosyl hydrolases revealed that Asp242 is fully conserved among all known family members whereas Asp243 is not, as shown in the partial multiple sequence alignment of Figure 7. This observation, in conjunction with the results from the tandem MS sequencing of the labeled peptide, permits the identification of Asp242 as the catalytic nucleophile. This assignment is in agreement with the results of both Legler and Fincher and thereby both unambiguously establishes the identity of the nucleophile within this family and demonstrates that the *β-N*-acetylglucosaminidases within this family operate via the same mechanism as that of the glucosidases.

Interestingly, members of this family are essentially of 2 different sizes: the smaller enzymes being approximately 340 residues long while the larger enzymes contain, on average, about 550 residues. The recently solved structure of the *Hordeum vulgare β*-glucan exohydrolase revealed this enzyme has a two-domain structure in which the active site is a cleft formed by the interface of the two domains (*12*). In light of this information and the sequence alignments, it is expected that the smaller enzymes within this family are composed of a single domain that is equivalent to the first domain of the larger enzymes, and therefore must have a somewhat different active site architecture. Of particular interest in this regard, Fincher has tentatively proposed Glu491, a residue that is positioned on the second domain, as the catalytic acid/base residue. This observation therefore raises a question regarding the identity of the acid/base catalyst in the single-domain enzymes. The bell-shaped pH/activity profile observed for ExoII indicates this enzyme bears a catalytic acid/base residue despite lacking this second domain. It is possible that ExoII and other single-domain family 3 enzymes bear an extended loop that partially occupies the position of the second domain in close contact with the active site and includes the catalytic acid/base residue. However, structural and kinetic studies must be undertaken to bear out such speculation. Another interesting difference noted in multiple sequence alignments between the *β-N*-acetylglucosaminidases and other members of this family is a region that lies at the end of *β*-strand *e*. The highly conserved residues KHF, found after the *β*-strand, have been shown to contain two cis-peptide bonds in the *Hordeum vulgare β*-glucan exohydrolase structure, and come in close contact with the C-3 and C-4 hydroxyls of the substrate. Immediately following this unusual feature is a stretch of 11 amino acids within a loop region that is well conserved

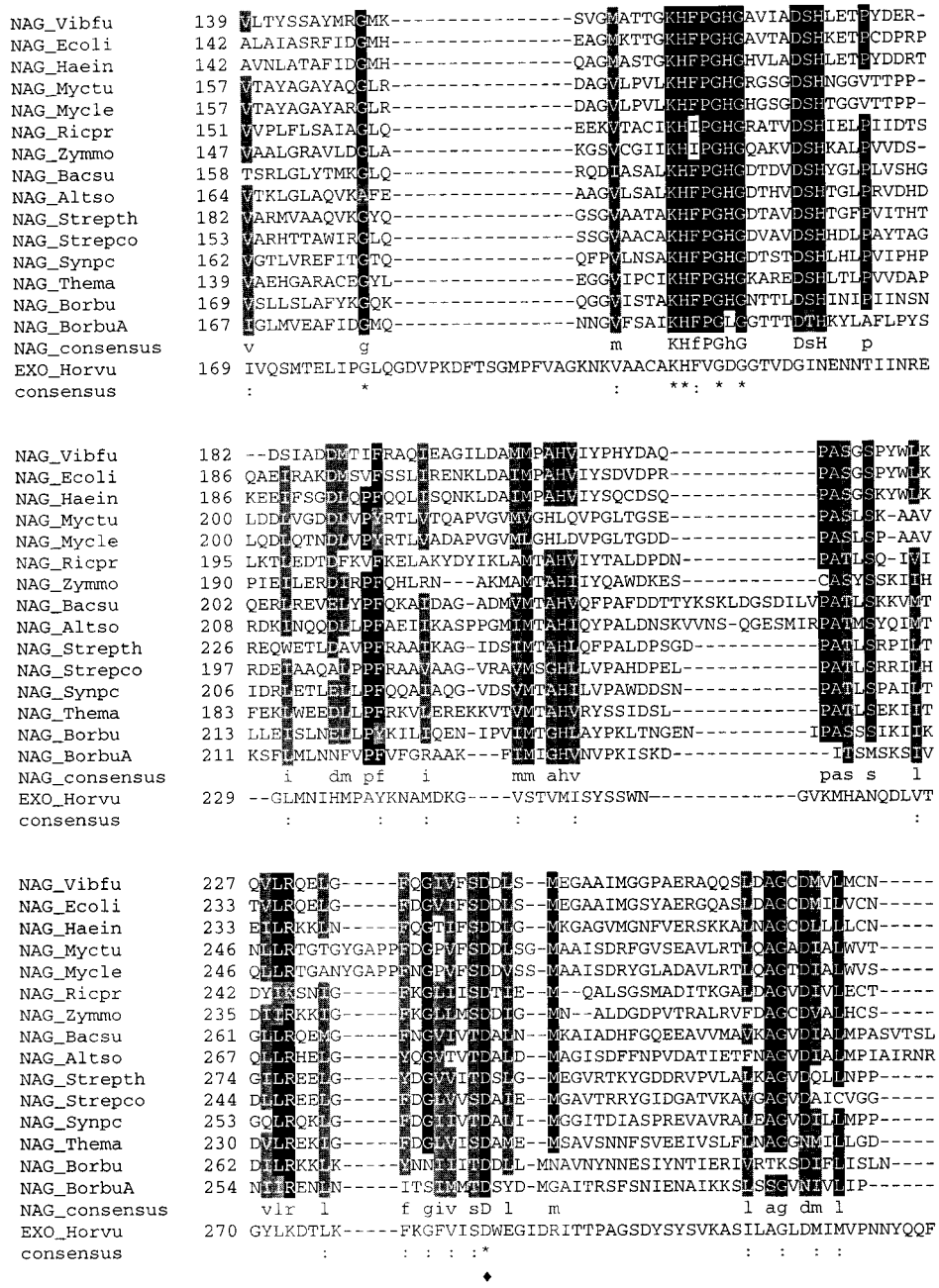


FIGURE 7: Partial multiple sequence alignment of the cloned and putative family 3 *β-N*-acetylglucosaminidases. Consensus of *β-N*-acetylglucosaminidases is shown below (NAG__consensus) with upper case letters indicating entirely conserved residues and lower case letters indicating similar residues. *β-N*-Acetylglucosaminidases were aligned using ClustalW (*34*). Dark shading indicates highly conserved residues, while light shading indicates conserved similar residues. Shading was performed using Version 3.21 of BoxShade. The sequence of *Hordeum vulgare β*-glucan exohydrolase was manually aligned in conjunction with use of ClustalW. The consensus of all sequences is shown (consensus) at bottom: (*) indicates entirely conserved residues, and (:) indicates similar residues. Numbers to the left denote residue positions. The abbreviations used, references to the published sequence, and data bank accession numbers are as follows: NAG__Vibfu, *N*-acetyl-*β*-D-glucosaminidase from *V. furnisii* (*6*) (GenBank identifier U52818); NAG__Ecoli, putative *N*-acetyl-*β*-D-glucosaminidase from *Escherichia coli* (*35*) (GenBank identifier AE000302); NAG__Haein, putative *N*-acetyl-*β*-D-glucosaminidase from *Haemophilus influenzae* (*36*) (GenBank identifier U32777); NAG__Myctu, putative *N*-acetyl-*β*-D-glucosaminidase from *Mycobacterium tuberculosis* (*37*) (GenBank identifier AL021929); NAG__Mycle, putative *N*-acetyl-*β*-D-glucosaminidase from *Mycobacterium leprae* (*38*) (GenBank identifier CAA18563); NAG__Ricpr, putative *N*-acetyl-*β*-D-glucosaminidase from *Rickettsia prowazekii* (*39*) (GenBank identifier CAA15141); NAG__Zymmo, putative *N*-acetyl-*β*-D-glucosaminidase from *Zymomonas mobiliz* (*40*) (GenBank identifier AF124757.1); NAG__Bacsu, putative *N*-acetyl-*β*-D-glucosaminidase from *Bacillus subtilis* (*41*) (GenBank identifier AB002150); NAG__Altso, *N*-acetyl-*β*-D-glucosaminidase from *Alteromonas* sp. (*42*) (GenBank identifier D17399); NAG__Strepth, *N*-acetyl-*β*-D-glucosaminidase from *Streptomyces thermoviolaceus* (*4*) (GenBank identifier AB008771); NAG__Strepco, putative *N*-acetyl-*β*-D-glucosaminidase from *Streptomyces coelicolor* (*43*) (GenBank identifier AL023702); NAG__Synpc, putative *N*-acetyl-*β*-D-glucosaminidase from *Synechocystis* sp. (*44*) (Genbank identifier D90914); NAG__Thema, putative *N*-acetyl-*β*-D-glucosaminidase from *Thermotoga maritima* (*45*) (Genbank identifier AE001748); NAG__Borbu, putative *N*-acetyl-*β*-D-glucosaminidase from *Borrelia burgdorferi* (*46*) (Genbank identifier AE001163); NAG__BorbuA, putative *N*-acetyl-*β*-D-glucosaminidase from *Borrelia burgdorferi* (*46*) (Genbank identifier AE001115); EXO__Horvu, *exo-β*-1,3-1,4-glucanase from *Hordeum vulgare* (*9*) (Genbank identifier AF102868). The nucleophile identified in this paper and the corresponding residue in *Hordeum vulgare exo-β*-1,3-1,4-glucanase are indicated by (◆).

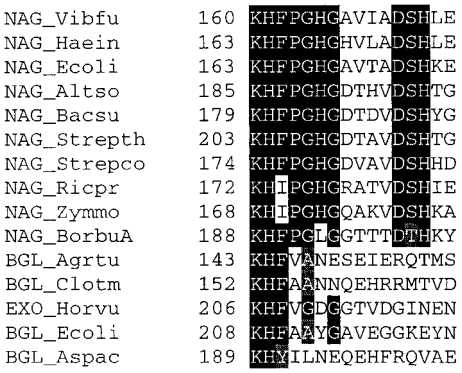


FIGURE 8: Partial multiple sequence alignment of 15 selected family 3 members including 10 β-*N*-acetylglucosaminidases. The region aligned extends from the end of β-strand *e* into a loop region containing the putative *N*-acetyl binding region of the β-*N*-acetylglucosaminidases. Alignments were performed using ClustalW (*34*). Dark shading indicates highly conserved residues, while light shading indicates conserved similar residues. Shading was performed using Version 3.21 of BoxShade. Abbreviations used, references to the published sequences, and data bank accession numbers can be found in the legend of Figure 7 or below: BGL_Agrtu, β-glucosidase from *Agrobacterium tumefaciens* (*47*) (Genbank identifier M59852); BGL_Clotm, β-glucosidase from *Clostridium thermocellum* (*48*) (Genbank identifier X15644); BGL_Ecoli, β-glucosidase from *E. coli* (*35*) (Genbank identifier AE000302); BGL_Aspac, β-glucosidase from *Aspergillus aculeatus* (*49*) (Genbank identifier D64088).

only among the β-*N*-acetylglucosaminidases and not at all within other members of this family (Figure 8). The function of this region can only be guessed at, although from examination of the *Hordeum vulgare* β-glucan exohydrolase structure it appears that in β-*N*-acetylglucosaminidases these residues may come into contact with the 2-acetamido group of the substrate. Basic local alignment sequence tool (BLAST) (*33*) analysis of this region with sequence databases revealed such a putative *N*-acetyl binding site in 15 sequences: 3 cloned family 3 β-*N*-acetylglucosaminidases (*Vibrio furnisii*, *Streptomyces thermoviolaceus*, and *Alteromonas* sp.), 5 open reading frames classified on the basis of global alignments as β-*N*-acetylglucosaminidases (*Escherichia coli*, *Haemophilus influenza*, *Mycobacterium tuberculosis*, *Streptomyces coelicolor*, and *Borrelia burgdorferi*), and 7 other open reading frames (*Zymomonas mobiliz*, *Rickettsia prowazekii*, *Mycobacterium leprae*, *Bacillus subtilus*, *Thermotoga maritima*, *Borrelia burgdorferi*, and *Synechocystic* sp.). We would like to propose this stretch of primary sequence [(KH-(F/I)-PGH-(G/L)-X-X-X-X-D(S/T)H] as a unique identifier for family 3 β-*N*-acetylglucosaminidases which may prove to be of utility given the large amount of information becoming available from genome sequencing work.

CONCLUSIONS

2-Acetamido-2-deoxy-5-fluoro-α-L-idopyranosyl fluoride is hydrolyzed by *Vibrio furnisii* β-*N*-acetylglucosaminidase as a slow Michaelian substrate with kinetic parameters of $K_m = 0.23$ mM and $k_{cat} = 0.0091$ s^{-1}. Incubation of ExoII with 5FIdNAcF permits the observation of a high steady-state population of 5-fluoro-IdNAc-enzyme intermediate, indicating this enzyme operates with a 'normal' double displacement mechanism, unlike β-*N*-acetylglucosaminidases from family 20. This is the first report on the observation of a covalent glycosyl-enzyme intermediate for any glycosyl hydrolase acting on substrates bearing a 2-acetamido group. In this case, the 5-fluoro-IdNAc-enzyme intermediate was sufficiently stable to allow identification of the active site nucleophile using tandem mass spectrometric analysis of proteolytic digests in conjunction with multiple sequence alignments. The assignment of the completely conserved Asp242 as the catalytic nucleophile of ExoII is in agreement with work conducted by Legler (*13*) and Fincher (*12*). Moreover, multiple sequence alignments of members of family 3 revealed a putative *N*-acetyl binding region contained within a short stretch of 14 amino acids [(KH-(F/I)-PGH-(G/L)-X-X-X-X-D(S/T)H] that can be used as a unique identifier for family 3 β-*N*-acetylglucosaminidases.

ACKNOWLEDGMENT

We thank Dr. Saul Roseman for providing the pXE18 plasmid containing the *exoII* gene. Thanks also go to Melanie Mah for excellent technical assistance in the preparation of purified protein. D.J.V. also thanks the Science Council of British Columbia and the Natural Sciences and Engineering Research Council of Canada for scholarships, and C.M. thanks the Swiss National Science Foundation for a fellowship.

REFERENCES

1. Mahuran, D. J. (1991) *Biochim. Biophys. Acta 1096*, 87−94.
2. Scigelova, M., and Crout, D. H. G. (1999) *Enzymol. Microb. Technol. 25*, 3−14.
3. Horsch, M., Mayer, C., Sennhauser, U., and Rast, D. M. (1997) *Phamacol. Ther. 76*, 187−218.
4. Tsujibo, H., Hatano, N., Mikama, T., Hirasawa, A., Miyamoto, K., and Inamori, Y. (1998) *Appl. Environ. Microbiol. 64*, 2920−2924.
5. Henrissat, B., and Bairoch, A. (1996) *Biochem. J. 316*, 695−696.
6. Chitlaru, E., and Roseman, S. (1996) *J. Biol. Chem. 271*, 33433−33439.
7. Legler, G., Sinnott, M. L., and Withers, S. G. (1980) *J. Chem. Soc., Perkins Trans. 2*, 1376−1388.
8. Hrmova, M., and Fincher, G. B. (1998) *Carbohydr. Res. 305*, 209−221.
9. Hrmova, M., Harvey, A. J., Wang, J., Shirley, N. J., Jones, G. P., Stone, B. A., Høj, P. B., and Fincher, G. B. (1996) *J. Biol. Chem. 271*, 5277−5286.
10. Legler, G., Roeser, K.-R., and Illig, H.-K. (1979) *Eur. J. Biochem. 101*, 85−92.
11. Davies, G., Sinnott, M. L., and Withers, S. G. (1998) in *Comprehensive Biological Catalysis*, pp 119−208, Academic Press, London.
12. Varghese, J. N., Hrmova, M., and Fincher, G. B. (1999) *Structure 7*, 179−190.
13. Bause, E., and Legler, G. (1980) *Biochim. Biophys. Acta 626*, 459−465.
14. Gebler, J. C., Aebersold, R., and Withers, S. G. (1992) *J. Biol. Chem. 267*, 11126−11130.
15. He, S., and Withers, S. G. (1997) *J. Biol. Chem. 272*, 24864−24867.
16. Tews, I., Perrakis, A., Oppenheim, A., Dauter, Z., Wilson, K. S., and Vorgias, C. E. (1996) *Nat. Stuct. Biol. 3*, 638−648.
17. Knapp, S., Vocadlo, D., Gao, Z., Kirk, B., Lou, J., and Withers, S. G. (1996) *J. Am. Chem. Soc. 118*, 6804−6805.
18. Drouillard, S., Armand, S., Davies, G. J., Vorgias, C. E., and Henrissat, B. (1997) *Biochem. J. 328*, 945−949.
19. Legler, G., and Bollhagen, R. (1992) *Carbohydr. Res. 233*, 113−123.
20. Kobayashi, S., Kiyosada, T., and Shoda, S.-I. (1997) *Tetrahedron Lett. 38*, 2111−2112.

21. Matsumura, I., and Kirsch, J. F. (1996) *Biochemistry 35*, 1881−1889.
22. Hashimoto, Y., Yamada, K., Motoshima, H., Omura, T., Yamada, H., Yasukochi, T., Miki, T., Ueda, T., and Imoto, T. (1996) *J. Biochem. 119*, 145−150.
23. Vocadlo, D. J., MacKenzie, L. F., He, S., Zeikus, G. J., and Withers, S. G. (1998) *Biochem. J. 335*, 449−455.
24. Wong, A. W., He, S., Grubb, J. H., Sly, W. S., and Withers, S. G. (1998) *J. Biol. Chem. 273*, 34057−34062.
25. Miao, S. C., Ziser, L., Aebersold, R., and Withers, S. G. (1994) *Biochemistry 33*, 7027−7032.
26. Howard, S., He, S., and Withers, S. G. (1998) *J. Biol. Chem. 273*, 2067−2072.
27. McCarter, J. D., and Withers, S. G. (1996) *J. Am. Chem. Soc. 118*, 241−242.
28. McCarter, J. D., and Withers, S. G. (1996) *J. Biol. Chem. 271*, 6889−6894.
29. Withers, S. G., and Aebersold, R. (1995) *Protein Sci. 4*, 361−372.
30. Ballardie, F. W., Capon, B., Dearie, W. M., and Foster, R. L. (1976) *Carbohydr. Res. 49*, 79−92.
31. Leatherbarrow, R. J. (1992) Erithacus Software Ltd., Staines, U.K.
32. Praly, J.-P., Brard, L., Descotes, G., and Toupet, L. (1989) *Tetrahedron 45*, 4141−4152.
33. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol. 215*, 403−410.
34. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res. 22*, 4673−4680.
35. Blattner, F. R., Plunkett, G. I., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collada-Vides, J., Glasner, F. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., and Shao, Y. (1997) *Science 277*, 1453−1474.
36. Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G. G., FitzHugh, W., Fields, C. A., Gocayne, J. D., Scott, J. D., Shirley, R., Liu, L. I., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Fritchman, J. L., Fuhrmann, J. L., Geoghagen, N. S., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O., and Venter, J. C. (1995) *Science 269*, 496−512.
37. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., III, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, S., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, S., Squares, S., Sqares, R., Sulston, J. E., Taylor, K., Whitehead, S., and Barrell, B. G. (1998) *Nature 393*, 537−544.
38. Eiglmeier, K., Honore, N., Woods, S. A., Caudron, B., and Cole, S. T. (1993) *Mol. Microbiol. 7*, 197−206.
39. Andersson, S. G., Zomorodipour, A., Andersson, J. O., Sicheritz-Ponten, T., Alsmark, U. C., Podowski, R. M., Naslund, A. K., Eriksson, A. S., Winkler, H. H., and Kurland, C. G. (1998) *Nature 396*, 133−140.
40. Lee, H. J., and Kang, H. S. (1999) GenBank electronic submission, Sequence analysis of 43D2 fosmid clone of *Zymomonas mobilis* ZM4; Accession Number AF124757.1.
41. Quirk, P. G., Guffanti, A. A., Clejan, S., Cheng, J., and Krulwich, T. A. (1994) *Biochim. Biophys. Acta 1186*(1−2), 27−34.
42. Tsujibo, H., Fujimoto, K., Tanno, H., Miyamoto, K., Imada, C., Okami, Y., and Inamori, Y. (1994) *Gene 146*, 111−115.
43. Redenbach, M., Kieser, H. M., Denapaite, D., Eichner, A., Cullum, J., Kinashi, H., and Hopwood, D. A. (1996) *Mol. Microbiol. 21*(1), 77−96.
44. Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., and Tabata, S. (1996) *DNA Res. 3*, 109−136.
45. Nelson, K. E., Clayton, R. A., Gill, S. R., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D., Nelson, W. C., Ketchum, K. A., McDonald, L., Utterback, T. R., Malek, J. A., Linher, K. D., Garrett, M. M., Stewart, A. M., Cotton, M. D., Pratt, M. S., Phillips, C. A., Richardson, D., Heidelberg, J., Sutton, G. G., Fleischmann, R. D., White, O., Salzberg, S. L., Smith, H. O., Venter, J. C., and Fraser, C. M. (1999) *Nature 399*, 323−329.
46. Fraser, C. M., Casjens, S., Huang, W. M., Sutton, G. G., Clayton, R. A., Lathigra, R., White, O., Ketchum, K. A., Dodson, R., Hickey, E. K., Gwinn, M., Dougherty, B., Tomb, J.-F., Fleischmann, R. D., Richardson, D., Peterson, J., Kerlavage, A. R., Quackenbush, J., Salzberg, S., Hanson, M., van-Vugt, R., Palmer, N., Adams, M. D., Gocayne, J. D., Weidman, J., Utterback, T., Watthey, L., McDonald, L., Artiach, P., Bowman, C., Garland, S., Fujii, C., Cotton, M. D., Horst, K., Roberts, K., Hatch, B., Smith, H. O., and Venter, J. C. (1997) *Nature 390*, 580−586.
47. Castle, L. A., Smith, K. D., and Morris, R. O. (1992) *J. Bacteriol. 174*, 1478−1486.
48. Grabnitz, F., Rucknagel, K. P., Seiss, M., and Staudenbauer, W. L. (1989) *Mol. Gen. Genet. 217*, 70−76.
49. Kawaguchi, T., Enoki, T., Tsurumaki, S., Sumitani, J., Ueda, M., Ooi, T., and Arai, M. (1996) *Gene 173*, 287−288.

BI991958D