Structural and mechanistic insight into the basis of mucopolysaccharidosis IIIB

Elizabeth Ficko-Blean[†], Keith A. Stubbs[‡], Oksana Nemirovsky[†], David J. Vocadlo[‡], and Alisdair B. Boraston^{†§}

[†]Department of Biochemistry and Microbiology, University of Victoria, P.O. Box 3055, Station CSC, Victoria, BC, Canada V8W 3P6; and

[‡]Department of Chemistry, Simon Fraser University, 8888 University Drive, Burnaby, BC, Canada V5A 1S6

Edited by Gregory A. Petsko, Brandeis University, Waltham, MA, and approved March 6, 2008 (received for review December 5, 2007)

Mucopolysaccharidosis III (MPS III) has four forms (A–D) that result from buildup of an improperly degraded glycosaminoglycan in lysosomes. MPS IIIB is attributable to the decreased activity of a lysosomal α -N-acetylglucosaminidase (NAGLU). Here, we describe the structure, catalytic mechanism, and inhibition of CpGH89 from *Clostridium perfringens*, a close bacterial homolog of NAGLU. The structure enables the generation of a homology model of NAGLU, an enzyme that has resisted structural studies despite having been studied for >20 years. This model reveals which mutations giving rise to MPS IIIB map to the active site and which map to regions distant from the active site. The identification of potent inhibitors of CpGH89 and the structures of these inhibitors in complex with the enzyme suggest small-molecule candidates for use as chemical chaperones. These studies therefore illuminate the genetic basis of MPS IIIB, provide a clear biochemical rationale for the necessary sequential action of heparan-degrading enzymes, and open the door to the design and optimization of chemical chaperones for treating MPS IIIB.

Clostridium perfringens | crystal structure | Sanfilippo | hexosaminidase | NAGLU

Mucopolysaccharidosis (MPS) is one group of autosomal recessive lysosomal storage disorders caused by the accumulation of undegraded glycosaminoglycans in lysosomes. MPS III, or Sanfilippo syndrome, is characterized by the impaired function of one of four lysosomal enzymes involved in the sequential degradation of heparan sulfate. MPS IIIA is caused by the reduced activity of heparan sulfate sulfamidase, MPS IIIB is caused by the reduced activity of an α -N-acetylglucosaminidase (NAGLU), MPS IIIC is caused by the reduced activity of acetyl-CoA: α -glucosaminide-N-acetyltransferase, and MPS IIID is caused by the reduced activity of N-acetyl-D-glucosamine 6-sulfatase (1). This tragic disease initially manifests as aggressiveness and hyperactivity, later progressing to mental retardation and, ultimately, CNS degeneration. There is currently no cure or effective treatment for this disease, and patients generally only survive into early adulthood.

NAGLU, a family 89 α -N-acetylglucosaminidase (www.cazy.org), is responsible for cleaving the glycosidic bond found in the backbone of heparan and thus plays a key role in heparan recycling. More than 100 different mutations in the *naglu* gene have been associated with the MPS IIIB phenotype, and biochemical studies have confirmed the deleterious effects of many of these mutations (1–9). Despite having been cloned >10 years ago, no structural or mechanistic data for NAGLU have been obtained, hindering the development of potential therapeutic strategies to treat patients suffering from MPS IIIB. To gain insight into the structure and function of this medically important enzyme, we cloned CpGH89, a bacterial family 89 glycoside hydrolase from *Clostridium perfringens*, to use as a model for NAGLU because they share a significant level of amino acid sequence identity ($\approx 30\%$). Here, we detail the structure of this enzyme and a homology model of the human enzyme. These structures enable us to map the positions of the known mutations causing MPS IIIB onto the structure. Further, we identify potent

inhibitors of this enzyme, obtain structures of the enzyme in complex with these inhibitors, and elucidate the catalytic mechanism of this enzyme. These studies should significantly facilitate the rational design of chemical chaperones that could be used to treat this deleterious genetic illness.

Results and Discussion

Activity and Structure of CpGH89. The genome of *C. perfringens*, ATCC 13124, contains an ORF (CPF.0859) that encodes a putative 2,095-aa protein. This protein is classified into glycoside hydrolase family 89 and shares significant amino acid sequence similarity with the related family 89 glycoside hydrolase from *Homo sapiens* known as NAGLU. After molecular dissection of the modular structure of the *C. perfringens* GH89, we were able to overproduce a truncated construct in *Escherichia coli* comprising residues 26–916 and purify the resulting polypeptide, which, for simplicity, is referred to here as CpGH89.

CpGH89 was screened for catalytic activity against a panel of synthetic substrates, and activity was detected only toward *para*-nitrophenyl α -N-acetyl-D-glucosaminide (pNP- α -GlcNAc), consistent with the known activity of NAGLU. This substrate produced a bell-shaped activity/pH profile, suggesting that two residues with kinetic pK_a values of ≈ 6.3 and ≈ 8.0 must be maintained in appropriate ionization states [see [supporting information \(SI\) Fig. S1](#)]. Further kinetic characterization at the optimum pH of 7.3 was carried out to determine the values of K_m (1.1 ± 0.1 mM), k_{cat} ($2.6 \pm 0.1 \times 10^{-1} s^{-1}$), and k_{cat}/K_m ($2.5 \pm 0.3 \times 10^{-1} mM^{-1}s^{-1}$) of the enzyme toward pNP- α -GlcNAc. The substrate dNP- α -GlcNAc (2,4-dinitrophenyl α -N-acetyl-D-glucosaminide) (10) is turned over significantly more quickly with corresponding values for K_m , k_{cat} , and k_{cat}/K_m of $7.4 \pm 1.1 \times 10^{-1}$ mM, $8.5 \pm 0.4 s^{-1}$, and $11.6 \pm 2.3 mM^{-1}s^{-1}$, respectively.

The structure of CpGH89 was solved by SIR with a HoCl₃ derivative (see *Materials and Methods*). A single molecule of CpGH89, which is composed of four distinct domains, was present in the asymmetric unit. The N-terminal domain (residues 26–155) is a putative family 32 carbohydrate-binding module (CBM) having the typical β -sandwich fold adopted by many CBMs (11). The catalytic region comprises a small α/β domain (residues 170–280), an elaborated (α/β)₈ barrel (residues 280–620), and an all α -helical domain that packs against the first three domains (residues 621–916) (Fig. 1A). The elaborations on the

Author contributions: E.F.-B. and A.B.B. designed research; E.F.-B., K.A.S., O.N., D.J.V., and A.B.B. performed research; E.F.-B., K.A.S., D.J.V., and A.B.B. analyzed data; and E.F.-B., D.J.V., and A.B.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org [PDB ID codes 2VCC (for the native structure), 2VCA (for the GLcNAc complex), 2VC9 (for the 2AcDNJ complex), and 2VCB (for the PUGNAc complex)].

[§]To whom correspondence should be addressed. E-mail: boraston@uvic.ca.

This article contains supporting information online at www.pnas.org/cgi/content/full/0711491105/DCSupplemental.

© 2008 by The National Academy of Sciences of the USA

Fig. 2. Structural location of naturally occurring mutations in NAGLU. (A) A cartoon representation of the homology model of NAGLU showing its overall fold. The coloring of the domains is as for CpGH89 in Fig. 1 A. (B) A structural overlay of the active site of CpGH89 and NAGLU. The catalytic residues are labeled. (C) A view of the ribbon trace of the NAGLU model with the structural location of the known mutations that lead to MPS IIIB. Sites of mutations are shown as spheres. Blue coloring indicates active site residues, whereas red coloring indicates nonactive site residues.

bound substrate, suggesting that this group is the general acid/base catalytic residue. Given that the enzyme is a retaining α -glycosidase, there is no precedent, and thus it does not appear likely that the 2-acetamido group could be productively involved in catalysis in a direct manner as a catalytic group. Indeed, there is no structural evidence to imply participation of the *N*-acetyl group. Therefore, it is most likely that CpGH89 and, by extension, all members of GH89 use a double-displacement retaining mechanism (see Fig. S2) (14, 15). In CpGH89, we propose Glu-483 as the general acid/base catalytic residue and Glu-601 as the catalytic nucleophile, residues that are conserved across all known members of the GH89 family (data not shown).

Insights into MPS IIIB. There is substantial genetic information describing mutations in the *naglu* gene that give rise to MPS IIIB and some biochemical characterization of these mutant NAGLU enzymes (7, 16). Further progress, however, has been hindered because no structure of NAGLU or any homologue is available that can be used to link this genetic information to the enzyme structure. CpGH89 has $\approx 30\%$ overall amino acid sequence identity to NAGLU (see Fig. S3), allowing for the construction of a complete model of NAGLU. This model shows the overall structural conservation (Fig. 2A), whereas an examination of the active site residues of CpGH89 overlaid with the predicted NAGLU active site shows that the architecture of the active sites is almost entirely conserved (Fig. 2B and Fig. S3). The only residue that is not conserved within the active site is His-482 in CpGH89, which is substituted by an asparagine in NAGLU (Fig. 2B). Accordingly, CpGH89 acts as an excellent model of NAGLU in terms of detailed substrate recognition and catalysis, as well as overall fold.

CpGH89 and, by extension, NAGLU are obligate *exo*-acting enzymes with active site architectures that appear to be unable to accommodate substrates in which the saccharides are modified. For example, processing of a sulfated terminal *N*-acetyl-D-glucosamine residue would not be possible, a prediction that, to the best of our knowledge, has not been experimentally demonstrated (17). This strict specificity necessitates a route to generate nonreducing terminal α -*N*-acetylglucosamine residues on heparan. Thus, the degradation and recycling of heparan in the human body occurs via the sequential action of three additional enzymes: heparan sulfate sulfamidase, acetylCoA: α -glucosaminidase-*N*-acetyltransferase, and *N*-acetylglucosamine 6-sulfatase (17). These enzymes are responsible for generating nonreducing terminal *N*-acetyl-D-glucosamine residues, and deficiency in the activity of each one of these three enzymes results in MPS IIIA, IIIC, or IIID, respectively. Given the strict structural requirement of the α -*N*-acetylglucosaminidase for unmodified nonreducing α -linked *N*-acetyl-D-glucosamine residues, it becomes clear how a deficiency in any of these four enzymes could stall heparan sulfate depolymerization.

The MPS IIIB phenotype is associated with numerous missense, nonsense, and deletion mutations, with the missense mutations far outnumbering the rest. Mapping the positions of these known missense mutations onto the model of NAGLU is revealing. These mutations are randomly scattered throughout the protein, with only four occurring at residues within the active site (Fig. 2C and Fig. S3). Remarkably, there are no mutations of the catalytic residues perhaps because these mutations are so deleterious as to be prenatally lethal. By analogy to other lysosomal storage disorders, such as Tay-Sachs disease (18), these mutations likely influence the protein by reducing its stability and resulting in less functional enzyme reaching the lysosome. Indeed, Yogalingam *et al.* (16) recently demonstrated that mutations in NAGLU do indeed affect the stability and transiting of the enzyme through the secretory pathway. It is a well described phenomenon that many point mutations in proteins cause changes in folding that abolish proper trafficking (19, 20). Ironically, in many cases, these physiologically disastrous mutations do not destroy the catalytic activity and may only cause slight folding anomalies (21).

One proposed approach to treating these diseases that stem from decreased protein stability involves the use of small molecules (22). Compounds called chemical chaperones have been described: They bind to the newly biosynthesized protein and thereby increase its stability. These “chaperoned” mutant proteins are more stable and can therefore be transported out of the endoplasmic reticulum, avoiding intracellular degradation (22–28). With this point in mind, we sought potential inhibitors that may aid in identifying, designing, and developing efficient chemical chaperones to treat MPS IIIB.

GH89 Inhibitors: Potential Chemical Chaperones? We assessed the binding of three known *N*-acetylglucosaminidase inhibitors: 2-acetamido-1,2-dideoxynojirimycin (2AcDNJ) (29), *O*-(2-acetamido-2-deoxy-D-glucopyranosylidene)amino-*N*-phenylcarbamate (PUGNAc) (30), and 6-acetamido-6-deoxycastanospermine (6AcCAS) (31) to CpGH89 by isothermal titration calorimetry (ITC) and kinetic methods (Fig. 3A and B; see Fig. S4 for 6AcCAS data). The K_i for PUGNAc was determined by Dixon plot analysis to be $6.2 \pm 0.1 \mu\text{M}$ (Fig. 3A Inset), which agreed with the dissociation constant (K_d) of $5.5 \pm 0.3 \mu\text{M}$ determined by ITC (Fig. 3A). The K_i - and ITC-determined K_d values for 6AcCAS were likewise in good agreement at $92 \pm 5 \text{ nM}$ and $90 \pm 20 \text{ nM}$, respectively (Fig. S4). Because the artificial substrates available for CpGH89 are turned over quite slowly, high concentrations of enzyme (67 nM) were required for kinetic assays, and this requirement defined the lower limit at which K_i values could be measured. However, the K_d for the tight binding inhibitor 2AcDNJ was determined by ITC to be $3.6 \pm 0.3 \text{ nM}$, indicating $\approx 1,000$ -fold tighter binding of this inhibitor to CpGH89 relative to PUGNAc. Further, a pattern of competitive

Fig. 3. Inhibitor binding by CpGH89. (A and B) Isotherms of CpGH89 binding to PUGNac (A) and 2AcDNJ (B) produced by ITC. (Upper) Raw heat measurements. (Lower) Integrated heats. The solid lines show the fit of a one site-binding model to the data. (A Inset) Dixon plot analysis of CpGH89 inhibition by PUGNac. The intersection point on the graph corresponding to the K_i (absolute value of the X value at the intersection) is indicated by an arrow. (C and D) Active site representations are shown for PUGNac (C) and 2AcDNJ (D). The blue mesh shows the maximum likelihood (43)/ σ_a (49)-weighted electron density maps contoured at $0.23 \text{ e}^-/\text{\AA}^3$ and $0.22 \text{ e}^-/\text{\AA}^3$ for 2AcDNJ and PUGNac, respectively. Key active site residues, including the putative catalytic residues Glu-483 and Glu-601, are shown in stick representation and colored gray. Ligands are shown in green stick representation. Putative hydrogen bonds between the protein and ligand identified by using the criteria of proper geometry and a distance cutoff of $3.2 \text{ \AA}$ are shown as dotted magenta lines. (E and F) Schematics showing the interactions within the active site of CpGH89 with PUGNac (E) and 2AcDNJ (F). A distance of $3.2 \text{ \AA}$ was used as the cutoff for significant hydrogen bonds. Water molecules are shown as shaded spheres. Protons on the amino acids are omitted for clarity.

inhibition is observed for PUGNac (Fig. 3A Inset) and 6AcCAS (Fig. S4B); because PUGNac and 2AcDNJ inhibitors bind to the same location (Fig. 3), it seems most likely that all three inhibitors bind in a competitive manner.

Diffraction data from crystals soaked with 2AcDNJ or PUGNac yielded excellent electron density for the sugar rings in the active site pocket (Fig. 3C and D). However, the phenyl ring substituent of PUGNac was so disordered as to render its electron density undetectable (Fig. 3C). Numerous polar interactions with the equatorial substituents at C3–C5 likely confer high specificity for the gluco-configured sugars (Fig. 3E and F). The basis for recognition of the 2-acetamido group also is clearly revealed in these complexes. A tight nonpolar pocket formed by the side chains

of Trp-366, Trp-685, Tyr-825, and Tyr-305 accommodates the *N*-acetyl group (Fig. 3C and D). Selectivity for the *N*-acetyl group of substrate and these inhibitors is likely imparted by the close van der Waals interactions between these aromatic residues and the substituent, as well as a favorably oriented $2.6\text{-\AA}$ hydrogen bond between the carbonyl and Tyr-305 (Fig. 3C–F). The structural basis for the comparatively weak binding of PUGNac is unclear, although it may result from difficulties of the active site in accommodating the phenylcarbamate group.

The structure of CpGH89 and the model of NAGLU clearly reveal that mutations leading to MPS IIIB are scattered throughout the tertiary structure of the enzyme, in keeping with the notion that the instability of mutated NAGLU can result in the MPS IIIB phenotype (16). Mutant lysosomal enzymes that are unstable, including point mutants of NAGLU, can be retained in the endoplasmic reticulum by quality control mechanisms and thereby fail to traffic to lysosomes (16, 20, 27). When used at low concentrations, inhibitors of lysosomal glycosidases can be used as “chemical chaperones” to bind to and stabilize these mutant enzymes, thereby facilitating their maturation and passage to lysosomes (23–28). Once in the lysosome, these mutant enzymes have catalytic activities that support normal cellular functioning (23–28). Accordingly, a chemical chaperone approach to treating MPS IIIB is likely viable, and our results indicate that 2AcDNJ and 6AcCAS are effective inhibitors that may provide an initial lead for developing such chemical chaperones. Indeed, both of these inhibitors also have been shown to inhibit NAGLU (32) *in vitro* with nanomolar potencies (K_i for 2AcDNJ = 450 nM and K_i for 6AcCAS = 87 nM), in keeping with the results we observe here for CpGH89. These data, in combination with the near-complete conservation of active site residues, strongly support the validity of using the structure of CpGH89 as a guide for inhibitor design. However, the use of these particular compounds may be limited because they are known to be potent inhibitors of several other human enzymes. Nevertheless, there remains the potential to generate more selective inhibitors (28, 33, 34).

In summary, the data presented here reveal the structural basis underlying the specificity of NAGLU for terminal nonreducing *N*-acetyl-D-glucosamine residues as well as its place within the glycosaminoglycan catabolic pathway. The extensive stereochemically specific interactions with the sugar hydroxyl groups, coupled with the sock-shaped architecture of the active site, strictly defines selectivity for the α -configured glycosidic linkage of unornamented terminal *N*-acetyl-D-glucosamine residues. Further, the collective structural, mechanistic, and inhibition data presented here provide a clear template that will facilitate the design of effective chemical chaperones that specifically target NAGLU.

Materials and Methods

All reagents are from Sigma–Aldrich unless otherwise stated.

Cloning, Protein Production, and Purification. The gene fragment comprising nucleotides 76–2748 of the *cpg89* gene (locus tag CPF.0859), which encodes a polypeptide with a putative N-terminal family 32 carbohydrate-binding module followed by the catalytic module, was PCR amplified from *C. perfringens* ATCC 13124 genomic DNA (Sigma) and cloned into pET-28a(+) (Novagen) via 5' and 3' NheI and XhoI restriction endonuclease sites to generate pCBM1GH89 by using standard PCR-cloning procedures. The polypeptide encoded by this recombinant gene comprises an N-terminal H6 tag, followed by a thrombin protease cleavage site fused to the N terminus of the CBM32/GH89 modules, and is referred to as CpGH89. “Mega-primer” PCR site-directed mutagenesis procedures were used to introduce the E483A and E601A substitutions. The DNA sequence fidelity of all constructs was verified by bidirectional sequencing. CpGH89 was produced in BL21 star (DE3) *E. coli* cells (Invitrogen) harboring the pCBM1GH89 and purified by immobilized metal affinity chromatography by using methods described previously (35).

Enzyme Kinetics. All steady-state kinetic studies were carried out at 25°C in a Cary/Varian 300 Bio UV-Visible Spectrophotometer. The pH dependence of CpGH89 was determined by using 133 nM enzyme and a pNP- α -GlcNAc concentration of 1 mM. Then, 0.1 M sodium acetate was used to buffer within the pH range 4.6–6.0, 0.1 M sodium phosphate for the pH range 6.0–8.0, and 0.1 M glycine-glycine for the pH range 8.0–9.8. At the transition between buffers, one or two pHs were overlapped to control for the effects of the buffer on the enzyme activity. Standard-reaction mixtures for the determination of kinetic constants were done at the optimal pH (0.1 M sodium phosphate, pH 7.3) in 750- μ l volumes containing 15.2 nM CpGH89 or 390 nM Glu483Ala, 10% BSA, and 0–2.5 mM pNP- α -GlcNAc. The 4-nitrophenolate production was measured at 400 nm over a 5-min time period. Rate of release was determined by linear regression over the linear period of release on the curve. All experiments were performed in triplicate. Michealis–Menten parameters were determined by nonlinear curve fitting by using Origin7. Similarly, determination of the kinetic constants for dNP- α -GlcNAc, synthesized as described previously (10), was done in 0.1 M sodium phosphate (pH 7.3) in 750- μ l volumes containing 240 nM CpGH89, 10% BSA, and 0–3.7 mM dNP- α -GlcNAc.

NMR Experiments. ^1H NMR spectroscopy (600 MHz Bruker AMX spectrometer) was used to identify the products of the enzyme-catalyzed reaction. The reaction was carried out in ≈ 0.6 ml [25°C, 0.1 M Gly-Gly (pH 7.7)] containing 3 mM dNP- α -GlcNAc. Initiation of the reaction was done by the addition of 30 μ l of a 70 μ M stock of the enzyme CpGH89. The hydrolysis of the dNP- α -GlcNAc was monitored until the reaction reached equilibrium. Spectral data were collected with 64,000 data points (64 k) over a spectral width of 7,211 Hz (12.01 ppm), with a relaxation delay of 1 s and 32 scans. An initial spectrum (referred to as time 0) containing substrate and buffer was acquired before the addition of enzyme.

Inhibitor-Binding Studies. The K_i value for CpGH89 binding to PUGNAc and 6AcCAS was determined by linear regression of Dixon plots for inhibitor concentrations ranging from 1/3 to 3 times K_i .

ITC was performed with a VP-ITC (MicroCal) as described previously (35). CpGH89 was dialyzed extensively against 0.1 M sodium phosphate (pH 7.3). Solid inhibitor was resuspended in buffer saved from the dialysis. Twenty-five 10- μ l injections of 80–100 μ M 2AcDNJ or 6AcCAS were titrated into 6–8 μ M protein. Similarly, 25 5- μ l injections of 1,700 μ M PUGNAc were titrated into 50 μ M CpGH89. The C values (36) were maintained between 10 and 1,000. Binding stoichiometry, enthalpy, and the equilibrium-association constant were determined by fitting the heat of dilution-corrected data to a one site-binding model. All values reported were averages, and standard deviations of experiments were performed in triplicate.

Crystallization, Data Collection, Structure Solution, and Refinement. CpGH89 in 20 mM Tris-HCl (pH 8.0) was digested with thrombin overnight at room temperature. Digested protein was applied to an S-200 Sephacryl high-resolution size-exclusion column, and fractions containing pure CpGH89 (as assessed by SDS/PAGE) were pooled and concentrated. Crystal trials were set up by using the hanging drop vapor diffusion method with both thrombin-treated and untreated CpGH89. Diffraction-quality crystals were obtained in 2.0 M $(\text{NH}_4)_2\text{SO}_4$ with 3% glycerol (no buffer). Both thrombin-treated and untreated CpGH89 yielded crystals.

The crystals were cryoprotected by a 10-s immersion in a solution contain-

ing 2 M ammonium sulfate and 30% glycerol, followed by freezing in a nitrogen Cryostream. Diffraction data were collected with a Rigaku R-Axis IV++ area detector coupled to an MM-002 x-ray generator with Osmic “blue” optics and an Oxford Cryostream 700. Data were processed with Crystal Clear/d*trek (37).

A holmium (III) derivative was obtained by soaking native CpGH89 crystals in mother liquor supplemented with 1 mM HoCl_3 for 3 days. This crystal was cryoprotected as for native crystals. The heavy atom substructure of two holmium sites was determined with the programs Shelx/C/D (38) by using the isomorphous differences between the derivative and a native dataset using data to a resolution of 4.5 Å. Refinement of heavy atom parameters and initial phasing to a resolution of 3.0 Å was performed with SHARP (39). Density modification with DM (40) was used to improve and extend the phases to 2.4 Å and ultimately yielded interpretable electron density maps. ARP/wARP (41) was able to build the backbone of $\approx 40\%$ of the molecule. Manual model building in COOT (42) allowed the backbone of the initial model to be corrected and extended to $\approx 65\%$ completeness with some side chains. This initial model was used as input into ARP/wARP, which was then able to build a model of $\approx 90\%$ completeness with $\approx 80\%$ of the side chains correctly docked. The model was manually completed in COOT, followed by refinement with REFMAC (43).

Inhibitor complexes (2-acetamido-1,2-dideoxynojirimycin and PUGNAc) and product complexes were obtained by adding excess inhibitor or substrate (pNP- α -GlcNAc) to mother liquor and soaking for a period of 1 h to up to 1 week. Crystals were cryoprotected as for the native data. These complexes were solved by using the native CpGH89 model as a starting point. Ligands were built into the models by using COOT. When necessary, appropriate REFMAC library files containing the restraints for the ligands were obtained by using the PRODRG Server (44).

In all datasets, the same 5% of the observations were flagged as “free” (45) and used to monitor refinement procedures. Water molecules were added by using the REFMAC implementation of ARP/wARP and inspected visually before deposition. See Table S1 for all data collection, refinement, and model validation statistics.

Homology Modeling. A homology model of NAGLU was constructed from the CpGH89-product complex coordinates by using methods essentially identical to those described previously (46) with DeepView and the Swiss-Model Automated Comparative Protein Modeling Server (47). The resulting model was subjected to 10 rounds of model idealization with REFMAC, and the output model was manually examined residue by residue. The model was validated by tools within COOT and PROCHECK (48).

Accession Codes. Coordinates and structure factors have been deposited with the PDB codes of 2VCC for the native structure, 2VCA for the GlcNAc complex, 2VC9 for the 2AcDNJ complex, and 2VCB for the PUGNAc complex.

ACKNOWLEDGMENTS. This work was supported by Canadian Institutes of Health Research Grants MOP 68913, ABB, and NIP79924 (to D.J.V.), Michael Smith Foundation for Health Research and Natural Sciences and Engineering Research Council of Canada Doctoral Fellowships (to E.F.-B.), a Canada Research Chair award in Chemical Glycobiology and a Michael Smith Foundation for Health Research Scholar award (to D.J.V.), and a Canada Research Chair award in Molecular Interactions and a Michael Smith Foundation for Health Research Scholar award (to A.B.B.).

- Yogalingam G, Hopwood JJ (2001) Molecular genetics of mucopolysaccharidosis type IIIA and IIIB: Diagnostic, clinical, and biological implications. *Hum Mutat* 18:264–281.
- Beesley CE, Jackson M, Young EP, Vellodi A, Winchester BG (2005) Molecular defects in Sanfilippo syndrome type B (mucopolysaccharidosis IIIB). *J Inher Metab Dis* 28:759–767.
- Beesley CE, Young EP, Vellodi A, Winchester BG (1998) Identification of 12 novel mutations in the alpha-N-acetylglucosaminidase gene in 14 patients with Sanfilippo syndrome type B (mucopolysaccharidosis IIIB). *J Med Genet* 35:910–914.
- Emre S, et al. (2002) Sanfilippo syndrome in Turkey: Identification of novel mutations in subtypes A and B. *Hum Mutat* 19:184–185.
- Tanaka A, Kimura M, Lan HT, Takaura N, Yamano T (2002) Molecular analysis of the alpha-N-acetylglucosaminidase gene in seven Japanese patients from six unrelated families with mucopolysaccharidosis IIIB (Sanfilippo type B), including two novel mutations. *J Hum Genet* 47:484–487.
- Tessitore A, et al. (2000) Molecular defects in the alpha-N-acetylglucosaminidase gene in Italian Sanfilippo type B patients. *Hum Genet* 107:568–576.
- Schmidtchen A, et al. (1998) NAGLU mutations underlying Sanfilippo syndrome type B. *Am J Hum Genet* 62:64–69.
- Bunge S, et al. (1999) Mucopolysaccharidosis type IIIB (Sanfilippo B): Identification of 18 novel alpha-N-acetylglucosaminidase gene mutations. *J Med Genet* 36:28–31.
- Weber B, et al. (1999) Sanfilippo type B syndrome (mucopolysaccharidosis IIIB): Allelic heterogeneity corresponds to the wide spectrum of clinical phenotypes. *Eur J Hum Genet* 7:34–44.
- Chen HM, Withers SG (2007) Facile synthesis of 2,4-dinitrophenyl alpha-D-glycopyranosides as chromogenic substrates for alpha-glycosidases. *ChemBiochem* 8:719–722.
- Boraston AB, Bolam DN, Gilbert HJ, Davies GJ (2004) Carbohydrate-binding modules: Fine tuning polysaccharide recognition. *Biochem J* 382:769–782.
- Zechel DL, et al. (2003) Mechanism, mutagenesis, and chemical rescue of a beta-mannosidase from *cellulomonas fimi*. *Biochemistry* 42:7195–7204.
- Vallmitjana M, et al. (2001) Mechanism of the family 1 beta-glucosidase from *Streptomyces* sp: Catalytic residues and kinetic studies. *Biochemistry* 40:5975–5982.
- Henrissat B, Davies G (1997) Structural and sequence-based classification of glycoside hydrolases. *Curr Opin Struct Biol* 7:637–644.
- Davies G, Henrissat B (1995) Structures and mechanisms of glycosyl hydrolases. *Structure* 3:853–859.
- Yogalingam G, Weber B, Meehan J, Rogers J, Hopwood JJ (2000) Mucopolysaccharidosis type IIIB: Characterisation and expression of wild-type and mutant recombinant alpha-N-acetylglucosaminidase and relationship with Sanfilippo phenotype in an attenuated patient. *Biochim Biophys Acta* 1502:415–425.

17. Hopwood JJ, Morris CP (1990) The mucopolysaccharidoses. Diagnosis, molecular genetics and treatment. *Mol Biol Med* 7:381–404.
18. Mahuran DJ (1999) Biochemical consequences of mutations causing the GM2 gangliosidosis. *Biochim Biophys Acta* 1455:105–138.
19. Bychkova VE, Ptitsyn OB (1995) Folding intermediates are involved in genetic diseases? *FEBS Lett* 359:6–8.
20. Ellgaard L, Molinari M, Helenius A (1999) Setting the standards: Quality control in the secretory pathway. *Science* 286:1882–1888.
21. Welch WJ, Brown CR (1996) Influence of molecular and chemical chaperones on protein folding. *Cell Stress Chaperones* 1:109–115.
22. Amaral MD (2006) Therapy through chaperones: Sense or antisense? Cystic fibrosis as a model disease. *J Inherited Metabol Dis* 29:477–487.
23. Lieberman RL, et al. (2007) Structure of acid beta-glucosidase with pharmacological chaperone provides insight into Gaucher disease. *Nat Chem Biol* 3:101–107.
24. Steet RA, et al. (2006) The iminosugar isofagomine increases the activity of N370S mutant acid beta-glucosidase in Gaucher fibroblasts by several mechanisms. *Proc Natl Acad Sci USA* 103:13813–13818.
25. Tropak MB, Reid SP, Guiral M, Withers SG, Mahuran D (2004) Pharmacological enhancement of beta-hexosaminidase activity in fibroblasts from adult Tay-Sachs and Sandhoff patients. *J Biol Chem* 279:13478–13487.
26. Sawkar AR, et al. (2002) Chemical chaperones increase the cellular activity of N370S beta-glucosidase: A therapeutic strategy for Gaucher disease. *Proc Natl Acad Sci USA* 99:15428–15433.
27. Fan JQ, Ishii S, Asano N, Suzuki Y (1999) Accelerated transport and maturation of lysosomal alpha-galactosidase A in Fabry lymphoblasts by an enzyme inhibitor. *Nat Med* 5:112–115.
28. Asano N, Kato A, Watson AA (2001) Therapeutic applications of sugar-mimicking glycosidase inhibitors. *Mini Rev Med Chem* 1:145–154.
29. Horsch M, Hoesch L, Vasella A, Rast DM (1991) N-acetylglucosaminono-1,5-lactone oxime and the corresponding (phenylcarbamoyl)oxime—novel and potent inhibitors of beta-N-acetylglucosaminidase. *Euro J Biochem* 197:815–818.
30. Beer D, Maloisel JL, Rast DM, Vasella A (1990) Synthesis of 2-acetamido-2-deoxy-D-gluconhydroxymolactone-derived and chitobionhydroxymolactone-derived N-phenylcarbamates, potential inhibitors of beta-N-acetylglucosaminidase. *Helv Chim Acta* 73:1918–1922.
31. Liu PS, Kang MS, Sunkara PS (1991) A potent inhibitor of beta-N-acetylglucosaminidase: 6-acetamido-6-deoxycastanospermine. *Tetrahedron Lett* 32:719–720.
32. Zhao KW, Neufeld EF (2000) Purification and characterization of recombinant human alpha-N-acetylglucosaminidase secreted by Chinese hamster ovary cells. *Protein Expr Purif* 19:202–211.
33. Macauley MS, Whitworth GE, Debowski AW, Chin D, Vocadlo DJ (2005) O-GlcNAcase uses substrate-assisted catalysis: Kinetic analysis and development of highly selective mechanism-inspired inhibitors. *J Biol Chem* 280:25313–25322.
34. Ogawa S, Kanto M, Suzuki Y (2007) Development and medical application of unsaturated carbaglycosylamine glycosidase inhibitors. *Mini Rev Med Chem* 7:679–691.
35. Ficko-Blean E, Boraston AB (2006) The interaction of a carbohydrate-binding module from a Clostridium perfringens N-acetyl-beta-hexosaminidase with its carbohydrate receptor. *J Biol Chem* 281:37748–37757.
36. Wiseman T, Williston S, Brandts JF, Lin LN (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal Biochem* 179:131–137.
37. Pflugrath JW (1999) The finer things in X-ray diffraction data collection. *Acta Crystallogr D* 55:1718–1725.
38. Schneider TR, Sheldrick GM (2002) Substructure solution with SHELXD. *Acta Crystallogr D* 58:1772–1779.
39. Bricogne G, Vonrhein C, Flensburg C, Schiltz M, Paciorek W (2003) Generation, representation and flow of phase information in structure determination: Recent developments in and around SHARP 2.0. *Acta Crystallogr D* 59:2023–2030.
40. Cowtan K, Main P (1998) Miscellaneous algorithms for density modification. *Acta Crystallogr D* 54:487–493.
41. Perrakis A, Morris R, Lamzin VS (1999) Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* 6:458–463.
42. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D* 60:2126–2132.
43. Murshudov GN, Vagin AA, Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D* 53:240–255.
44. van Aalten DM, et al. (1996) PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. *J Comput Aided Mol Des* 10:255–262.
45. Brunger AT (1993) Assessment of phase accuracy by cross validation: The free R value. Methods and applications. *Acta Crystallogr D* 49:24–36.
46. Rempel BP, Clarke LA, Withers SG (2005) A homology model for human alpha-l-iduronidase: Insights into human disease. *Mol Genet Metab* 85:28–37.
47. Schwede T, Kopp J, Guex N, Peitsch MC (2003) SWISS-MODEL: An automated protein homology-modeling server. *Nucleic Acids Res* 31:3381–3385.
48. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) Procheck: A program to check the stereochemical quality of protein structures. *J Appl Crystallogr* 26:283–291.
49. Read RJ (1986) Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Crystallogr A* 42:140–149.