Crystallographic Evidence for Substrate-assisted Catalysis in a Bacterial β-Hexosaminidase*

Received for publication, December 8, 2000, and in revised form, December 21, 2000
Published, JBC Papers in Press, December 21, 2000, DOI 10.1074/jbc.M011067200

Brian L. Mark‡‡, David J. Vocadlo¶¶, Spencer Knapp**, Barbara L. Triggs-Raine‡‡, Stephen G. Withers‡‡, and Michael N. G. James$$$ From the %Medical Research Group in Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada, the **Department of Chemistry, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada, the $$$Department of Chemistry, Rutgers University, New Brunswick, New Jersey 08903, and the ‡‡Department of Biochemistry and Medical Genetics, University of Manitoba, Winnipeg, Manitoba R3E 0W3, Canada

β-Hexosaminidase, a family 20 glycosyl hydrolase, catalyzes the removal of β-1,4-linked N-acetyhexosamine residues from oligosaccharides and their conjugates. Heritable deficiency of this enzyme results in various forms of GalNAc-β(1,4)-(N-acetylneuraminic acid (2,3)]-Gal-β(1,4)-Glc-ceramide gangliosidosis, including Tay-Sachs disease. We have determined the x-ray crystal structure of a β-hexosaminidase from Streptomyces plicatus to 2.2 Å resolution (Protein Data Bank code 1HP4). β-Hexosaminidases are believed to use a substrate-assisted catalytic mechanism that generates a cyclic oxazolinium ion intermediate. We have solved and refined a complex between the cyclic intermediate analogue N-acetylglucosamine-thiazoline and β-hexosaminidase from S. plicatus to 2.1 Å resolution (Protein Data Bank code 1HP5). Difference Fourier analysis revealed the pyranose ring of N-acetylglucosamine-thiazoline bound in the enzyme active site with a conformation close to that of a C1 chair. A tryptophan-lined hydrophobic pocket envelops the thiazoline ring, protecting it from solvolysis at the iminium ion carbon. Within this pocket, Tyr390 and Asp713 appear important for positioning the 2-acetamido group of the substrate for nucleophilic attack at the anomeric center and for dispersing the positive charge distributed into the oxazolinium ring upon cyclization. This complex provides decisive structural evidence for substrate-assisted catalysis and the formation of a covalent, cyclic intermediate in family 20 β-hexosaminidases.

Carbohydrates are involved in many diverse biological functions including cell structural integrity, energy storage, patho-

gen defense and invasion mechanisms, viral penetration, and cellular signaling. Therefore, a large number of enzymes dedicated to carbohydrate metabolism have evolved. Enzymes specifically responsible for carbohydrate catabolism are collectively referred to as glycosyl hydrolases and have been classified into 77 families based on amino acid sequence similarity (1–3). Three-dimensional structures are known for representatives of 30 of the families. Although there are differences in chain length and domain structure between proteins of a single family, all proteins of a family hydrolyze the glycosidic bond with the same stereocchemical outcome (4).

Family 20 includes the β-N-acetyhexosaminidases (β-hexosaminidases)1 (EC 3.2.1.52), enzymes that catalyze the removal of terminal β-1,4 linked N-acetyhexosamine residues from the nonreducing ends of oligosaccharides and their conjugates. In humans, there are two major β-hexosaminidase isoforms: HexA and HexB. HexA is a heterodimer of subunits α (encoded by HEXA) and β (encoded by HEXB), whereas HexB is a homodimer of β subunits. HexA is essential for degrading GalNAc-β(1,4)-(N-acetylneuraminic acid (2,3)]-Gal-β(1,4)-Glc-ceramide ganglioside; the biological importance of HexA activity is illustrated by the fatal neurodegenerative disorders that result from its heritable deficiency (5). Mutations in HEXA or HEXB cause Tay-Sachs and Sandhoff disease, respectively. These genetic diseases have made the human β-hexosaminidase isoforms the subject of much research. A substantial amount of genetic and biochemical information is available for these isoenzymes (5), but detailed information about their catalytic mechanism is limited. Mechanistic studies have been primarily limited by the difficulties in producing sufficient amounts of recombinant enzyme needed for kinetic analysis (6, 7); however, recent improvements in expression and purification procedures have allowed more accurate kinetic measurements to be made (8). Crystals of human HexB have been grown (9); however, attempts at solving its three-dimensional structure have not been successful. Nonetheless, much insight into the mechanism of human HexA and HexB has been provided by structural and functional studies carried out on related family 20 glycosyl hydrolases (10–12).

Stereocchemical outcome studies on the family 20 chitobiase from Serratia marcescens (13) and human β-hexosaminidase (14) demonstrated that this family operates via a retaining

---

1 The abbreviations used are: β-hexosaminidase, β-N-acetyhexosaminidase; HexA and HexB, human β-hexosaminidase A and B, respectively; SpHEX, S. plicatus β-hexosaminidase; SmCHB, S. marcescens chitobiase; NAG, N-acetylgalactosamine; rms, root mean square; MAD, multiwavelength anomalous diffraction.
catus (21). The 55-kDa enzyme, referred to as SpHEX, is a highly active and stable glycosyl hydrolase that functions over a broad pH range. Co-crystallization of SpHEX with the cyclic intermediate analogue N-acetylglucosamine (NAG)-thiazoline (Fig. 1B and Ref. 19), and subsequent crystallographic analysis has provided decisive structural evidence for a substrate-assisted catalytic mechanism involving 2-acetamido group participation, resulting in the formation of a covalent, cyclic intermediate (Fig. 1).

**EXPERIMENTAL PROCEDURES**

**Protein Expression and Purification—Escherichia coli strain JM109** was used for plasmid amplification, and plasmid purification was carried out using Qiagen purification systems. Restriction enzymes and Vent DNA polymerase were from New England Biolabs. T4 DNA ligase was from Roche Molecular Biochemicals. All cloning procedures are described in Ref. 22. SpHEX is a 506-amino acid protein having a predicted molecular mass of 55010 Da (GenBank accession number AF063001). It was expressed as a recombinant, N-terminal His7-tagged fusion protein. Briefly, the plasmid pHEX-1.8 (11) contained the SpHEX open reading frame. The first 100 base pairs of the 5′-end of the SpHEX open reading frame was amplified by the polymerase chain reaction using the sense primer (5′-GGAAATTCATTACCATCATCATCAT- CATCATCATAACGCGGCACGCGACGCGAAG-3′) and the antisense primer (5′-TGGCGCGCGCGGCGGTCCGCGGGCAGGCGGG-3′). This polymerase chain reaction product was restriction digested with Ascl and Ndel for ligation into the final expression plasmid. To obtain the remaining 1.7-kilobase pair fragment of the SpHEX open reading frame, a further aliquot of pHEX-1.8 was restriction digested with Ascl and BamHI. The 100-base-pair (Ndel/Acl) and 1.7-kilobase pair (Ascl/BamHI) fragments were then ligated into the T7 expression plasmid pET-3a (Novagen) that had been linearized by digestion with Ndel and BamHI. The ligated product resulted in the expression plasmid p3AHEX-1.8 whose sequence was verified prior to use in fusion protein expression.

The His7-SpHEX fusion protein was expressed in *E. coli* strain BL21 (DE3). Transformed cells were grown at 37 °C to an *A*<sub>600</sub> = 0.5 and then induced with 0.4 mM isopropyl-1-thio-β-D-galactopyranoside for 3 h at 25 °C. Cells were pelleted by centrifugation, resuspended in a lysis buffer (20 mM Tris-Cl, pH 8.0, 300 mM NaCl, 10 mM imidazole, 10 mM β-mercaptoethanol) and lysed by French press. After centrifugation at 20,000 × *g* for 1 h, the supernatant was loaded onto a nickel-nitrilotriacetic acid superflo (Qiagen) column equilibrated with lysis buffer. Once loaded, the column was washed with the lysis buffer supplemented with 80 mM imidazole (pH 8.0). The fusion protein was eluted from the column using lysis buffer supplemented with 250 mM imidazole, pH 8.0, and precipitated with 50% ammonium sulfate for storage at 4 °C. Aliquots of the precipitated protein were routinely resuspended and dialyzed twice against 50 mM trisodium citrate, pH 6.0, 300 mM NaCl, and 0.5 mM dithiothreitol and then concentrated to ~10 mg/ml with a Millipore concentrator. Approximately 40–60 mg of pure fusion protein was routinely obtained per liter of culture. Electrospray ionization mass spectrometric analysis using a VG Quattro triple quadrupole mass spectrometer (VG Biotech, Altrinningham, UK) determined the mass of the purified fusion protein to be 56,054 Da, in good agreement with the theoretical mass of 56,049 Da.

Seleno-Met-substituted His7-SpHEX was expressed in *E. coli* strain BL21 (DE3) pLYS S using the method described in Ref. 23. Transformed cells were grown at 37 °C in M9 minimal medium until mid-log phase growth was reached. The culture was then supplemented with 0.5 mM Lys, 0.8 mM Thr, 0.6 mM Phe, 0.5 mM Leu, 0.8 mM Ile, and 0.8 mM Val to inhibit endogenous Met biosynthesis. After a 30-min incubation, the culture was further supplemented with 0.25 mM seleno-Met and induced with 0.5 mM isopropyl-1-thio-β-D-galactopyranoside for 10 h. Seleno-Met substituted His7-SpHEX was purified in the same manner as native His7-SpHEX except that the protein was dialyzed against 3 mM dithiothreitol before concentrating to avoid selenium oxidation. Electrospray mass spectrometric analysis verified that all 6 Met residues in the 512-amino acid SpHEX protein had been substituted with seleno-Met. All purified fusion protein was visualized for purity by SDS-polyacrylamide gel electrophoresis.

**Crystallographic and Data Collection**—Both native and seleno-Met substituted His7-SpHEX crystallized in the hexagonal space group P6<sub>3</sub>,22 within 2 weeks by vapor diffusion at room temperature. The mother liquor consisted of 2.2 M ammonium sulfate, 100 mM trisodium citrate, pH 6.0, and 20–25% glycerol. Hanging drops were set up by
mixing an aliquot of SpHEX (concentrated to 10 mg/ml) with an equal amount of the mother liquor. Crystals of His-SpHEX in complex with NAG-thiazoline were obtained by co-crystallization of the native protein (from which dithiothreitol had been removed by dialysis) with 2–5 mM NAG-thiazoline. Diffraction data for a MAD phasing experiment were collected at the Advanced Photon Source, BioCARS sector beamline BM-14-C and BM-14-D on native and seleno-Met substituted His-SpHEX crystals flash cooled to 100 K, respectively (see Table I). Diffraction data from crystals of the complex between His-SpHEX and NAG-thiazoline were collected at Stanford Synchrotron Radiation Laboratory, beamline 9-2 (see Table I). All diffraction data were processed using DENZO and SCALEPACK (24).

Structure Determination and Refinement—A solution to the crystal structure of the protein was obtained by a MAD phasing experiment performed on seleno-Met-substituted protein crystals (25). A combination of data derived from the MAD phasing experiment at beamline BM-14-D with data collected from native SpHEX crystals at beamline BM-14-C allowed for the determination of the three-dimensional structure of S. plicatus β-hexosaminidase to 2.2 Å resolution. Although the SpHEX crystals diffracted to slightly higher resolution than 2.2 Å, data collection was restricted to this resolution to avoid excessive data refection caused by spot overlap. The program SOLVE (26) was used for local scaling of the data and to calculate the anomalous and dispersive differences needed to find selenium sites and to determine phase probability distributions. Patterson maps, calculated from the anomalous and dispersive differences, allowed us to find clearly five of the six selenium atoms present in the SpHEX structure. The missing selenium atom was part of the initiation Met whose position could not be determined because of disorder of the first 14 residues of the His-tagged N terminus.

Electron density maps, generated using structure factor phases obtained from the MAD phasing experiment (initial figure of merit 0.8), were improved only slightly by solvent flattening using Density Modification (Fig. 2 and Ref. 27). Map boundaries were extended beyond the CCP4 asymmetric unit using EXTEND (28) and skeletonized using MAPMAN (29). A molecular model of the enzyme was built from the skeletonized map using O (30). Residues 8–512 were readily fit into the density as one continuous chain. Coordinates for the small molecules glycerol and SO₄²⁻ were obtained through the HIC-UP world wide web site, and their geometries were optimized by X-PLOR (31) prior to use in model building.

The molecular model of native SpHEX was refined using a maximum likelihood target function during both simulated annealing and conjugate gradient minimization as implemented in Crystallography and Nuclear Magnetic Resonance System (32). Prior to refinement, 10% of the diffraction data was randomly flagged for cross-validation using the free R factor. After each round of refinement, the model was manually inspected with O using 2Fᵦ – F and Fᵦ – F maps. The final refinement statistics for the model reflect the high quality data (see Table I).

NAG-thiazoline Complex—An Fᵦ – F map, used to visualize NAG-thiazoline in the active site, was obtained using structure factor phases calculated from the native SpHEX model that had been positioned into the unit cell of the NAG-thiazoline complex using rigid body refinement followed by conjugate gradient minimization. Solvent molecules were removed from the model before placing it into the new cell and were relocated during later rounds of refinement. Any waters found in the active site were deleted until the NAG-thiazoline had been modeled into the electron density ascribed to it. The initial NAG-thiazoline model and its geometrical parameters were based on the x-ray crystal structure of N-acetylglactosamine-thiazoline (GalNAc-thiazoline). Refinement of the NAG-thiazoline complex was carried out using Crystallography and Nuclear Magnetic Resonance System as described for the native SpHEX model above. The final refinement statistics are presented in Table I.

Coordinates—The coordinates and structure factors have been deposited into the Protein Data Bank (native SpHEX Protein Data Bank code 1HP4; SpHEX/NAG-thiazoline complex Protein Data Bank code 1HP5).

RESULTS AND DISCUSSION

Structure of β-Hexosaminidase—Excellent crystallographic data (Table I) produced easily interpretable electron density maps into which a model of SpHEX was built (Fig. 2). The enzyme is a kidney shaped, two-domain protein having overall dimensions of ~68 × 58 × 56 Å (Fig. 3). The two domains of SpHEX have a similar fold to domains II (residues 214–335) and III (residues 336–818) of SmCBH (Fig. 4); however, significant deviations between the two structures exist. The most striking structural difference between SpHEX and SmCBH is the absence in SpHEX of two of the four domains that compose SmCBH (Fig. 4). This results in a solvent-exposed active site at the C-terminal end of the (βα)₇ barrel forming domain II. Such a solvent-exposed active site appears to explain why β-hexosaminidases, such as human β-hexosaminidase A, can accommodate large glycoconjugates like GalNAc-β(1,4)-N-acetyll neuraminic acid (2,3)-Gal-β(1,4)-Glc-ceramide ganglioside.

Domain I of SpHEX is composed of residues 1–151. As in SmCBH, this domain has an α/β topology consisting of a solvent exposed, seven-stranded anti-parallel β-sheet that buries two, roughly parallel, α-helices (Fig. 3). Similar topologies have been found in metalloproteinases (10) and collagenases.

D. J. Vocadlo and S. G. Withers, unpublished data.
The amino acid sequence identity between SpHEX and SmCHB is lowest throughout the entire family. Such conservation suggests a functional requirement for this domain by family 20 glycosyl hydrolases.

Domain II of SpHEX is composed of residues 151–512 and is folded into a (βα)₈ barrel with the active site of the enzyme residing at the C termini of the 8 β-strands of the barrel. This domain is homologous to domain III in SmCHB, and a structure-based sequence alignment demonstrated there to be a 29.5% sequence identity between the two domains, where 236 of the 256 Cα atoms of the two homologous domains have a rms difference of only 1.34 Å. A multiple sequence alignment of all family 20 glycosyl hydrolases indicates there to be a common feature among many family 20 glycosyl hydrolases, including the human isoforms (10, 11). Unlike the basic (βα)₈ barrel motif, domain II of SpHEX contains three major loop structures that extend from the C termini of three of the 8 β-strands of the barrel. First, loop Lp7 replaces helix α7 as described above (Fig. 3). Second, a 36-amino acid loop, Lp2, extends from the C terminus of strand β2 and contains a short helical segment that packs against and stabilizes the third major loop Lp3. Lp3 is a 41-amino acid loop that extends from the C terminus of strand β3 and contains a helical segment that is complimentary to and packs against the helical segment found in Lp2 (Fig. 3). There is only one disulfide bond in SpHEX (Cys⁴⁰⁶–Cys⁴⁰⁸), and its presence close to the base of Lp3 may help to stabilize the conformation of this loop. Lp3 and Lp7 act in concert to form the hydrophobic faces of sugar binding site 1 described below (Fig. 5). There are two homologous loops in SmCHB; however, they are longer and perform an additional function by interacting with a domain not present in SpHEX (SmCHB domain I) (Fig. 4). Finally, an extra helix continues on from helix α8 of the (βα)₈ barrel to complete the C terminus of SpHEX. This extra helix stabilizes domains I and II with respect to each other (Fig. 3). It is interesting to observe that the relative orientation of domains I and II of SpHEX is the same as the homologous domains II and III in SmCHB.

The Complex with NAG-Thiazoline: Mechanistic Implications—According to our x-ray structure of SpHEX and that of the SmCHB-chitobiose complex (10), family 20 glycosyl hydrolases do not appear to contain a side chain in a position suitable to act as a catalytic nucleophile that would stabilize developing oxocarbonium ion character. Instead, it has been observed that, in the conformation bound by the enzyme, the C2 acetamido oxygen of the nonreducing sugar in subsite –1 is held within 3 Å of its C-1 anomic carbon. When in this position, it is

### Table I

| Crystal information | SpHEX | Se-Met-SpHEX | SpHEX:NGT |
|---------------------|-------|-------------|-----------|
| Crystallographic statistics |       |             |           |
| Data set | Monochromatic | Edge | Peak | Remote | Monochromatic |
| Detector | ADSC Q4 | ADSC Q4 | ADSC Q4 | ADSC Q4 | ADSC Q4 |
| Wavelength (Å) | 1.00 | 0.9797 | 0.9795 | 0.9496 | 1.03 |
| Resolution (Å) | 40–2.20 | 30–2.20 | 30–2.20 | 30–2.20 | 40–2.10 |
| High resolution (Å) | 2.28–2.20 | 2.28–2.20 | 2.28–2.20 | 2.28–2.20 | 2.16–2.10 |
| Total observations | 416806 | 432049 | 432434 | 431524 | 361546 |
| Unique reflections | 47455 | 47697 | 47690 | 47648 | 52266 |
| Completeness (%) | 58.5 (98.9) | 64.8 (99.3) | 64.8 (99.0) | 64.0 (99.1) | 64.4 (99.2) |
| R<sub>free</sub> | 0.031 (0.067) | 0.061 (0.126) | 0.056 (0.124) | 0.044 (0.091) | 0.084 (0.263) |
| R<sub>work</sub> | 0.18 | 0.21 | 0.22 | 0.18 | 0.20 |
| Number of atoms | 3864 | 3864 | 3864 | 3864 | 3864 |
| Protein | 14 | 14 | 14 | 14 | 14 |
| Water | 270 | 270 | 270 | 270 | 270 |
| Average B (Å²) | 28.7 | 28.7 | 28.7 | 28.7 | 28.7 |
| Rmsd from ideal geometry | Bond lengths (Å) | 0.0052 | 0.0052 | 1.31 |
| % most favored | 89.4 | 89.4 | 88.9 | 88.9 | 88.9 |
| % additionally allowed | 10.6 | 10.6 | 11.1 | 11.1 | 11.1 |

*Calculated by treating Bijvoet pairs as equivalent.

<sup>a</sup> R<sub>free</sub> = Σ(|I(h)−I(h)|)/ΣI(h), where I(h) is the h<sup>th</sup> intensity measurement and I(h) is the weighted mean of all measurements of I(h).

<sup>b</sup> R<sub>work</sub> = Σ(|I(h)−I(h)|)/ΣI(h).

<sup>c</sup> Ramachandran plot.

<sup>d</sup> Bond angles (°).

<sup>e</sup> % most favoured.

<sup>f</sup> % additionally allowed.

<sup>g</sup> Regions defined by PROCHECK (27).
believed that the acetamido oxygen can act as a nucleophile and attack the anomeric center to form a cyclic NAG-oxazolinium ion intermediate (10).

We have determined the three-dimensional structure of an analogue of the proposed NAG-oxazolinium ion intermediate bound to SpHEX. Because NAG-oxazolinium itself is too hydrolytically unstable for use in structural studies, a relatively stable analogue, NAG-thiazoline, has been synthesized and shown to be a potent competitive inhibitor of jack bean β-hexosaminidase (Kᵰ = 280 nM) (19). NAG-thiazoline also acts as an excellent competitive inhibitor of both SpHEX and human β-hexosaminidase B.

Fig. 6 shows NAG-thiazoline bound in the SpHEX active site and the quality of the electron density into which it was modeled. Excluding O-4 and O-6 because of differences in C-4 chirality and enzyme packing effects, respectively, the remaining atoms in NAG-thiazoline had an rms difference of only 0.071 Å compared with the equivalent atoms in the small molecule structure of GalNAc-thiazoline. NAG-thiazoline was bound in the −1 subsite of SpHEX and adopts a conformation that is close to a 4C₁ chair, although the current data do not exclude small distortions toward a sofa or skew boat conformation. There are no significant changes in the SpHEX structure upon binding NAG-thiazoline except for a slight opening of the active site pocket. Fig. 7 clearly shows Trp residues 344, 361, and 442 of the −1 subsite of SpHEX and the homologous residues in SmCHB (Trp 616, Trp 639, and Trp 737). Val 276 (SpHEX) and Val 293 (SmCHB) are on loop Lp3, whereas Trp 696 (SpHEX) and Trp 750 (SmCHB) are on loop Lp7 and compose the hydrophobic sides of the +1 subsite. The nonreducing sugar of chitobiose superposes onto NAG-thiazoline bound to subsite −1. The reducing sugar of chitobiose, bound to subsite +1 and rotated 90° relative to the nonreducing sugar, superposes onto the glycerol bound to subsite +1 in SpHEX.
hydrophobic pocket is highly reminiscent of that found around the catalytic nucleophile in the structures of glycosyl-enzyme intermediates in "normal" retaining glycosidases. In both cases such an environment would protect the intermediate from solvolysis via unwanted pathways. Importantly, the conformation of the sugar in this intermediate is a C4′ chair in both covalent glycosyl-enzyme and cyclic oxazoline.

Numerous hydrogen-bonding interactions lock NAG-thiazoline into the active site of SpHEX and disperse the positive charge distributed into the thiazoline ring upon cyclization (Fig. 6B). These include at least one hydrogen bond to every hydroxyl group on the pyranose ring. However, no hydrogen bonds to the ring oxygen O-5 are evident; indeed, a hydrogen bond to O-5 would be counter-catalytic because it would decrease the extent of lone pair donation by O-5 to the antibonding orbital of the scissile bond (34, 35).

NAG-thiazoline is held in place particularly strongly by Arg162, which forms hydrogen bonds to both O-3 and O-4 of the inhibitor. The mutation R162H results in a 40-fold decrease in $K_m$ relative to wild type SpHEX and a 5-fold decrease in $V_{\text{max}}$ when assayed using 4-methylumbelliferyl-$\beta$-N-acetylgalcosaminide (11). The resultant 200-fold decrease in $V_{\text{max}}/K_m$ confirms that this residue is involved in stabilization of the transition states occurring along the reaction coordinate. The analogous mutation in the a-subunit of human HexA (R178H) is associated with the B1 variant form of Tay-Sachs disease in which the enzyme appears to be normally folded and processed but lacks sufficient enzymatic activity and thus results in disease (36, 37). Recently, the mutation R211K (homologous to Arg178 of the a-subunit of human HexA) was created in human HexB (8). The mutation resulted in a 10-fold increase in $K_m$, paralleling the findings with SpHEX (Arg162). Furthermore, the $k_{\text{cat}}/K_m$ value for the R211K mutation was 500-fold less than that of the wild type enzyme, suggesting that it may serve a more important role in transition state stabilization than its counterpart in SpHEX (8).

Two particularly important hydrogen-bonding interactions are formed with the thiazoline ring of NAG-thiazoline when it binds to SpHEX. First, the OH of Tyr393 donates a hydrogen bond to the sulfur atom of the thiazoline ring. In the substrate complex such a hydrogen bond would orient the carbonyl oxygen into position for nucleophilic attack on the anomeric carbon C-1. A similar role is envisioned for Tyr569 of SmCHB (10). Second, upon formation of the cyclic intermediate, the nitrogen atom N-2 develops a positive charge and SpHEX appears to stabilize this positive charge by delocalizing it through a hydrogen-bonding network between Asp313, Asp246, and the main chain NH group of Met247. This is seen in the two short hydrogen bonds of 2.5 and 2.4 Å from the nitrogen N-2 of the thiazoline ring and the carboxylate oxygens of Asp313 and Asp246, respectively (Fig. 6). These short hydrogen bond distances indicate that the carboxylate of Asp313 is likely deprotonated and possesses a delocalized negative charge during catalysis.

The other key residue in the active site of retaining glycosidases is the acid/base catalyst, which adopts a dual role, functioning to protonate the departing aglycone in the first step and then to deprotonate the incoming water in the second step. In the structure of the complex of SmCHB with chitobiose, a 2.9 Å hydrogen bond was seen between the glycosidic oxygen of chitobiose and Glu540, leading to the assignment of Glu540 as the acid catalyst (10). Comparative molecular modeling combined with site-directed mutagenesis and kinetic studies of SpHEX and human $\beta$-hexosaminidase subunits $\alpha$ and $\beta$ have shown Glu314, Glu233, and Glu355 to be homologous to SmCHB Glu540, respectively (10, 11, 38, 39). The mutation E314Q in SpHEX decreases both $V_{\text{max}}$ and $K_m$ for 4-methylumbelliferyl-$\beta$-N-acetylglucosaminide by 296- and 7-fold, respectively, confirming an important role for this residue in catalysis (11). Superposition of the crystal structures of SpHEX and SmCHB confirms that Glu314 of SpHEX is indeed positioned within the active site such that it too would make a hydrogen bond to the glycosidic oxygen of the superimposed chitobiose model (Fig. 7).

The second and final step in the double displacement mechanism is the hydrolysis of the intermediate by general base-catalyzed attack of water at the anomeric center C-1, resulting in overall retention of the anomeric configuration. Figs. 5 and 8 show the position of a glycerol molecule bound in the +1 subsite. This glycerol superimposes onto half of the pyranose ring of chitobiose and suggests that subsite +1 in SpHEX causes the sugar in this subsite to be twisted ~90° relative to the sugar bound in subsite −1 (Fig. 5). Furthermore, one of the hydroxyl groups of this glycerol is within 3.4 Å of the anomeric C-1 of NAG-thiazoline and forms a hydrogen-bonding interaction with the carboxylate of the general acid/base Glu314. We postulate that this hydroxyl group occupies the position that an incoming water molecule would take to nucleophilically attack C-1, thereby hydrolyzing the oxazolinium ion intermediate, with release of $\beta$-N-acetylglucosamine. Absorption of the proton from water by Glu314 is assisted by a hydrogen-bonding network formed between its carboxylate group, the imidazole nitrogens of His359, the carboxylate of Asp191 and the main chain NH group of Asp192 (Fig. 8). The active site water molecule seen in the SmCHB structure, and proposed to be the reactant species (10), is indeed conserved in the SpHEX structure and is indicated in Figs. 6 and 8 as WAT. However, this water molecule is buried within the active site of both structures, and it seems more plausible that the incoming water enters directly from the bulk solvent after departure of the aglycone rather than occupying this site first. The role of buried

Fig. 6. NAG-thiazoline bound to SpHEX. A, electron density for NAG-thiazoline. The refined model is drawn as sticks with carbon atoms in purple, nitrogen atoms in blue, oxygen atoms in red, and the sulfur atom in yellow. The map was calculated as a $|F_o| - |F_c|$ simulated annealing omit map as implemented in Crystallography and Nuclear Magnetic Resonance System (32). B, the SpHEX active site architecture showing hydrogen-bonding interactions with NAG-thiazoline (NGT). NAG-thiazoline is in a full C4′ conformation. Asp313 and Tyr393 are showing hydrogen-bonding interactions with NAG-thiazoline (Fig. 6). These include at least one hydrogen bond to every group of this glycerol is within 3.4 Å of the anomeric C-1 of NAG-thiazoline and forms a hydrogen-bonding interaction with the carboxylate of the general acid/base Glu314. We postulate that this hydroxyl group occupies the position that an incoming water molecule would take to nucleophilically attack C-1, thereby hydrolyzing the oxazolinium ion intermediate, with release of $\beta$-N-acetylglucosamine. Absorption of the proton from water by Glu314 is assisted by a hydrogen-bonding network formed between its carboxylate group, the imidazole nitrogens of His359, the carboxylate of Asp191 and the main chain NH group of Asp192 (Fig. 8). The active site water molecule seen in the SmCHB structure, and proposed to be the reactant species (10), is indeed conserved in the SpHEX structure and is indicated in Figs. 6 and 8 as WAT. However, this water molecule is buried within the active site of both structures, and it seems more plausible that the incoming water enters directly from the bulk solvent after departure of the aglycone rather than occupying this site first.

\[^3\]M. Joshi, personal communication.
water is unclear, but structured waters that mediate the binding of sugars with proteins are quite common and may provide some of the flexibility required to accommodate substrates of both gluco and galacto configuration.

A β-retaining mechanism utilizing acetamido group participation in family 20 β-hexosaminidases and chitobiases is consistent with observations from glycosyl hydrolases from the functionally related family 18 (17). In this family, there is also no apparent enzyme nucleophile, and crystallographic analysis of the family 18 plant chitinase hevamine in complex with the chitinase inhibitor allosamidin suggests that a similar cyclic reaction intermediate is formed in chitinases by C2-acetamido group participation (18, 40). Further examples of enzymes possibly utilizing substrate assisted catalysis include soluble lytic transglycosylase (41), and goose lysozyme (42). Hence, it appears that substrate-assisted catalysis is a common feature between glycosyl hydrolase families 18 and 20 and potentially other families.

**Mechanistic Conclusion**—A combination of the results from this study, in which the structure of a complex with an intermediate analogue is presented, with those from a previous study of the structure of the substrate (chitobiose) complex with SmCHB allows interesting insights into the reaction mechanism and particularly into the substrate conformational changes that occur along the reaction coordinate.

The substrate binds to the enzyme with the sugar in the −1 subsite in a distorted sofa/boat conformation, as seen in the bound chitobiose structure (Fig. 7). This places the scissile bond in a pseudo-axial orientation similar to that seen for the complex of lysozyme with NAM-NAG-NAM bound as a product (43). Such a conformation allows atoms C-5, O-5, C-1, and C-2 of the sugar in the −1 site to adopt the coplanar configuration.
required for effective overlap of the nonbonding lone pair of electrons on O-5 with the antibonding orbital at the electron-deficient anomeric center of the oxacarbenium ion. This conformation not only satisfies the requirements of stereoelectronic theory, it also obeys the principle of least nuclear motion and the need to minimize 1,3-diaxial repulsive interactions between the approaching nucleophile and H3 and H5 of the substrate (15, 35, 44). A similar conformational distortion of the analogous sugar has been observed in a nonhydrolyzable thioligosaccharide mimic of cellulose bound to endoglucanase I from family 7 (45). Upon cleavage of the glycosidic bond, with concerted proton donation from Glu134, the sugar ring relaxes to the C2 chair conformation, as evidenced by the structure of the sp(3)Hex-thiazoline complex. Hydrolysis of this intermediate then follows a similar conformational itinerary, with formation of a product complex in a skew boat conformation, and finally product release. A very similar conformational itinerary has been shown for a normal retaining β-hexosaminidase in which a covalent glycosyl-enzyme intermediate is formed.

Interestingly, these crystal structures reveal that as the bound substrate proceeds along the reaction coordinate to yield the enzyme-bound product, the greatest nuclear motion of heavy atoms occurs at C-1, as shown in Fig. 9. As the reaction proceeds, the C-1 atom scribes an arc from its initial position (position 1) as it breaks a covalent bond to the glycosidic oxygen to form a new bond with the acetamido oxygen (position 3). Approximately halfway along this arc is the transition state where C-1, C-2, C-5, and O-5 are coplanar (position 2).

Acknowledgments—We thank K. Ng, M. Fraser, W. Wolodko, E. Bergman, and the staff of BioCARS and Stanford Synchrotron Radiation Laboratory for assistance with x-ray data collection; L. Burke for performing mass spectrometric analyses; and Dr. Brian Patrick for performing the X-ray structure determination of GalNAc-thiazoline.

REFERENCES

1. Davies, G., and Henrisatt, B. (1996) Structure 3, 853–859
2. Henrisatt, B. (1991) Biochem. J. 280, 309–316
3. Henrisatt, B., and Bairoch, A. (1993) Biochem. J. 293, 781–788
4. Henrisatt, B., and Davies, G. (1997) Curr. Opin. Struct. Biol. 7, 637–644
5. Gravel, R. A., Clark, J. T. R., Kahack, M. M., Mahuran, D., Sandhoff, K., and Suzuki, K. (1995) in The Metabolic Basis of Inherited Disease (Scrivner, C. R., Beaudet, A. L., Sly, W. S., and Valle, D., eds) 7th Ed., pp. 1807–1839, McGraw-Hill Inc., New York
6. Cao, Z., Petroulakis, E., Sals, T., and Triggs-Raine, B. (1997) J. Biol. Chem. 272, 14975–14982
7. Brown, C. A., and Mahuran, D. J. (1993) Am. J. Hum. Genet. 53, 497–508
8. Hou, Y., Vocadlo, D., Withers, S., and Mahuran, D. (2000) Biochemistry 39, 6216–6227
9. Church, W. B., Swenson, L., James, M. N., and Mahuran, D. (1992) J. Mol. Biol. 227, 577–580
10. Tews, I., Perras, A., Oppenheim, A., Dauter, Z., Wilson, K. S., and Vorgias, C. E. (1996) Nat. Struct. Biol. 3, 638–648
11. Mark, B. L., Wasney, G. A., Sals, T. J., Khan, A. R., Cao, Z., Robbins, P. W., James, M. N., and Triggs-Raine, B. L. (1998) J. Biol. Chem. 273, 19618–19624
12. Prag, G., Papaniokolou, Y., Tavlas, G., Vorgias, C. E., Petroulakis, K., and Oppenheim, A. B. (2000) J. Mol. Biol. 300, 611–617
13. Drouillard, S., Armand, S., Davies, G. J., Vorgias, C. E., and Henrisatt, B. (1997) Biochem. J. 328, 945–949
14. Lai, E. C., and Withers, S. G. (1994) Biochemistry 33, 14743–14749
15. Sinnott, M. L. (1990) Chem. Rev. 90, 1171–1202
16. McCarter, J. D., and Withers, S. G. (1994) Curr. Opin. Struct. Biol. 4, 885–892
17. Tews, I., Tewfik, A. A., Schobinger-Poenitz, N., Wilson, K. S., and Dijkstra, B. W. (1997) J. Am. Chem. Soc. 119, 7654–7659
18. Tewfik, A. A., van Scheltinga, A. C., Armand, S., Kalk, K. H., Isogai, A., Henrisatt, B., and Dijkstra, B. W. (1995) Biochemistry 34, 16519–16523
19. Knapp, S., Vocadlo, D., Cao, Z., Tewfik, A. A., and Withers, S. G. (1996) J. Am. Chem. Soc. 118, 6804–6805
20. Legler, G., and Bellahgen, R. (1997) Carbohydr. Res. 233, 113–123
21. Robbins, P. W., Overby, K., Albright, C., Benfield, B., and Pero, J. J. (1992) Gene (Amst.) 111, 69–76
22. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
23. Van Duyne, G. D., Standaert, R. F., Karpus, P. A., Schreiber, S. L., and Clardy, J. (1993) J. Mol. Biol. 229, 105–124
24. Gao, Z., and Withers, S. G. (1994) Biochemistry 33, 5374–5381
25. Hendrickson, W. A. (1991) Science 254, 51–58
26. Terwilliger, T. C., and Berendzen, J. (1999) Acta Crystallogr. Sect. D Biol. Crystallogr. 55, 849–861
27. Cowtan, K. (1999) Acta Crystallogr. Sect. D Biol. Crystallogr. 55, 1555–1567
28. Collaborative Computational Project Number 4. (1994) Acta Crystallogr. Sect. D Biol. Crystallogr. 50, 760–763
29. Kleywegt, G. J., and Jones, T. A. (1996) Acta Crystallogr. Sect. D Biol. Crystallogr. 52, 826–828
30. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallogr. Sect. A 47, 110–119
31. Brugger, A. T. (1993) XPLOR: A System for X-ray Crystallography and NMR, Yale University Press, New Haven, CT
32. Brugger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, S. J., Kuksa, T., Milenkoski, B., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D Biol. Crystallogr. 54, 905–921
33. Guex, N., and Peitsch, M. C. (1997) Electrophoresis 18, 2714–2723
34. Sidhu, G., Withers, S. G., Nguyen, N. T., McIntosh, L. P., Ziser, L., and Brayer, G. D. (1999) Biochemistry 38, 5346–5354
35. Kirby, A. J. (1984) Acc. Chem. Res. 17, 305–311
