Structural Rationale for Low-Nanomolar Binding of Transition State Mimics to a Family GH3 β-D-Glucan Glucohydrolase from Barley[†,‡]

Maria Hrmova,*[,§] Victor A. Streltsov,[||] Brian J. Smith,[⊥] Andrea Vasella,[#] Joseph N. Varghese,[||] and Geoffrey B. Fincher[§]

School of Agriculture and Wine and Australian Centre for Plant Functional Genomics, University of Adelaide, Waite Campus, Glen Osmond, SA 5064, Australia, Division of Molecular and Health Technologies, Commonwealth Scientific and Industrial Research Organization, Victoria 3052, Australia, The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria 3050, Australia, and Laboratorium für Organische Chemie, ETH-Hönggerberg, CH-8092 Zürich, Switzerland

Received July 27, 2005; Revised Manuscript Received September 22, 2005

ABSTRACT: The interactions of a transition state mimic anilinomethyl glucoimidazole (AmGlcIm), with a K_i constant of 0.6×10^{-9} M and a Gibbs free energy value of −53.5 kJ/mol, with a family GH3 β-D-glucan glucohydrolase from barley have been analyzed crystallographically and by ab initio quantum mechanical modeling. AmGlcIm binds 3 times more tightly to the β-D-glucan glucohydrolase than a previously investigated phenyl glucoimidazole. In the enzyme−AmGlcIm complex, an additional residue, Tyr253, and a water molecule positioned between subsites −1 and +1 are recruited for binding. Analyses of the two binary complexes reveal the following. (i) An intricate network exists in which hydrogen bonds between the enzyme's catalytic pocket residues Lys206, His207, Tyr253, Asp285, and Glu491 and the glucoimidazoles are shorter by 0.15−0.53 Å, compared with distances of hydrogen bonds in the Michaelis complex. (ii) The "glucose" moiety of the glucoimidazoles adopts a ^{4}E conformation that is vital for the low-nanomolar binding. (iii) The N1 atoms of the glucoimidazoles are positioned nearly optimally for in-line protonation by the Oε1 atom of the catalytic acid/base Glu491. (iv) The enzyme derives binding energies from both glycone and aglycone components of the glucoimidazoles. (iv) The prevalent libration motion of the two domains of the enzyme could play a significant role during induced fit closure in the active site. (v) Modeling based on the structural data predicts that protons could be positioned on the N1 atoms of the glucoimidazoles, and the catalytic acid/base Glu491 could carry an overall negative charge. (vi) The enzyme−AmGlcIm complex reveals the likely structure of an early transition state during hydrolysis. Finally, the high-resolution structures enabled us to define minimal structures of oligosaccharides attached to Asn221, Asn498, and Asn600 N-glycosylation sites.

Barley β-D-glucan glucohydrolase is a two-domain enzyme consisting of an $(\alpha/\beta)_8$ TIM barrel and an $(\alpha/\beta)_6$ sandwich, connected by a helixlike linker (*1*). The enzyme belongs to a large group of family GH3 glycoside hydrolases, which currently includes more than 820 GenBank/GenPep entries (*2*). The barley enzyme represents the only completely determined three-dimensional (3D)[1] structure in the GH3 family of hydrolases, although the structure of the $(\alpha/\beta)_8$ TIM barrel domain 1 of an *N*-acetyl-β-glucosaminidase from *Vibrio cholerae* has recently been reported in its free and *N*-acetyl-β-D-glucosamine (GlcNAc)-complexed forms (PDB entries 1TR9 and 1Y65, respectively; J. Gorman and L. Shapiro, New York Structural Genomics Research, New York, NY). The barley β-D-glucan glucohydrolase catalyzes hydrolytic removal of nonreducing glucosyl end residues from a broad range of β-D-glucans and β-D-glucosides (*3*). The enzyme has two glucosyl binding subsites (*4*, *5*), designated −1 and +1, positioned in a 13 Å deep active site pocket that is located at the interface between the two domains (*1*).

The phenyl glucoimidazole (PheGlcIm) and anilinomethyl glucoimidazole (AmGlcIm) used in this work are excellent chemical tools for mimicking transition states of glycoside hydrolases (*5*−*8*). These inhibitors belong to the C2-substituted gluco-configured tetrahydroimidazopyridine fam-

[†] This research was supported by grants from the Australian Research Council to G.B.F., J.N.V., and M.H., from the Australian Partnership for Advanced Computing Facility to B.J.S., and from the Australian Synchrotron Research Program to J.N.V. and M.H. (Advanced Photon Source), and V.A.S. (Photon Factory). Use of synchrotrons was supported by the Australian Synchrotron Research Program, which is funded by the Commonwealth of Australia under the Major National Research Facilities Program.

[‡] The coordinates for the β-D-glucan glucohydrolase−PheGlcIm and −AmGlcIm complexes have been deposited in the Protein Data Bank as entries 1X38 and 1X39, respectively.

* To whom correspondence should be addressed. E-mail: maria.hrmova@adelaide.edu.au. Telephone: +61-8-8303-7280. Fax: +61-8-8303-7109.

[§] University of Adelaide.

[||] Commonwealth Scientific and Industrial Research Organization.

[⊥] The Walter and Eliza Hall Institute of Medical Research.

[#] ETH-Hönggerberg.

[1] Abbreviations: AmGlcIm, anilinomethyl glucoimidazole; E, envelope; F_c and F_o, calculated and observed X-ray structure factor amplitudes, respectively; GlcIm, glucoimidazole; GH, glycoside hydrolases; GlcNAc, *N*-acetyl-β-D-glucosamine; H, half-chair; HKL, indices of reflections; PDB, Protein Data Bank; PheGlcIm, phenyl glucoimidazole; rmsd, root-mean-square deviation; TLS, translation−libration−screw; 3D, three-dimensional.

10.1021/bi0514818 CCC: $30.25
Published on Web 11/18/2005

ily of inhibitors that possess a nonhydrolyzable glycosidic C−N bond between C10 and N1 of the "sugar" and imidazole moieties, and their anomeric carbons are sp^2-hybridized and double-bonded with the exocyclic N1 atoms (*9*, *10*). The conformation of these inhibitors is analogous to those of gluconojiritetrazoles, which in the solid state or in D_2O adopt a $^4H_3$ conformation that is similar to an envelope (^4E) (*6*, *11*). These inhibitors also possess characteristic charge distributions in their structures that could resemble charge distributions during the formation of transition species in the process of hydrolysis. Finally, the C2-substituted gluco-configured tetrahydroimidazopyridines are powerful tools for identification of protonation trajectories of glycoside hydrolases, which may be either anti or syn, depending on the position of the catalytic acid/base versus substrate (*6*, *12*).

In our previous crystallographic work, conduritol B epoxide, 2,4-dinitrophenyl 2-deoxy-2-fluoro-β-D-glucopyranoside, and 4^I,4^{III},4^V-*S*-trithiocellohexaose were used to identify catalytic amino acid residues and to define three key stable intermediates in the catalytic sequence; these intermediates represent a covalent glycosyl−enzyme intermediate, a Michaelis complex, and an enzyme−product complex (*4*). We have also analyzed binding interactions of a transition state mimic PheGlcIm in the active site of the barley β-D-glucan glucohydrolase, and defined a hypothetical transition state structure at a resolution of 2.62 Å (*5*). In the β-D-glucan glucohydrolase−PheGlcIm complex, it was found that the enzyme derives substrate binding energy mainly from the glycone portion of the ligand at the −1 subsite, although the aglycone portion at the +1 subsite might also contribute (*5*).

To rationalize interactions of the enzyme with a second-generation transition state mimic that binds more tightly to the enzyme than the previously investigated generation of glucoimidazoles (*5*, *13*), and to investigate in detail binding of the inhibitor at the +1 subsite, a 3D structure of the β-D-glucan glucohydrolase in complex with AmGlcIm has now been determined at 1.80 Å resolution. The inhibitor AmGlcIm binds 3 times more tightly to the β-D-glucan glucohydrolase than PheGlcIm. Thus, the β-D-glucan glucohydrolase−AmGlcIm complex that could resemble an early transition state has been studied in detail in this work and compared with the enzyme−*S*-cellobioside moiety, or the Michaelis complexes. Further, the crystal structure of the enzyme−PheGlcIm complex determined previously (*5*) was revised using the data now refined to 1.70 Å resolution, and this structure was also compared with the enzyme−AmGlcIm complex. Last, the high-resolution structures enabled us to define the minimal structures of oligosaccharides attached to Asn498 and Asn600 N-glycosylation sites, and these represent an additional β-D-xylosyl sugar residue that is attached by a 1,2-β-linked O-glycosidic linkage to the second GlcNAc moiety of the oligosaccharides at both the Asn498 and Asn600 N-glycosylation sites. On the other hand, the oligosaccharide motif attached to the Asn221 N-glycosylation site contains simply the GlcNAcβ4→1GlcNAcβ4→1Man minimal structure.

EXPERIMENTAL PROCEDURES

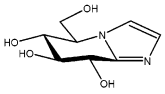
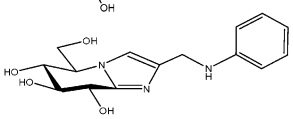
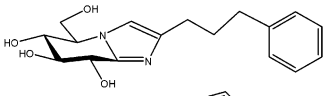
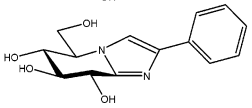
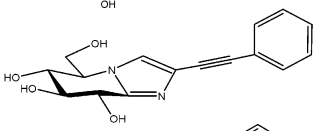
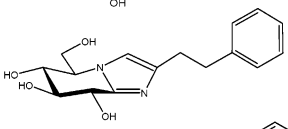
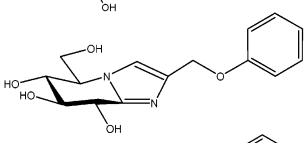
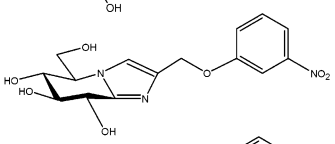
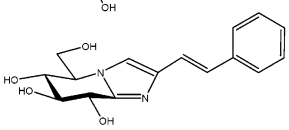
Materials, β-D-Glucan Glucohydrolase, Isoenzyme ExoI Isolation, Enzyme Inactivation by Glucoimidazoles, and Inhibition Parameter (K_i, ΔG) Calculations. These procedures were used as previously detailed (*5*). Briefly, the enzyme inactivation was monitored at 30 °C by incubating 8 μM enzyme in 100 mM sodium acetate buffer (pH 5.25) containing 0.83−6.6 mM 4-nitrophenyl β-D-glucopyranoside as a substrate, 160 μg/mL BSA, and 0−200 nM glucoimidazoles. Each inhibitor was tested at five to six concentrations that were 0.4−3 times greater than the K_i (defined as a dissociation constant of the enzyme−inhibitor complex) values, each in duplicate. The enzyme activity in the presence of inhibitors was monitored at 410 nm, and Dixon plots (inhibitor concentrations vs $1/v$) were used to determine the final K_i values. Substrate hydrolysis during the determination of K_i values never exceeded 10% of the initial substrate concentrations. The inhibition constants K_i were determined by using nonlinear regression analyses and GraFit as specified previously (*5*).

The chemical synthesis of the inhibitors has recently been described (*10*); their chemical structures are depicted in Table 1, and their chemical names are as follows: GlcIm, (5*R*,6*R*,7*S*,8*S*)-5,6,7,8-tetrahydro-5-(hydroxymethyl)imidazo[1,2-*a*]-pyridine-6,7,8-triol; PheGlcIm, (5*R*,6*R*,7*S*,8*S*)-5,6,7,8-tetrahydro-5-(hydroxymethyl)-3-phenylimidazo[1,2-*a*]pyridine-6,7,8-triol; phenethyl GlcIm, (5*R*,6*R*,7*S*,8*S*)-5,6,7,8-tetrahydro-5-(hydroxymethyl)-3-phenethylimidazo[1,2-*a*]pyridine-6,7,8-triol; phenpropyl GlcIm, (5*R*,6*R*,7*S*,8*S*)-5,6,7,8-tetrahydro-5-(hydroxymethyl)-3-(3-phenylpropyl)imidazo[1,2-*a*]pyridine-6,7,8-triol; styryl GlcIm, (5*R*,6*R*,7*S*,8*S*)-5,6,7,8-tetrahydro-5-(hydroxymethyl)-3-styrylimidazo[1,2-*a*]pyridine-6,7,8-triol; phenethynyl GlcIm, (5*R*,6*R*,7*S*,8*S*)-5,6,7,8-tetrahydro-5-(hydroxymethyl)-3-(2-phenylethynyl)imidazo[1,2-*a*]pyridine-6,7,8-triol; phenoxymethyl GlcIm, (5*R*,6*R*,7*S*,8*S*)-5,6,7,8-tetrahydro-5-(hydroxymethyl)-3-(phenoxymethyl)imidazo[1,2-*a*]pyridine-6,7,8-triol; nitrophenoxymethyl GlcIm, (5*R*,6*R*,7*S*,8*S*)-5,6,7,8-tetrahydro-5-(hydroxymethyl)-3-[(3-nitrophenoxy)methyl]imidazo[1,2-*a*]pyridine-6,7,8-triol; and AmGlcIm, (5*R*,6*R*,7*S*,8*S*)-3-(anilinomethyl)-5,6,7,8-tetrahydro-5-(hydroxymethyl)imidazo[1,2-*a*]pyridine-6,7,8-triol.

β-D-Glucan Glucohydrolase Crystallization and Soaking Experiments. Crystals of β-D-glucan glucohydrolase isoenzyme ExoI (*14*) were washed in 100 mM Hepes-NaOH buffer (pH 7.0) containing 1.2% (w/v) polyethylene glycol 400 and 1.7 M ammonium sulfate (solution A), and dissolved by adding an equal volume of 100 mM Hepes-NaOH buffer (pH 7.0) at 4 ± 2 °C. The protein solution, adjusted to 8 mg/mL, was microseeded with previously prepared crystals that were approximately 20−30 μm in the longest dimension (*14*). The next generation of crystals grew for approximately 4 weeks at 4 ± 2 °C, and several crystals that were 120−150 μm in the longest dimension were obtained for soaking experiments.

Data Collection and Structure Determination of the β-D-Glucan Glucohydrolase−PheGlcIm and −AmGlcIm Complexes. Enzyme crystals that were up to 150 μm in the longest dimensions were transferred into fresh solution A, to which either PheGlcIm or AmGlcIm was added directly in solid form; i.e., 0.01 mg of PheGlcIm or AmGlcI was added to 10 μL of the solution A. The final concentrations of the inhibitors corresponded to approximately 2.5 mM. After being soaked for ∼5−10 min at 4 ± 2 °C, the crystals were cryoprotected with 20% (v/v) glycerol in solution A and mounted on synchrotron goniometers in a stream of N_2 gas

Table 1: Inhibition Parameters of β-D-Glucan Glucohydrolase with Glucoimidazoles

Inhibitor	K_i[a] (M)	ΔG[b] (kJ mol^{-1})	Chemical structure
GlcIm[c]	60 x 10^{-9}	-41.9	
AmGlcIm	0.6 x 10^{-9}	-53.5	
Phenpropyl GlcIm	1.3 x 10^{-9}	-51.5	
PheGlcIm[c]	1.7 x 10^{-9}	-50.9	
Phenethynyl GlcIm	1.7 x 10^{-9}	-50.9	
Phenethyl GlcIm[c]	1.8 x 10^{-9}	-50.7	
Phenoxymethyl GlcIm	2.0 x 10^{-9}	-50.5	
Nitrophenoxymethyl GlcIm	2.0 x 10^{-9}	-50.5	
Styryl GlcIm	2.2 x 10^{-9}	-50.2	

[a] Using 4-nitrophenyl β-D-glucopyranoside as a substrate with a K_m of 1.4 mM (*42*). [b] Calculated according to the equation $\Delta G = -RT \ln(1/K_i)$ (*43*). [c] Inhibition data for GlcIm, PheGlcIm, and phenethyl GlcIm taken from ref *5*.

at 100 K (Oxford Instruments, Oxford, England). An X-ray diffraction data set for the enzyme−PheGlcIm complex was collected using an undulator beamline BioCARS 14-ID-B at the University of Chicago Advanced Photon Source (Argonne, IL), which is equipped with a bent cylindrical Si-mirror (Rh coating) diamond(111) double-bounce monochromator and a focusing to the MARCCD-165 detector mirror. The X-ray diffraction data set for the enzyme−AmGlcIm complex was collected using a multipole wiggler beamline BL5 at the Photon Factory in Japan, which is fitted with a collimating mirror, double-crystal Si(111) monochromator, and a focusing to the ADSC Quantum 315 CCD detector mirror. Both data sets were collected at 0.5−1° oscillations throughout the range of 180−360°. The data were processed using the DENZO/SCALEPACK HKL suite of programs (*15*). Autoindexing determined that space groups of crystals were consistent with primitive tetragonal space group $P4_32_12$. The structures of the enzyme−PheGlcIm and enzyme−AmGlcIm complexes were refined using CCP4 REFMAC5 (*16*). The starting model for the β-D-glucan glucohydrolase−PheGlcIm complex was the previously determined crystal structure of the enzyme with bound PheGlcIm [PDB entry 1LQ2 (*5*)], without the ligand and water molecules included. The starting model for the β-D-glucan glucohydrolase−AmGlcIm complex was the refined high-resolution structure of the enzyme−PheGlcIm complex obtained in the work presented here. The iterative model building using XtalView (*17*) and model refinement using REFMAC5 (*16*) allowed tracing of all the residues in the crystal structures. The electron densities for the ligands were well-defined in the active site regions at the 3σ level, and the water molecules were located automatically with CCP4 ARP for the level higher that 2σ in the $m|F_o| - |F_c|$ maps. The water molecules were retained if they satisfied hydrogen bonding criteria and if their $2m|F_o| - D|F_c|$ electron density maps were confirmed after the refinement. During model building and refinement, 5% of the data were flagged for cross validation to monitor the progress of the refinement, using R_{free} statistics (*18*). PROCHECK (*19*) was used to check the geometrical quality of the models. Ramachandran plots (*19*) showed that 90 and 91% of the residues for the enzyme−AmGlcIm and enzyme−PheGlcIm complexes, respectively, were found in the most favorable regions of the plots, and these statistics indicated that the qualities of the structures were good.

Table 2: Data Collection and Refinement Statistics of β-D-Glucan Glucohydrolase−PheGlcIm and −AmGlcIm Complexes

	β-D-glucan glucohydrolase−PheGlcIm complex	β-D-glucan glucohydrolase−AmGlcIm complex
resolution range (highest-shell) (Å)	34.30−1.70 (1.74−1.70)	48.30−1.80 (1.86−1.80)
unit cell dimensions [$a = b$, c (Å)]	100.558, 182.414	99.991, 184.386
no. of unique observations	98154	87042
multiplicity	12.8	14.0
R_{merge} [a,b] (%)	5.4 (63.1)	13.2 (98.1)
$\langle I/\sigma(I)\rangle$ [b]	42.3 (3.4)	23.5 (2.9)
data completeness[b] (%)	99.8 (98.3)	100 (100)
R_{work} [b] (%)	16.0 (27.1)	15.9 (20.6)
R_{free} [b,c] (%)	18.3 (32.0)	18.5 (25.8)
rmsd for bonds (Å), angles (deg)	0.011, 1.326	0.010, 1.264
overall coordinate error (based on R_{free}) (Å)	0.08	0.09

[a] $R_{merge} = 100[\sum(I_i - \langle I\rangle)^2/\sum I_i^2]$, summed over all independent reflections. [b] Data for the highest-resolution shell are in parentheses. [c] Represents approximately 5% of the data.

Quantum Mechanical Calculations. Standard ab initio molecular orbital calculations (*20*) were performed using GAMESS-US (*21*). Geometry minimizations were performed on active site models of the β-D-glucan glucohydrolase. The side chains of Asp95, Asp285, and Glu491 were represented as acetate molecules, and Arg158 was represented as a guanidinium moiety, Lys206 as a methylamine, His207 as an imidazole, and Trp434 as an indole. All atoms of the inhibitor were included. Restrained geometry optimizations at the HF/3-21G level were applied, in which non-hydrogen atoms were restrained to positions observed in the X-ray crystallographic analyses. Atoms were permitted to move 0.50 Å without restraints. Beyond this distance, a harmonic restraint with a force constant of 0.5 aJ Å^{-2} (mdyn Å^{-1}) was applied. Four models that differed in the protonation states of the catalytic acid/base Glu491 or the ligands were considered. The acetate molecule representing Glu491 was either ionized (formal charge of −1) or neutral with a proton bonded to either the Oε1 or the Oε2 atom. The ligand was either neutral or protonated at the N1 atom (formal charge of +1). Model **1** had Glu491 and the ligand ionized; model **2** had Glu491 neutral and the ligand ionized, while models **3** and **4** had both Glu491 and the ligand neutral, with the proton bound to the Oε1 and Oε2 atoms, respectively. In models **3** and **4**, the HOCO dihedral angle formally equals 0°. In model **3**, the distance between Oε1 and H atoms was constrained to 0.97 Å, which represents the distance in the HF/3-21G-minimized geometry of acetic acid.

RESULTS

Inhibition of β-D-Glucan Glucohydrolase by Glucoimidazole Inhibitors. Dissociation constants of the enzyme−inhibitor complex (K_i) and calculated Gibbs free energy (ΔG) values for a range of glucoimidazole inhibitors are summarized in Table 1. The values of the K_i constants varied from 60 to 0.6 × 10^{-9} M, and in all cases, the compounds exhibited competitive inhibition (data not shown) with respect to 4-nitrophenyl β-D-glucopyranoside as a substrate; thus, the $v = (V_{max}S)/[K_m(1 + I/K_i) + S]$ equation was used for calculations of K_i constants. The inhibition experiments showed clearly that the introduction of a phenyl substituent onto the tetrahydroimidazopyridine core structure of the inhibitors *caused* a decrease in the K_i values by approximately 20−100-fold (Table 1). The most effective inhibitor from the studied glucoimidazole series was AmGlcIm, which was selected for a further X-ray crystallographic analysis. The crystal structure of the enzyme−PheGlcIm complex (*5*) was also re-evaluated to higher resolution.

Refinements of β-D-Glucan Glucohydrolase−PheGlcIm and −AmGlcIm Complexes. A new batch of protein crystals was prepared by redissolving previously crystallized protein (*14*), in attempts to obtain high-resolution crystals of β-D-glucan glucohydrolase that could be used for collection of the data for the enzyme−PheGlcIm and enzyme−AmGlcIm binary complexes. The resolution of 2.62 Å that was achieved in our previous work (*5*) was *improved* to 1.70−1.80 Å in the investigation presented here (Table 2).

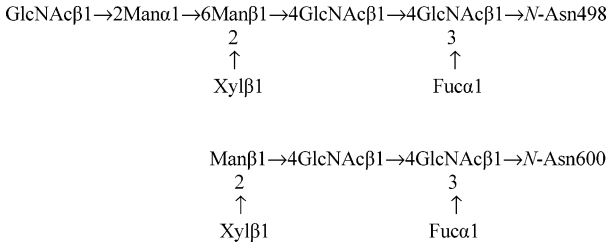
The structure of the β-D-glucan glucohydrolase−PheGlcIm complex consists of 602 amino acid residues, 931 bound water molecules, one SO_4^{2-} ion, two glycerol molecules, and one PheGlcIm molecule. The two glycerol molecules were found to adopt alternative conformations with different temperature factors in this structure. The structure of the β-D-glucan glucohydrolase−AmGlcIm complex is composed of 602 residues, 874 bound water molecules, two SO_4^{2-} ions, three glycerol molecules, and one AmGlcIm molecule.

The final refinement statistics for both enzyme−inhibitor complexes are summarized in Table 2, and the final R_{work} and R_{free} values for the full range of 34.3−1.70 and 48.3−1.80 Å data were 16.0 and 18.3% for the enzyme−PheGlcIm complex and 15.9 and 18.5% for the enzyme−AmGlcIm complex, respectively. Following the convergence in the standard model refinement, a further improvement of more than 2% in the R_{work}/R_{free} ratio was achieved by refining the two domains of the enzyme (domain 1, residues 1−357; domain 2, residues 374−559), as two independent anisotropic domains with a translation−libration−screw (TLS) motion (*22*), as implemented in CCP4 REFMAC5 (*16*).

The high-resolution data in conjunction with a two-domain-like arrangement of the β-D-glucan glucohydrolase provided an opportunity to define the fine details of *rigid body motions* of the two independent domains of the enzyme (*22*−*24*). It was observed in both inhibitor−enzyme complexes that librational motion **L** tensors of domains 1 and 2 were at least one order of a magnitude higher than those of translational motion **T** and screw motion **S** tensors (data not shown). Thus, we concluded that **L** motion tensors exhibited anisotropy and that the two **T** and **S** tensors were less anisotropic and significant, which indicated that translational and screw motions of domains 1 and 2 were limited. The significant anisotropy in the **L** tensors of domains 1 and 2 could be explained by the concerted rotational movement

of domains 1 and 2 toward and away from each other. This mutual librational movement of both domains was also observed in a high-resolution structure of the β-D-glucan glucohydrolase with bound glucose molecule, in which the librational **L** tensor values had marginally but significantly higher values compared to the values in the inhibitor−enzyme complexes presented in this work (A. Peisley, M. Hrmova, V. Streltsov, and J. N. Varghese, unpublished data). Thus, it could be concluded that libration is the prevalent type of motion linking the two domains in the β-D-glucan glucohydrolase and that this motion could play a significant role during induced fit closure of the active site during inhibitor binding. Because the two catalytic residues, Asp285 and Glu491, are located on the two separate domains, it might also be suggested that the two domains move toward and away from each other along the reaction sequence, as the enzyme proceeds through successive catalytic events.

Further, an additional β-D-xylosyl sugar residue was discovered, attached by a 1,2-β-linked O-glycosidic linkage to the second GlcNAc moiety of the oligosaccharides at both the Asn498 and Asn600 N-glycosylation sites. The 1,2-β-bonded glycosyl residue was previously suggested in the β-D-glucan glucohydrolase, but was not modeled in the electron density maps (*1*). The Asn498 and Asn600 N-glycosylation sites either were disordered in the previous crystals or contained no detectable oligosaccharides or partially detectable oligosaccharides (*1*, *4*, *5*, *25*). The oligosaccharides attached to Asn498 and Asn600 of the barley β-D-glucan glucohydrolase isoenzyme ExoI have the following minimal structure:



where GlcNAc, Man, Fuc, and Xyl represent *N*-acetyl-β-D-glucosamine, mannosyl, fucosyl, and xylosyl residues, respectively. It is interesting to note that the oligosaccharide motif attached to the third N-glycosylation site Asn221 contains simply GlcNAcβ4→1GlcNAcβ4→1Man in both refined structures, and these moieties could only be observed in the high-resolution enzyme−ligand complexes refined in this work. The fact that only GlcNAcβ4→1GlcNAcβ4→1Man moieties are found at the N-glycosylation site Asn221, and not branched oligosaccharide motifs as at the Asn498 and Asn600 N-glycosylation sites, would not be entirely unexpected. The reason for this is that the first GlcNAc moiety of the GlcNAcβ4→1GlcNAcβ4→1Man motif is positioned approximately 16 Å from the Trp286 and Trp434 residues, which delineate the entry into the active site pocket of the enzyme. Thus, the presence of fewer sugar moieties at the Asn221 N-glycosylation site could be fully justified, as more oligosaccharides could easily obstruct the entry into the active site pocket. It appears rather challenging to explain a somewhat minimal or limited N-glycosylation pattern of the sugar moieties at the Asn221 site, based on the available structural data, although it is well-known that the extent of N-glycosylation affects expression, structure, stability, antigenicity, and biological function of proteins in plants and other eukaryotes (*26*). It therefore will be important in the future to determine if the presence of fewer sugars at the Asn221 site is linked to a limited biosynthesis of the sugar moieties at this site, and/or is linked to a selective deglycosylation or trimming of an originally fully developed N-glycosylation site by specific exoglycosidases (*26*).

Finally, the unit cell dimensions of the crystals with the two inhibitors studied here *differed significantly* from each other (Table 2). In the unit cell of the crystal with AmGlcIm, the dimensions of the *a*-axis and *b*-axis were reduced by 0.57 Å, and simultaneously, the *c*-axis was increased by 1.95 Å, compared with the unit cell of the crystal with PheGlcIm. These alterations corresponded to approximately 0.6 and 1% changes, compared with the "usual" dimensions of the unit cells (*5*). It is worth noting that the volumes of the unit cells of the two crystals were almost identical and that the difference between the two was less than 0.04% (Table 2). Comparisons of the positions of the enzyme−AmGlcIm and enzyme−PheGlcIm complexes in the unit cells indicated that the molecules adopted very similar positions in the unit cells, and that the only difference in their positions was a translational movement of the enzyme molecule in the crystal with AmGlcIm by approximately 0.5 Å alongside the *c*-axis, without significant rotational repositioning. It is important to point out that all the diffraction data collected so far on the in-house rotating anode or synchrotron X-ray sources were measured under similar experimental physicochemical conditions in 100 mM Hepes-NaOH (pH 7.0) and a temperature of ∼100 K. The changes in unit cell dimension therefore do not reflect different experimental conditions. The changes in the dimensions of the unit cell axes of the enzyme crystal with the bound inhibitor could be translated into significant changes of the crystal dimensions, and it is remarkable that these changes did not result in crystal cracking or shattering. We concluded that during binding of the inhibitor in the active site of the enzyme, the crystal's unit cell dimensions have changed but have not led to crystal anisomorphism or to changes in the space group.

3D Structures of β-D-Glucan Glucohydrolase−PheGlcIm and −AmGlcIm Complexes. The overall fold of the enzyme and the position of the bound AmGlcIm inhibitor at the interface between the two domains are shown in Figure 1A. The interactions of β-D-glucan glucohydrolase with PheGlcIm and AmGlcIm in the two high-resolution structures with overall coordinate errors of less than 0.1 Å (Table 2) indicated that the enzyme interacted similarly with the two structurally related transition state mimics (Figure 2); the overall root-mean-square deviation (rmsd) value between the two structures was 0.127 Å. The glucoimidazole and phenyl or aniline moieties of the PheGlcIm and AmGlcIm inhibitors were bound at the −1 and +1 subsites of the active site, respectively, and in both enzyme−ligand complexes, a network of 12−13 mono- and cooperative bi- and tridentate hydrogen bonds (*27*, *28*) of less than 3.27 Å was formed (Figures 2 and 3A). Of particular interest are the separations between Oδ1 and Oδ2 of Asp95 and the C6OH and C7OH groups of the tetrahydroimidazopyridine moieties [corresponding to the C6OH and C4OH groups of the glucopyranosyl residue at subsite −1 (*4*)], between Nϵ2 of His207 and the C8OH group, between the catalytic nucleophile Oδ2

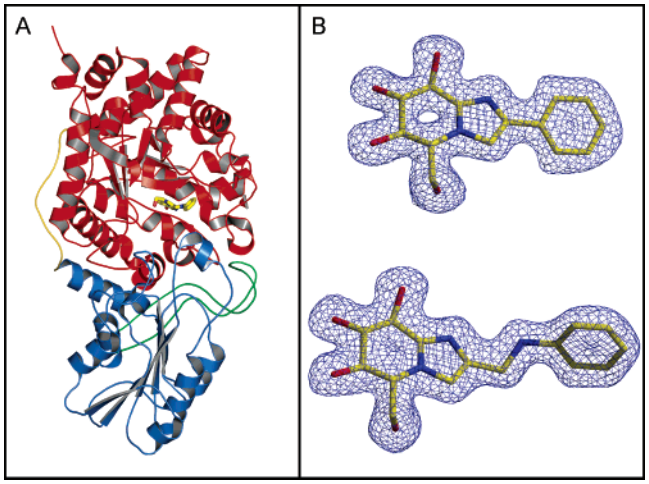


FIGURE 1: Ribbon representation of β-D-glucan glucohydrolase with bound AmGlcIm and the electron density maps of PheGlcIm and AmGlcIm. (A) Ribbon representation of β-D-glucan glucohydrolase with bound AmGlcIm, where domain 1 (residues 1−357), the linker (residues 358−373), domain 2 (residues 374−559), and the COOH-terminal antiparallel loop (residues 560−602) are colored red, yellow, blue, and green, respectively. The atoms of the ligand, AmGlcIm, are colored yellow (carbons), blue (nitrogens), and red (oxygens). (B) Derived $2m|F_o| - D|F_c|$ electron density maps contoured at 1σ for PheGlcIm and AmGlcIm, where the atoms are colored as indicated for AmGlcIm in panel A. The associated electron density maps of the inhibitors are colored blue. This figure was prepared with XTalView (*17*) and PyMol (*44*).

of Asp285 and the C9OH group, and between the Oϵ1 group of acid/base catalyst Glu491 and the N1 atom of the tetrahydroimidazopyridine moieties in both enzyme−inhibitor complexes. All these separations correspond to short interactions of 2.53−2.70 Å (Figures 2 and 3A). Further, there were three (Arg158, Lys206, and Trp434) and four (Arg158, Lys206, Trp434, and Tyr253) other residues directly involved in the binding of PheGlcIm and AmPheIm, respectively. Finally, residues Gly57 and Glu220 were bound to the glucoimidazole inhibitors through water-mediated hydrogen bonds.

Overall, there are seven water molecules in the active site of the enzyme−PheGlcIm and −AmGlcIm complexes, from which five water molecules are associated with the catalytic acid/base Glu491 and another acidic residue Glu220 (Figures 2−4). These two residues are highly conserved in the GH3 family of hydrolases (*29*), and while the former residue is the catalytic acid/base, the latter residue most likely plays a role in coordinating the water molecules around the acid/base catalyst (*4*). These two acidic amino acids might cooperate with Arg291 and guide the associated water molecules to their target location during hydrolysis of substrate molecules (Figure 4). At least four of the solvent molecules are highly conserved in their positions and have been previously observed in the enzyme−2-deoxy-2-fluoro-α-glucopyranosyl and enzyme−PheGlcIm complexes (*4*, *5*).

Comparison of the enzyme−PheGlcIm and −AmGlcIm complexes indicates that in the second complex the newly identified interacting residue Tyr253 and water molecule W3 participate in the binding of AmGlcIm through the hydrogen bonding interactions with the NH group of the aniline moiety of AmGlcIm at the separation of 3.07 Å (Figures 2 and 3A). Water molecule W3 makes a hydrogen bond contact with Oϵ2 of Glu220 and is one of the five water molecules directly surrounding Glu220.

The next interesting feature of the enzyme−AmGlcIm complex is that Arg291 is rotated away from the aniline moiety of AmGlcIm by approximately 100°, compared with the orientation of this residue in the enzyme−PheGlcIm complex; the CG−CD−NE−CZ dihedral angles in the enzyme−PheGlcIm and −AmGlcIm complexes are 173.3° and 73.8°, respectively (Figure 3A). The rotation of the guanidine moiety of Arg291 in the enzyme−AmGlcIm complex also results in tilting of this moiety toward Glu220, and thus a salt link with a separation of 2.86 Å is formed between the NH1 group of Arg291 and Oϵ1 of Glu220. In the enzyme−PheGlcIm complex, this separation is 3.21 Å. Another residue that has been significantly reoriented in the enzyme−AmGlcIm complex is Glu287, which has its Oϵ1 atom 5.09 Å from the NH2 group of Arg291; the corresponding distance in the enzyme−PheGlcIm complex is 3.24 Å (Figure 3A).

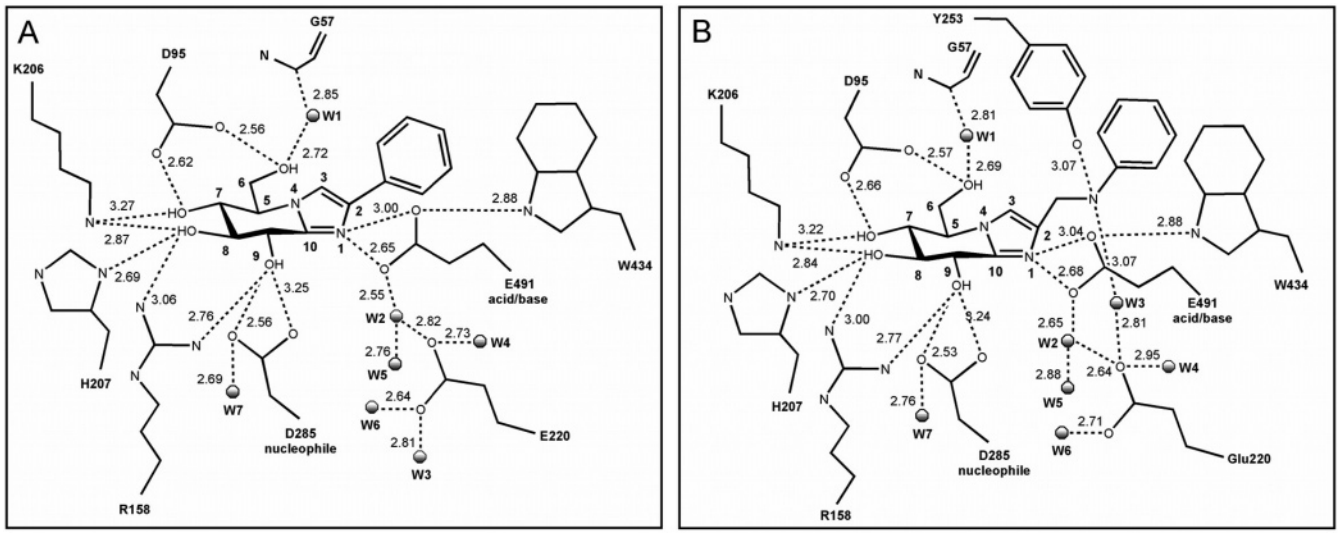


FIGURE 2: Hydrogen bonding interactions of PheGlcIm (A) and AmGlcIm (B) with β-D-glucan glucohydrolase. PheGlcIm is shown in the ^{4}E conformation with atomic numbering of the C atoms of the tetrahydroimidazopyridine moiety. The dashed lines represent hydrogen bonding interactions among the ligand, water molecules, and amino acid residues.

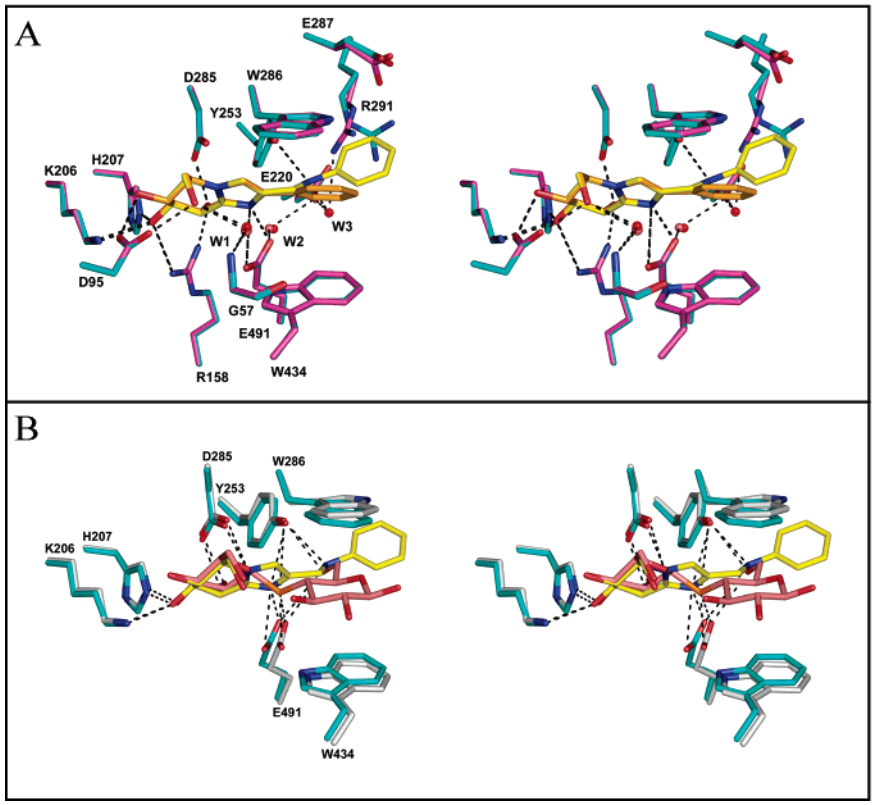


FIGURE 3: Stereo representations of β-D-glucan glucohydrolase active site interactions with PheGlcIm, AmGlcIm, and the *S*-cellobioside moiety. (A) The enzyme−PheGlcIm complex is superposed with the enzyme−AmGlcIm complex, where PheGlcIm is colored orange and the interacting residues are colored magenta. The AmGlcIm ligand is colored yellow (carbons), blue (nitrogens), and red (oxygens), and the interacting residues are colored cyan (carbons), blue (nitrogens), and red (oxygens). The two water molecules (W1 and W2) in the enzyme−PheGlcIm complex and the three water molecules (W1−W3) in the enzyme−AmGlcIm complex are colored pink and red, respectively. The distances of the separations of <3.20 Å are represented by dashed lines (for the distances, cf. Figure 1). The two structures were superposed using the LSQ-Explicit tool as implemented in O (*45*) over the C_α atoms of residues 1−602, with a rmsd value of 0.127 Å. (B) The enzyme−AmGlcIm complex is superposed over the enzyme−*S*-cellobioside moiety complex. The *S*-cellobioside moiety is colored pink (carbons), orange (sulfur), and red (oxygens), and the interacting residues are colored gray (carbons), blue (nitrogens), and red (oxygens). The color coding of the enzyme−AmGlcIm complex is as indicated in panel A. The seven separations, from a total of nine, represented by dashed lines in the enzyme−AmGlcIm complex among residues Lys206, His207, Tyr253, Asp285, and Glu491 are 0.17−0.53 Å shorter than in the enzyme−*S*-cellobioside moiety complex (for the distances, cf. Figure 1). The two structures were superposed as specified above over the C_α atoms of residues 1−602, with a rmsd value of 0.392 Å. This figure was prepared with PyMol (*44*).

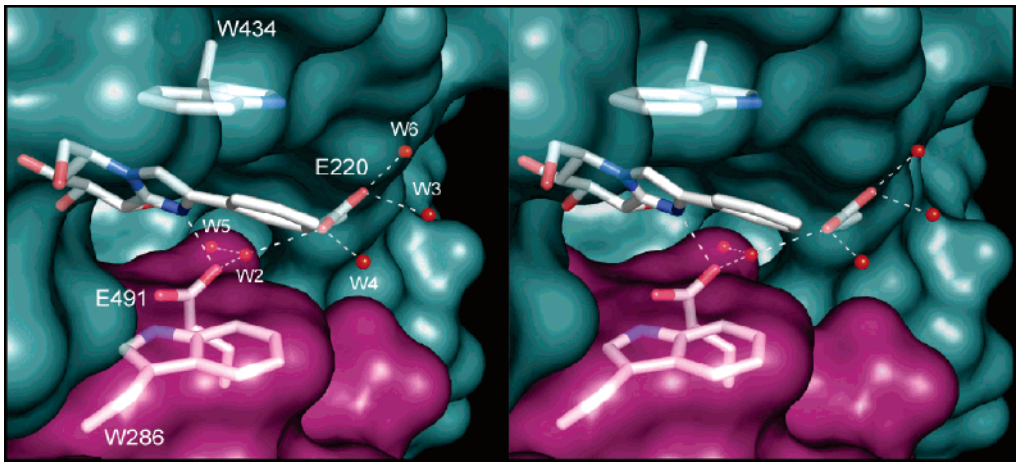


FIGURE 4: Stereo representation of the water channel in the active site of β-D-glucan glucohydrolase involved in the hydrolytic cycle. The two acidic residues (Glu491 and Glu220), the two aromatic residues (Trp286 and Trp434), and the PheGlcIm molecule are presented as tubes, and atoms are colored gray (carbons), blue (nitrogens), and red (oxygens). The five water molecules (W2−W6) depicted as red spheres are connected with each other; Glu491 and Glu220 are connected through 2.55−2.82 Å hydrogen bonding interactions (depicted as white dashed lines), and are directed to the N1 atom of the tetrahydroimidazopyridine moiety, which mimics the glycosidic oxygen. The separation between Oϵ1 of Glu491 and N1 of PheGlcIm is 2.65 Å. This figure was prepared with PyMol (*44*).

The positive electron densities at the 3σ level in the $2m|F_o| - D|F_c|$ electron density maps for PheGlcIm and AmGlcIm during the initial stages of refinements were very clearly defined (data not shown). The final derived electron density maps contoured at 1σ accentuate the continuous nature of the electron densities for each inhibitor and the presence of

ring puckers for the sugar and aromatic moieties (Figure 1B). In agreement with the previously determined enzyme−PheGlcIm complex (*5*), the sugar components of PheGlcIm and AmGlcIm positioned at the −1 subsite of the active site are locked into the ^{4}E conformations with the coplanar C8, C9, C10, N1, and C5 atoms (Figure 3). While the sugar and imidazole moieties of both inhibitors are almost completely superposable, the phenyl ring of PheGlcIm and the aniline moiety of AmGlcIm, positioned at the +1 subsite, adopt quite different orientations, which are dictated by their distinct chemical structures (Figure 3A). The tilt of the phenyl substituent of PheGlcIm by approximately 45° with respect to the planes of the side chains of Trp286 and Trp434 creates a network of cooperative "edge-to-face" interactions between a π-electron system of the indole moiety of Trp286 and the C6 edge of the phenyl ring of PheGlcIm, and between the π-electron system of the phenyl ring of PheGlcIm and the CH_2 moiety of the Trp434 residue. On the other hand, the aniline moiety of AmGlcIm, due to the presence of an extra C−NH group, protrudes further through the Trp286−Trp434 "coin slot" and emerges well beyond the boundaries of the two Trp residues, and is also tilted by ∼45°. However, the tilting is in the opposite direction compared to that of the phenyl ring in the enzyme−PheGlcIm complex (Figure 3A). The protrusion of the aniline moiety of AmGlcIm from the Trp286−Trp434 coin slot still allows the π-system of the phenyl part of the aniline moiety of AmGlcIm to make a polar interaction with the Nϵ1 atom of the pyrrole component of Trp286. The Trp286 residue is the only interacting partner that makes contacts with the phenyl part of AmGlcIm, but despite the lack of interactions, the electron density for the aniline moiety is well-defined (Figure 1B).

The superposition of the β-D-glucan glucohydrolase−AmGlcIm complex onto the 3D structure of the enzyme with a bound *S*-cellobioside moiety (*4*) is shown in Figure 3B. From a total of nine hydrogen bonds, the seven separations between enzyme's catalytic pocket residues Lys206, His207, Tyr253, Asp285, and Glu491 and AmGlcIm are shorter (Figure 3B). In particular, Oϵ1 of Glu491 is 0.31 Å closer to N1 of AmGlcIm, which corresponds to the glycosidic "heteroatom", than Oϵ1 of Glu491 is to the S atom of the *S*-cellobioside moiety. This comparison indicates that the N1 atom of AmGlcIm is positioned nearly optimally for in-line protonation by Oϵ1 of Glu491. Similarly, Oδ1 and Oδ2 of Asp285, the OH group of Tyr253, the NZ group of Lys206, and Nϵ2 of His207 are all 0.17−0.53 Å closer to the corresponding atoms of AmGlcIm than are the corresponding groups of the β-D-glucan glucohydrolase to *S*-cellobioside moiety (Figure 3B). The very similar shortenings of these interactions were also obvious from the comparisons of the enzyme−PheGlcIm and enzyme−*S*-cellobioside moiety complexes (data not shown).

Quantum Mechanical Modeling. Investigation of the protonation states of the catalytic acid/base Glu491 and the N1 atoms of the glucoimidazole moieties of the two inhibitors that correspond to the glycosidic O atoms was addressed through quantum mechanical modeling using the high-resolution enzyme−PheGlcIm and enzyme−AmGlcIm inhibitor structures (Table 3). The distances between the N1 atom of each of the inhibitors and the Oϵ1 atom of Glu491 showed excellent agreement with the distances obtained from the X-ray crystallographic analyses for both models **1** and **3**, where the difference between theory and experiment was less than 0.05 Å (Table 3). In both models, a rotation about the C−C bond of the acetate group (cf. Experimental Procedures) causes a minor disagreement in the distances between the N1 and Oϵ2 atoms; the difference ranges from 0.26 to 0.37 Å (Table 3). The rotation of the C−C bond of the acetate group occurs because the Cγ atom of the catalytic acid/base Glu491 is not tethered during the minimization of the active site model. It is important to point out that in the complete β-D-glucan glucohydrolase structure the position of the Oϵ2 atom of Glu491 is determined in a context of this residue to the remainder of the protein, whereas in the gas phase during the modeling, no such restraints exist. This observation might suggest the presence of a small repulsive force between the N1 and Oϵ2 atoms in the inhibitor−enzyme complex. In models **2** and **4**, the distances between the N1 and Oϵ1 atoms differ from the experimentally observed distances by approximately 0.7 Å. In model **2**, the N1−Oϵ2 distance is too long (0.7−0.8 Å), while in model **4**, this distance is too short, by approximately 0.4 Å (Table 3).

Table 3: Calculated Geometries of β-D-Glucan Glucohydrolase−PheGlcIm and − AmGlcIm Models Evaluated by ab Initio Molecular Orbital Calculations

			distance (Å)	
	Q_{ligand}[a]	Q_{Glu491}[b]	R(N1−Oϵ1)[c]	R(N1−Oϵ2)[d]
enzyme−PheGlcIm model				
1	+1	−1	2.63	3.26
2	0	−1	3.31	3.68
3[e,f]	0	0	2.64	3.37
4[e]	0	0	3.30	2.58
experiment			2.65	3.00
enzyme−AmGlcIm model				
1	+1	−1	2.64	3.32
2	0	−1	3.44	3.86
3[e,f]	0	0	2.64	3.35
4[e]	0	0	3.38	2.58
experiment			2.68	3.04

[a] Formal charge on N1 of the ligand. [b] Formal charge on Glu491 acetate (cf. Experimental Procedures). [c] Distance between Oϵ1 and N1 of the ligand. [d] Distance between Oϵ2 and N1 of the ligand. [e] The formal HOCO angle is 0°. [f] The distance between Oϵ1 and H atoms was constrained to 0.97 Å.

DISCUSSION

Transition state analogues or transition state mimics have been used to study the nature of putative substrate transition states and the mechanistic aspects of oligo- and polysaccharide distortions by glycosylhydrolases (*30*, *31*). The first successful attempts to synthesize tetrahydroimidazopyridine inhibitors with K_i constants in the low nanomolar ranges and with flattened sugar chair conformations containing an exocyclic N atom were reported (*32*, *33*), and it was later suggested that these inhibitors could mimic the planar glycosyl−oxycarbenium ion transition states (*6*, *13*). Recent advances in transition state inhibitor design indicated that members of the C2-substituted gluco-configured tetrahydroimidazopyridines could be very powerful inhibitors for β-retaining glycoside hydrolases (*9*, *10*). Notenboom et al. (*8*), following earlier discussions on the development and formation of transition states in glycoside hydrolases (*6* and references therein), concluded that five main factors could contribute to it. These factors are charge distribution, a

trigonal anomeric center, a half-chair (or close to) conformation, an appropriate relative configuration of OH groups, and an ability of the glycosidic oxygen to be directionally protonated. It has also been suggested that in glycoside hydrolases all these phenomena occur predominantly at the −1 subsite of the catalytic site, where the two catalytic amino acid residues are spatially allocated during the hydrolytic event, and where hydrogen bonding is consistent with the theory that distortions into the transition state geometry along the reaction sequence occur at the −1 subsite (*34*, *35*). However, it is possible that interactions from the residues positioned at the neighboring −2 subsite or alternatively at the +1 subsite could participate in transition state development and formation (*7*, *8*, *35*). In the work presented here, we investigated binding interactions of a barley β-D-glucan glucohydrolase in complex with the AmGlcIm inhibitor, in which its anilinomethyl substituent could protrude into the +1 subsite of the active site region.

In our previous crystallographic work, a hydrolytic reaction pathway for the β-D-glucan glucohydrolase was proposed (*4*). The first stage of the hydrolytic pathway along the reaction trajectory represents formation of the Michaelis complex, which is preceded by the release of a noncovalently bound glucose molecule from the stable enzyme−product complex. It is almost certain that this glucose is displaced from the −1 subsite by the incoming substrate molecule before the next hydrolytic event takes place. The second stage of the hydrolytic pathway involves C1−O glycosidic bond cleavage, which proceeds through a double-displacement mechanism and the formation of a relatively stable covalent glycosyl−enzyme intermediate (*4*). The chemical identity of glucose in the enzyme active site pocket and its occupancy of one in the crystal have recently been determined by evaporative light scattering, gas chromatography, and mass spectrometry analyses after the release of glucose from enzyme crystals (M. Hrmova and G. B. Fincher, unpublished data). It has thus been confirmed by this combined quantitative analytical approach that the glucose molecule indeed has an occupancy of 1 in the crystal of the enzyme, and this finding now directly parallels our earlier crystallographic observations (*1*, *4*).

We have previously shown for the β-D-glucan glucohydrolase with bound PheGlcIm inhibitor that this enzyme−inhibitor complex resembles a hypothetical transition state during the hydrolytic cycle of the enzyme (*5*). In this complex, the inhibitor forms a well-defined electron density in the active site and the "sugar" moiety of the inhibitor adopts the ^{4}E conformation (*5*). It was also shown that the PheGlcIm inhibitor satisfies the definition of a transition state mimic (*35*−*37*), because there is a strong correlation between binding of an inhibitor and apparent pK_{HA} constants of the catalytic machinery of the enzyme. Finally, it was suggested that the enzyme might derive substrate binding energy from the aglycone portion of the PheGlcIm inhibitor (*5*). However, the resolution of 2.62 Å with a coordination error of 0.30 Å did not allow us to specify precisely the fine structural details of the enzyme−PheGlcIm complex.

Against this background, and to obtain more precise insights into the 3D structure of putative substrate transition states during hydrolysis catalyzed by the barley β-D-glucan glucohydrolase, PheGlcIm and the second-generation transition state mimic AmGlcIm were soaked into the crystals, and X-ray diffraction data were collected at 1.70−1.80 Å resolution. The apparent pK_{HA} constants of the N1 atoms of PheGlcIm and AmGlcIm are 4.99 and 5.62, respectively (*10*, *13*), indicating that under the conditions of the X-ray data collection the N1 atoms of the imidazole moieties of the inhibitors might be nonprotonated or neutral. The K_i values for PheGlcIm and AmGlcIm were 1.7×10^{-9} and 0.6×10^{-9} M, respectively, and represent relatively low values compared with those so far observed for retaining β-glucosidases and other β-glucoside hydrolases (*9*, *10*). The low nanomolar K_i constants confirm that the enzyme follows an antiprotonation trajectory during hydrolysis (*5*). The binding of AmGlcIm to the enzyme is approximately 5×10^5 times tighter than that of laminarin, which is the most efficient ground state nonsynthetic substrate so far found for this enzyme (*3*), and is 5×10^7 times tighter than glucose, which represents the hydrolytic product that is always found in the active site of the enzyme purified from barley seedlings at an occupancy of 1 in the enzyme crystal (*1*, *4*).

The key message that emerges from the crystallographic analyses of the β-D-glucan glucohydrolase−PheGlcIm and β-D-glucan glucohydrolase−AmGlcIm complexes is that there is an intricate network in which seven vital hydrogen bonds between the enzyme's catalytic pocket residues Lys206, His207, Tyr253, Asp285, and Glu491 and the corresponding atoms of the glucoimidazoles are shorter by 0.15−0.53 Å, compared with distances of hydrogen bonds in the enzyme−*S*-cellobioside moiety or the Michaelis complexes (Figure 3B; *4*). Some of the residues such as Asp285 and Glu491 in both complexes make very short hydrogen bonding interactions of around 2.53 Å directly with the inhibitors or through water-mediated hydrogen bonds and could correspond to low-barrier hydrogen bonds (*38*). Comparison of the two enzyme−PheGlcIm and enzyme−AmGlcIm complexes indicates first that in the enzyme−AmGlcIm complex an additional residue (Tyr253) makes a hydrogen bond interaction with the NH group of the anilinomethyl moiety of AmGlcIm, which might mimic the position of the C′6OH group of the *S*-cellobiosyl moiety at the +1 subsite (*4*). This observation suggests that the aniline moiety of AmGlcIm could mimic the aglycone portion of the substrate, and the presence of a hydrophilic, hydrogen bond-accepting component in the aglycone portion of the inhibitor is likely to be important for transition state development. Second, an extra water molecule is involved in binding of AmGlcIm. The two additional binding partners are positioned at the interface between subsites −1 and +1, and are not present in the enzyme−PheGlcIm complex (*5*). The new contributing water molecule (Figures 2 and 3) fills an otherwise vacant space in the active site pocket of the enzyme−PheGlcIm complex, and thus this water molecule could be linked directly to the overall binding affinity of AmGlcIm. The calculated ΔG value for AmGlcIm is 2.4 kJ/mol lower than it is for PheGlcIm. This value can be compared to the value of 4.6 kJ/mol that corresponds to a typical loss of a single nonionic hydrogen bond (*39*).

Finally, significant differences in distances among the "ancillary" catalytic residues Glu220, Arg291, and Glu287 have been observed in the two enzyme−inhibitor complexes that could directly affect protonation states of the catalytic residues. The latter residues could influence the overall charge distribution during hypothetical transition state de-

velopment. Thus, there appears to be differences in the mobility of amino acid residues involved in binding of PheGlcIm and AmGlcIm.

In summary, these comparisons indicate that all the criteria suggested for the development of transition states (*8*, *34*, *40*) have been fulfilled in the two binary β-D-glucan glucohydrolase−transition state mimic complexes, and thus, these compounds could represent true transition state analogues. It is suggested that the β-D-glucan glucohydrolase−AmGlcIm complex is likely to represent an early transition state, or the one that precedes the first transition state with a distorted $^4H_3$ conformation, which is close to the ^4E conformation. According to the current interpretation of the hydrolytic mechanism of retaining β-glucosidases and other β-glucoside hydrolases, it is expected that an early transition state would be closer to a distorted glucopyranosyl ring conformation. On the other hand, the late transition state or the one that comes after the second transition state might be closer to an intermediate that is developed after a deglycosylation step during hydrolysis (*12*, *34*, *35*). It is therefore likely that if AmGlcIm binds to the enzyme at both the −1 and +1 subsites and its sugar component adopts the ^4E conformation, then a structure similar to this is also adopted by the nonhydrolyzed substrate before the hydrolytic event takes place.

Investigation of protonation states of the catalytic acid/base Glu491 and the N1 atoms of the glucoimidazole moieties at the resolutions achieved in this work could not be deduced from the crystal structures of the complexes themselves. Only high-resolution atomic structures of less than 1 Å could provide unambiguous assignments of proton positions (*7*, *41*). In one such study with the *Bacillus agaradhearens* Cel5A 1,4-β-D-glucan endohydrolase (EC 3.2.1.1), it was shown that the proton is shared between the acid/base catalyst Glu139 and the exocyclic "glycosidic" N1 atom of the transition state mimic cellobiose-derived imidazole (*7*). An alternative approach to addressing the protonation states of the acid/base Glu491 and the N1 atoms of bound PheGlcIm and AmGlcIm is through quantum mechanical modeling, where suitable protonation models can be investigated in silico (Table 3). The lack of agreement with the experimental data for models **2** and **4** and the excellent agreement with the crystallographic experiment described in this study with models **1** and **3** indicate that the protonation state of the ligands is best represented by models where the catalytic acid/base Glu491 is negatively charged and the N1 atom is protonated (or positively charged), or where both interactive partners are neutral. Thus, model **1** and model **3** differ only in the position of a proton, which is either on the N1 atom of the ligand or on the Oϵ1 atom of the catalytic acid/base Glu491. This proton could potentially be shared between the Oϵ1 atom of Glu491 and the N1 atom of the inhibitors. In the gas phase at the HF/3-21G level of theory, model **1** lies approximately 30 kJ/mol lower in energy than model **3** (data not shown), although these calculations by themselves are not sufficient to distinguish between models **1** and **3**. It should be noted that the quantum mechanical predictions described in this work, based on the high-resolution crystal structures, support previously published computational data, which were based on the lower-resolution crystallographic data (*5*). The differences between the two sets of distances for models **1**−**4** account for less than 0.01−0.33 Å in each instance. However, it should also be noted that in the earlier quantum mechanical calculations, in which higher-level calculations were reported (*5*), the difference in energy between models **1** and **3** was significantly smaller than that calculated in this study.

The new interaction partners involved in binding of the aglycone portion of AmGlcIm at the interface between the −1 and +1 subsites suggest potentially novel organosynthetic avenues for future improvements in transition state analogue design (*10*, *12*, *40*). These considerations could be important for the design of highly efficient tailor-made transition state mimics of the barley β-D-glucan glucohydrolase, where at the +1 subsite the enzyme could harness binding energy from the aglycone portions of improved mimics. It might also be suggested that various classes of β-retaining glycoside hydrolases perform catalysis via slightly different and distinct transition state geometries, and therefore, in each instance, the design of transition state mimicry could be exclusive for a given hydrolase, as opposed to generalized transition state mimicry (*35*, *41*). This design could ultimately lead to specific transition state mimics as inhibitors of glycoside hydrolases that could be used as herbicides, and which could control vital biological events in life cycles of economically important embryophytes, where barley is classified.

ACKNOWLEDGMENT

We thank Dr. Harry Tong for advice during data collection at the BioCARS 14-ID beamline of the Advance Photon Source and Dr. Naohiro Matsugaki and Professor Soichi Wakatsuki at beamline BL5 of the Photon Factory.

REFERENCES

1. Varghese, J. N., Hrmova, M., and Fincher, G. B. (1999) Three-dimensional structure of a barley β-D-glucan exohydrolase: A family 3 glycosyl hydrolase, *Structure 7*, 179−190.
2. Coutinho, P. M., and Henrissat, B. (1999) Carbohydrate-active enzymes: An integrated database approach, in *Recent Advances in Carbohydrate Bioengineering* (Gilbert, H. J., Davies, G., Henrissat, B., and Svensson, B., Eds.) pp 3−12, The Royal Society of Chemistry, Cambridge, U.K.
3. Hrmova, M., and Fincher, G. B. (1998) Barley β-D-glucan exohydrolases. Substrate specificity and kinetic properties, *Carbohydr. Res. 305*, 209−221.
4. Hrmova, M., Varghese, J. N., De Gori, R., Smith, B. J., Driguez, H., and Fincher, G. B. (2001) Catalytic mechanisms and reaction intermediates along the hydrolytic pathway of plant β-D-glucan glucohydrolase, *Structure 9*, 1005−1016.
5. Hrmova, M., De Gori, R., Smith, B. J., Vasella, A., Varghese, J. N., and Fincher, G. B. (2004) Three-dimensional structure of the barley β-D-glucan glucohydrolase in complex with a transition state mimic, *J. Biol. Chem. 279*, 4970−4980.
6. Heightman, T. D., and Vasella, A. T. (1999) Recent insights into inhibition, structure, and mechanism of configuration-retaining glycosidases, *Angew. Chem., Int. Ed. 38*, 750−770.
7. Varrot, A., Schülein, M., Pipelier, M., Vasella, A., and Davies, G. J. (1999) Lateral protonation of a glycosidase inhibitor: Structure of the *Bacillus agaradhearens* Cel5A in complex with a cellobio-derived imidazole at 0.97 Å resolution, *J. Am. Chem. Soc. 121*, 2621−2622.
8. Notenboom, V., Williams, S. J., Hoos, R., Withers, S. G., and Rose, D. R. (2000) Detailed structural analysis of glycosidase/inhibitor interactions: Complexes of Cex from *Cellulomonas fimi* with xylobiose-derived aza-sugars, *Biochemistry 39*, 11553−11563.
9. Terinek, M., and Vasella, A. (2004) Synthesis of tetrahydropyridoimidazole-2-acetates: Effect of carboxy and methoxycarbonyl groups at C(2) on the inhibition of some β- and α-glycosidases, *Helv. Chim. Acta 87*, 3035−3049.

10. Shanmugasundaram, G., and Vasella, A. (2005) Synthesis of new *C*(2)-substituted *gluco*-configured tetrahydroimidazopyridines and their evaluation as glucosidase inhibitors, *Helv. Chim. Acta 88*, 2593−2602.
11. Ermert, P., Vasella, A., Weber, M., Rupitz, K., and Withers, S. G. (1993) Transition state analogue inhibitors of glycosidases are configurationally selective: A study with nojiritetrazoles, a new class of glycosidase inhibitors, *Carbohydr. Res. 250*, 113−128.
12. Nerinckx, W., Desmet, T., Piens, K., and Claeyssens, M. (2005) An elaboration on the *syn-anti* proton donor concept of glycoside hydrolases: Electrostatic stabilization of the transition state as a general strategy, *FEMS Lett. 302*, 302−312.
13. Panday, N., Canac, Y., and Vasella, A. (2000) Very strong inhibition of glucosidases by C(2)-substituted tetrahydroimidazopyridines, *Helv. Chim. Acta 83*, 58−79.
14. Hrmova, M., Varghese, J. N., Høj, P. B., and Fincher, G. B. (1998) Crystallization and preliminary X-ray analysis of β-glucan exohydrolase isoenzyme ExoI from barley (*Hordeum vulgare*), *Acta Crystallogr. D54*, 687−689.
15. Otwinowski, Z., and Minor, W. (1996) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
16. Collaborative Computational Project Number 4 (1994) The CCP4 suite: Programs for protein crystallography, *Acta Crystallogr. D50*, 760−763.
17. McRee, D. E. (1999) XtalView/Xfit: A versatile program for manipulating atomic coordinates and electron density, *J. Struct. Biol. 125*, 156−165.
18. Brünger, A. T. (1992) The Free *R* value: A novel statistical quantity for assessing the accuracy of crystal structures, *Nature 355*, 472−474.
19. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: A program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
20. Hehre, W. J., Radom, L., Schleyer, P. v. R., and Pople, J. A. (1986) *Ab Initio Molecular Orbital Theory*, pp 1−548, John Wiley & Sons, New York.
21. Schmidt, M. W., Baldridge, K. K., Boatz, J. A., Elbert, S. T., Gordon, M. S., Jensen, J. J., Koseki, S., Matsunaga, N., Nguyen, K. A., Su, S., Windus, T. L., Dupuis, M., and Montgomery, J. A. (1993) General atomic and molecular electronic structure system, *J. Comput. Chem. 14*, 1347−1363.
22. Winn, M. D., Isupov, M. N., and Murshudov, G. N. (2001) Use of TLS parameters to model anisotropic displacements in macromolecular refinement, *Acta Crystallogr. 57*, 122−133.
23. Yousef, M. S., Fabiola, F., Gattis, J. L., Somasundaram, T., and Chapman, M. S. (2002) Refinement of the arginine kinase transition-state analogue complex at 1.2 Å resolution: Mechanistic insights, *Acta Crystallogr. D58*, 2009−2017.
24. Painter, J., and Merritt, E. A. (2005) A molecular viewer for the analysis of TLS rigid-body motion in macromolecules, *Acta Crystallogr. D61*, 465−471.
25. Hrmova, M., De Gori, R., Smith, B. J., Fairweather, J. K., Driguez, H., Varghese, J. N., and Fincher, G. B. (2002) Structural basis for a broad specificity in higher plant β-D-glucan glucohydrolases, *Plant Cell 14*, 1−22.
26. Rayon, C., Lerouge, P., and Faye, L. (1998) The protein *N*-glycosylation in plants, *J. Exp. Bot. 49*, 1463−1472.
27. Quiocho, F. A. (1986) Carbohydrate-binding proteins: Tertiary structures and protein-sugar interactions, *Annu. Rev. Biochem. 55*, 287−315.
28. Fernandez-Alonso, M. C., Canada, F. J., Jimenez-Barbero, J., and Cuevas, G. (2005) Molecular recognition of saccharides by proteins. Insights on the origin of the carbohydrate-aromatic interactions, *J. Am. Chem. Soc. 127*, 7379−7386.
29. Harvey, A. J., Hrmova, M., De Gori, R., Varghese, J. N., and Fincher, G. B (2000) Comparative modeling of the three-dimensional structures of family 3 glycoside hydrolases, *Proteins: Struct., Funct., Genet. 41*, 257−269.
30. Schramm, V. L. (1998) Enzymatic transition states and transition state analogue design, *Annu. Rev. Biochem. 67*, 693−720.
31. Varfolomeev, S. D., Gariev, I. A., and Uporov, I. V. (2005) Catalytic sites of hydrolases: Structures and catalytic cycles, *Russ. Chem. Rev. 74*, 61−76.
32. Tatsuta, K., Miura, S., Ohta, S., and Gunji, H. (1995) Total syntheses of de-branched nagstatin and its analogs having glycosidase inhibiting activities, *Tetrahedron Lett. 36*, 1085−1088.
33. Tatsuta, K., Miura, S., Ohta, S., and Gunji, H. (1995) Syntheses and glycosidase inhibiting activities of nagstatin analogs, *J. Antibiot. 48*, 286−288.
34. Zechel, D. L., and Withers, S. G. (2000) Glycosidase mechanisms: Anatomy of a finely tuned catalyst, *Acc. Chem. Res. 33*, 11−18.
35. Vasella, A., Davies, G. J., and Böhm, M. (2002) Glycosidase mechanisms, *Curr. Opin. Chem. Biol. 6*, 619−629.
36. Vocadlo, D. J., Wicki, J., Rupitz, K., and Withers, S. G. (2002) Mechanism of *Thermoanaerobacterium saccharolyticum* β-xylosidase: Kinetic studies, *Biochemistry 41*, 9236−9246.
37. Wolfenden, R., and Snider, M. (2001) The depth of chemical time and the power of enzymes as catalysts, *Acc. Chem. Res. 34*, 938−945.
38. Cleland, W. W., Frey, P. A., and Gerl, J. A. (1998) The low barrier hydrogen bond in enzymatic catalysis, *J. Biol. Chem. 273*, 25529−25532.
39. Fersht, A. R., Shi, J. P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Cortes, P., Waye, M. M. Y., and Winter, G. (1985) Hydrogen bonding and biological specificity analysed by protein engineering, *Nature 314*, 235−238.
40. Lillelund, V. H., Jensen, H. H., Liang, X., and Bols, M. (2002) Recent developments of transition-state analogue glycosidase inhibitors of non-natural product origin, *Chem. Rev. 102*, 515−553.
41. Varrot, A., and Davies, G. J. (2003) Direct experimental observation of the hydrogen-bonding network of a glycosidase along its reaction coordinate revealed by atomic resolution analyses of endoglucanase Cel5A, *Acta Crystallogr. D59*, 447−452.
42. Hrmova, M., Harvey, A. J., Wang, J., Shirley, N. J., Jones, G. P., Stone, B. A., Høj, P. B., and Fincher, G. B. (1996) Barley β-D-glucan exohydrolases with β-D-glucosidase activity. Purification and determination of primary structure from a cDNA clone, *J. Biol. Chem. 271*, 5277−5286.
43. Fersht, A. (1999) *Structure and Mechanism in Protein Science*, pp 1−631, W. H. Freeman and Co., New York.
44. DeLano, W. L. (2002) *The PyMOL User's Manual*, DeLano Scientific, San Carlos, CA.
45. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. A47*, 110−119.

BI0514818