Active Site and Laminarin Binding in Glycoside Hydrolase Family 55*

Received for publication, November 5, 2014, and in revised form, February 27, 2015. Published, JBC Papers in Press, March 9, 2015, DOI 10.1074/jbc.M114.623579

Christopher M. Bianchetti†§¶, Taichi E. Takasuka*†, Sam Deutsch†, Hannah S. Udel†, Eric J. Yik**, Lai F. Bergeman‡, and Brian G. Fox†‡

From the †Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, Wisconsin 53706, the ‡Department of Chemistry, University of Wisconsin-Oshkosh, Oshkosh, Wisconsin 54901, the §Department of Agriculture, Hokkaido University, Sapporo, 060-8589, Japan, the ††Joint Genome Institute, Walnut Creek, California 94598, and the ‡‡Department of Chemistry and Biochemistry, California State University-Fullerton, Fullerton, California 92831

Background: SacteLam55A is a GH55 enzyme from highly cellulolytic Streptomyces sp. SirexAA-E.

Results: Substrate-bound structures identify residues involved in binding, catalysis, enforcement of reaction specificity, and possibly processivity.

Conclusion: Natural GH55 are exo-β,1,3-glucanases with a broad range of temperature and pH optima.

Significance: Experimental annotation of GH phylogenetic space by use of bioinformatics, high throughput cell-free translation, biochemical assay, and structure determination is feasible.

The Carbohydrate Active Enzyme (CAZy) database indicates that glycoside hydrolase family 55 (GH55) contains both endo- and exo-β,1,3-glucanases. The founding structure in the GH55 is PcLam55A from the white rot fungus Phanerochaete chrysosporium (Ishida, T., Fushinobu, S., Kawai, R., Kitaoka, M., Igarashi, K., and Samejima, M. (2009) Crystal structure of Phanerochaete chrysosporium. J. Biol. Chem. 284, 10100–10109). Here, we present high resolution crystal structures of bacterial SacteLam55A from the highly cellulolytic Streptomyces sp. SirexAA-E with bound substrates and product. These structures, along with mutagenesis and kinetic studies, implicate Glu-502 as the catalytic acid (as proposed earlier for Glu-663 in PcLam55A) and a proton relay network of four residues in activating water as the nucleophile. Further, a set of conserved aromatic residues that define the active site apparently enforce an exo-glucanase reactivity as demonstrated by exhaustive hydrolysis reactions with purified laminariligosaccharides. Two additional aromatic residues that line the substrate-binding channel show substrate-dependent conformational flexibility that may promote processive reactivity of the bound oligosaccharide in the bacterial enzymes. Gene synthesis carried out on ~30% of the GH55 family gave 34 active enzymes (19% functional coverage of the nonredundant members of GH55). These active enzymes reacted with only laminarin from a panel of 10 different soluble and insoluble polysaccharides and displayed a broad range of specific activities and optima for pH and temperature. Application of this experimental method provides a new, systematic way to annotate glycoside hydrolase phylogenetic space for functional properties.

Microbial communities that have symbiotic relationships with biomass-harvesting insects are now recognized to be a rich source of microbes with diverse metabolic and biosynthetic capabilities (1, 2). In addition to the well-documented cellulolytic activity of microbes that inhabit the digestive tracks of termites (3), newer paradigms for insect/microbe symbiotic interactions have been recently described in leaf cutter ants (1, 2, 4, 5), pine beetles (6), and wood wasps (7). For example, the highly invasive pinewood-boring wasp Sirex noctilio deposits a fungal/bacterial community into the pine tree when it lays eggs (8). If left unchecked, this insect/microbe community can cause massive destruction of pine forests (9). The white rot fungus Amelostereum aerolatum was the originally identified member of this community (10). Recently, however, a highly cellulolytic and hemicellulolytic Actinomycete, Streptomyces sp. SirexAA-E (SirexAA-E), was isolated (8). Genomic, transcriptomic, and biochemical studies showed that this microbe secretes an array of glycoside hydrolases (GHS)3 (7, 11, 12) and oxidative enzymes including AA10 lytic polysaccharide monooxygenases (6), peroxidases (7), and the first characterized caffeoyl-CoA intradiol dioxygenase (13) during growth on plant biomass.

The gene SACTE_4363, encoding a GH family 55 protein described hereafter as SacteLam55A, was of interest because it

* This work was supported by Department of Energy Great Lakes Bioenergy Research Center BER Office of Science Grant DE-FC02-07ER64494. The use of Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science Contract W-31-109-ENG-38. Use of the Life Science Collaborative Access Team at the Advanced Photon Source was supported by the College of Agricultural and Life Sciences, Department of Biochemistry, the Graduate School of the University of Wisconsin, the Michigan Economic Development Corporation, and Michigan Technology Tri-Corridor Grant 085P1000817.

† These authors contributed equally to this work.

‡ To whom correspondence should be addressed: Dept. of Biochemistry, 433 Babcock Dr., University of Wisconsin, Madison, WI 53706. Tel.: 608-262-9708; Fax: 608-262-3453; E-mail: bgfox@biochem.wisc.edu.

3 The abbreviations used are: GH, glycoside hydrolase; BTP, bis-tris propane; GH55, GH family 55; L2, laminaribiose; L3, laminaritriose; L4, laminaritetraose; L5, laminaripentaose; L6, laminarihexaose; RMSD, root mean square deviation; PDB, Protein Data Bank; DNS, dinitrilosalic acid; CHES, 2-(cyclohexylamino)ethanesulfonic acid; SBS, secondary binding site.
was up-regulated and secreted into the medium when Sir-
exAA-E was grown on cellulose, xylan, and various pretreated switchgrass samples but was not detected in the secreted pro-
teome when SirexaoAA-E was grown on glucose (7). Further-
more, although there are two genes encoding predicted GH55s in the 
SirexAA-E genome (SACTE_0324 and SACTE_4363), only SACTE_4363 (encoding SacteLam55A) was up-regulated 
during growth on biomass.

GH55 members are known as laminarinases because of their 
ability to hydrolyze laminarin, a β-1,3-glucan with occasional 
β-1,6 branching found in brown algae such as Laminaria digi-
tata (14, 15). They may also react with other β-1,3-glucans 
found in fungal cell walls (16), in callose-containing substruc-
tures of plant cell walls (17, 18), and in other sources. GH55 
enzymes hydrolyze β-1,3-glucan bonds (19) with inversion of 
stereochemistry at the anomeric carbon (20). The family is 
thought to contain members with either exo-β-1,3-glucanase 
(EC 3.2.1.58) or endo-β-1,3-glucanase (EC 3.2.1.39) activities. 
With the exception of single report on the cloning and 
sequence analysis of GluA from Arthrobacter sp. NHB-1 (21), 
previous published work on GH55 has been on fungal members 
(19, 20, 22–27). This includes the GH55 from Phanerochaete 
chrysosporium, previously named PcLam55A (20), whose 
structure has also been determined. However, the identifica-
tion of the catalytic residues, residues involved in substrate 
binding, and origins and extent of specificity of reaction in this 
family have not yet been established.

In this work, we used a combination of gene synthesis, cell-
free protein translation, catalytic assays, and x-ray crystallogra-
phy to provide a correlated biochemical and structural charac-
terization of the GH55 family. High resolution structures of 
substrate- and product-bound SacteLam55A revealed an 
extended substrate-binding cleft with six or more subsites for 
sugar binding formed at the interface of two right handed 
β-helical domains. Mutations guided by the structures support 
the identification of glutamate as the catalytic acid (Glu-502 in 
SacteLam55A) and a proton relay network that activates water 
to serve as the catalytic base. These structures also gave insight 
into residues that enforce an exolytic reaction, interactions that 
potentially support a processive hydrolysis of the bound oligo-
saccharide, and the ability of the enzyme to accommodate sub-
strates with β-1,6 branching as observed in natural β-1,3-glucan 
classes from diverse sources.

**EXPERIMENTAL PROCEDURES**

**Materials**—Laminaribiose (L2), laminaritriose (L3), laminar-
itetraose (L4), laminaripentaose (L5), laminarihexaose (L6), 
litchenan from Icelandic moss (mixed β-1,3–1,4 glucan), and 
D-mannan were purchased from Megazyme (Wicklow, Ire-
land). Sigmacell-20, insoluble β-1,3-glucan (degree of polymer-
ization greater than 5000), soluble laminarin from L. digita-
ta (β-1,3-glucan backbone with partial β-1,6 branching, degree of 
polymerization ~25, roughly one β-1,6 branch per polymer, and 
an ~3:1 ratio of polymers terminated with mannitol (M-ser-
ies) relative to polymers terminated with glucose (G-series); 
Refs. 14 and 15), beechwood xylan, wheat arabinoxylan, and 
lucost bean gum were purchased from Sigma-Aldrich. β-Pus-
tulan (β-1,6-glucan) from *Lasallia pustulata* with median 
degree of polymerization of ~100 was purchased from Invivo-
Gen (San Diego, CA). Soluble compounds (glucose, L3–L6) 
were used to determine structures of substrate-bound 
SacteLam55A E502A; laminarin and other insoluble polysac-
charides were used in functional screening of the enzymes 
produced by cell-free translation.

**Cloning, Expression, and Purification of SacteLam55A**—The 
SACTE_4363 gene (Uniprot G2NF9) encodes a 605-resi-
due polypeptide including a twin arginine translocation extra-
cellular secretion signal peptide (29) and the catalytic portion of 
the protein. SACTE_4363 was cloned into the *Escherichia coli* 
expression vector pVP67K (Center for Eukaryotic Structural 
Genomics, Madison, WI) using a two-step PCR method. 
pVP67K, a derivative of pVP68K described previously (30– 
32), produces a fusion protein that contains a tobacco etch virus 
protease cleavable N-terminal His8 tag. Polymerase incomplete 
primer extension (33) primer pairs were designed to amplify 
the GH55 domain without the predicted signal peptide, designated 
here as SacteLam55A (residues 45–605). The amplified DNA 
also contained partial overlap with flanking sequences that are 
present in pVP67K to assist in cloning (30). The amplification 
reaction mixture was transformed into *E. coli* 10G competent 
cells (Lucigen, Middleton, WI), and the nucleotide sequences of 
several recovered SacteLam55A-pVP67K plasmids were veri-
fied at the University of Wisconsin Biotechnology Center DNA 
sequencing facility. A sequence-verified plasmid was used as 
the starting template to generate nine additional plasmids con-
taining mutations in the SacteLam55A gene by PCR-based 
mutagenesis using primer pairs that contain either single or 
double nucleotide substitutions. The incorporation of the 
desired mutations was verified by nucleotide sequencing.

Sequence-verified plasmids were transformed into either 
*E. coli* B834 (selenomethionine labeling) or *E. coli* BL21 (DE3), 
respectively, for heterologous expression (31, 34). A single col-
ony from an overnight transformation was inoculated into 1 ml of 
chemically defined noninducing medium containing 50 
mg/liter kanamycin and 34 mg/liter chloramphenicol for 8 h 
at 37 °C (31). The 1-ml starter culture was inoculated into 50 
ml of the same noninducing medium. After 16 h at 25 °C, the 
50-ml culture was transferred into 1 liter of auto-induction 
medium containing 50 mg/liter kanamycin and 34 mg/liter 
chloramphenicol and incubated for 25 h at 25 °C. The auto-
inducing medium used to express selenomethionine-labeled 
SacteLam55A contained 0.01 g/ml of methionine (67 μM) and 
0.125 mg/ml of selenomethionine (637 μM), whereas the 
medium used to express the unlabelled, mutated forms of 
SacteLam55A contained 0.2 mg/ml of methionine (1.34 mM). 
Cells were grown in the auto-inducing medium for 25 h at 
25 °C, harvested by centrifugation at 5,000 × g for 15 min, and 
then stored at −80 °C.

All mutated forms of SacteLam55A were purified using an 
identical protocol. Pelleted cells were resuspended in 100 mM 
MOPS, pH 7.5, containing 500 mM NaCl and 10 mM imidazole 
and lysed by sonication. The cell lysate was clarified by cen-
trifugation at 20,000 × g for 60 min and loaded onto a 1.6-cm-
diameter × 2.5-cm HisTrap HP column (GE Healthcare) equil-
ibrated with 10 column volumes of 100 mM MOPS, pH 7.5, 
containing 500 mM NaCl and 10 mM imidazole. The proteins
were eluted with 100 mM MOPS, pH 7.5, containing 500 mM NaCl and 500 mM imidazole. The most pure fractions, as judged by visual inspection of an SDS-PAGE gel, were pooled and concentrated using a Vivaspin 20 (Sartorius AG, Germany). Residual imidazole was removed by loading the concentrated sample onto a 2.6-cm-diameter × 30-cm HisPrep desalting column (GE Healthcare) that had been equilibrated with 10 mM MOPS, pH 7.5, containing 50 mM NaCl. The N-terminal His tag was removed by treatment with tobacco etch virus protease, and the cleaved tag, uncleaved fusion enzyme, and tobacco etch virus protease were separated by subtractive nickel affinity chromatography (32). Tag-free SacteLam55A was concentrated to 20 mg/ml by ultrafiltration and exchanged into 10 mM MOPS, pH 7.0, containing 50 mM NaCl. The protein concentration was determined from absorbance at 280 nm using a molar absorption coefficient calculated from the amino acid sequence of the purified enzyme (35).

Bioinformatics—GH55 sequences were obtained from the Carbohydrate-Active enZyme database (CAZy) (36). The GH55 catalytic domain was identified by alignment of the SacteLam55A and PcLam55A (20) structures. After the boundaries of the GH55 catalytic domain were identified, 177 nonredundant GH55 sequences were trimmed to contain only the catalytic domain, and phylogenetic trees were generated from amino acid sequence alignments using Clustal Omega and maximum likelihood methods (37). All percentages reported herein are relative to this nonredundant set of sequences.

Gene Synthesis and Cell-free Protein Translation—GH55 family members were selected from phylogenetic trees, and the protein sequences were reverse translated with the goals of optimizing codon usage for protein expression and removal of sequence features commonly associated with gene synthesis failure (low/high GC, homopolymers, repeats, hairpins, etc.). The optimized sequences were then partitioned into intermediate fragments of <1000 bp and broken into overlapping oligonucleotides using GeneDesign (38). Oligonucleotides (150mers on average) were then passed to an automated assembly pipeline that included pooling, dilution, and two rounds of polymerase chain assembly performed by acoustic deposition using the Echo 555 instrument (Labcyte, CA). The resulting intermediate fragments were verified for correct assembly and purified using Ranger technology (Coastal Genomics, Burnaby, Canada) and cloned into a standard entry vector. For every target DNA fragment, eight colonies were sequence-verified using PACBIO RSII (Pacific Biosciences, Menlo Park, CA) and analyzed for direction of insertion and mutations using an internal variant detection algorithm. For final gene assembly, the corresponding error-free fragments were PCR-amplified, purified, and assembled with pEU-JGI using Gibson assembly (39). pEU-JGI is a derivative of pEU-HSBC that contains Smal restriction sites flanking the SgfI and Pmel delimited sacB/chloramphenicol acetyltransferase selection cassette (40). The full-length genes underwent an additional round of sequence verification, and the correct clones were archived in 96-well plates ready for DNA purification and cell-free translation.

Sequence-verified plasmids were purified as previously described, and the purified plasmids were adjusted to 1 μg/μl for use in coupled cell-free transcription and translation reactions (40). Translation reactions were analyzed by SDS-PAGE, and the concentration of each translated protein was quantitated using Gel Dock (Bio-Rad).

Enzyme Assays with Purified Enzyme—Purified SacteLam55A and mutated forms of the enzyme were screened for activity using the DNA assay with the substrates β-1,3-glucan, β-1,6-glucan, lichenan, and laminarin. The standard assay was performed in 100 μl of 50 mM phosphate buffer, pH 6.0, containing 10 mg/ml of the substrate at 50 °C. The reaction was initiated by addition of ~0.1 μg of purified enzyme and was incubated for up to 1 h at 50 °C. Reducing end products were determined using the dinitrosalicylic acid (DNS) assay (40, 41). After the desired time of reaction, 1 volume of the assay mixture was added to 2 volumes of DNS reagent, mixed, and then heated at 95 °C for 5 min. The amount of released reducing sugar was determined by monitoring the absorbance at 540 nm using glucose as a standard.

Steady-state kinetic parameters were estimated using varied amounts of soluble laminarin as the substrate (0, 1, 2, 5, 10, 15, and 20 mg/ml). In these reactions, ~0.02–0.1 μg of enzyme was mixed with varied amounts of laminarin up to 20 mg/ml in 100 μl of 50 mM phosphate buffer, pH 6.0, at 50 °C. Samples were taken from the assay mixture at 0, 1, 2, 4, 8, and 10 min to estimate the initial rate for reducing sugar formation, v, at each substrate concentration and to assess the time period over which the reaction rate was linear. A blank containing no enzyme was also prepared and used as a correction for the absorption of the DNS reagent and nonenzymatic hydrolysis. The initial rate, v, was determined from 2-min reactions for SacteLam55A carried out in triplicate, and reactions up to 10 min in duration were carried out in triplicate for mutated forms of the enzyme. In the reaction conditions described above, ~2–10% of the total substrate was converted to reducing sugar products depending on the activity of the enzyme. For the lowest activity mutated forms of SacteLam55A, the concentration of enzyme used in the reaction was increased by ~5-fold relative to SacteLam55A. Kinetic parameters were estimated by nonlinear least squares fitting to the expression $v = V_{\text{max}} [S] / (K_m + [S])$ using Prism 6.0 (GraphPad, La Jolla, CA).

End products (glucose and L2) from complete hydrolysis of pure laminariligosaccharides L2–L5 were separated by HPLC using a Rezex RPM-oligosaccharide column (200 × 10 mm) (Phenomenex, Torrance, CA) and detected by change in refractive index. Degassed Milli-Q water was used as the mobile phase with a flow rate of 0.3 ml/min at 85 °C. The area of each product peak was obtained by integration using EZstart software (Shimadzu Scientific Instruments, Columbia, MD) and compared with the areas given by injection of known amounts of glucose and L2 to estimate the molar amounts of products formed.

To determine the activity of the GH55 enzymes produced by cell-free translation (40), 10 μl of an individual translation reaction was added to a total reaction volume of 100 μl containing 10 mg/ml of either Sigmacell-20, β-1,3-glucan, β-1,6-glucan, lichenan, laminarin, beechwood xylan, arabinoxylan, d-mannan, or locust bean gum in 50 mM phosphate buffer, pH 6.0. The reactions were carried out at 50 °C for up to 24 h. A unit of activity is defined as the release of 1 μmol of reducing sugar per
minute from laminarin as determined by using the DNS assay (40, 41). For calculation of specific activity, the concentration of each translated protein was determined using SDS-PAGE and Gel Dock (Bio-Rad).

Temperature optima were determined by carrying out enzyme reactions from 30 to 90 °C in 10 °C increments. The reaction was started by mixing 4 μl of an individual translation reaction with 10 mg/ml of laminarin in 50 mM phosphate buffer, pH 6.0. After 10 min, reducing end products were determined using the DNS assay (40, 41). The pH optima were determined at 50 °C using the same enzyme and substrate concentrations as above for the temperature optima studies, but the buffer was varied as follows: 50 mM sodium citrate (pH 4.0 and 5.0); 50 mM phosphate (pH 6.0); Tris-HCl (pH 7.0 and 8.0); 50 mM CHES (pH 9.0 and 10.0); and 50 mM CAPS (pH 11.0). After the temperature optimum was determined, specific activities were estimated by reacting 4 μl of translation reaction with 10 mg/ml laminarin in 50 mM phosphate buffer, pH 6.0, for 10 min. The concentration of enzyme in the translation reactions was determined using SDS-PAGE and Gel Dock (Bio-Rad). Enzymes that did not show activity after 10 min were incubated up to 24 h and then assayed for products.

Crystalization—All purified SacteLam55A preparations were screened for crystallization using a Mosquito nanoliter liquid handling robot (TTP LabTech, Cambridge, MA). Crystals of SacteLam55A and SacteLam55A E502A were observed in the IndexHT (Hampton Research, Aliso Viejo, CA) and PACT premier HT (Molecular Dimensions, Altamonte Springs, FL) screens, respectively, and were further optimized. Crystals used to determine the SacteLam55A structure were grown using hanging drop vapor diffusion by mixing 1 μl of the protein solution (20 mg/ml) with 1 μl of 27% PEG 3350, 200 mM MgCl₂, and 100 mM BTP, pH 7.0, at 20 °C. Crystals used to determine the substrate-bound structures were grown by mixing 1 μl of SacteLam55A E502A with the following: (glucose bound) 1 μl of 20% PEG 3350, 125 mM sodium formate, 75 mM glucose, and 100 mM BTP, pH 6.5; (laminaritriose bound) 1 μl of 20% PEG 3350, 75 mM sodium formate, 25 mM laminaritriose, and 100 mM BTP, pH 6.0; (laminaritetraose bound) 1 μl of 22% PEG 3350, 50 mM sodium formate, 25 mM laminaritetraose, and 100 mM BTP, pH 6.5; and (laminarihexaose bound) 1 μl of 20% PEG 3350, 200 mM sodium acetate, 25 mM laminarihexaose, and 100 mM BTP, pH 6.0. Crystals were directly frozen in liquid N₂ after the addition of 15% (v/v) 1,2-ethanediol to the buffer, pH 6.0. After 10 min, reducing end products were determined using the DNS assay (40, 41). The pH optima were determined at 50 °C using the same enzyme and substrate concentrations as above for the temperature optima studies, but the buffer was varied as follows: 50 mM sodium citrate (pH 4.0 and 5.0); 50 mM phosphate (pH 6.0); Tris-HCl (pH 7.0 and 8.0); 50 mM CHES (pH 9.0 and 10.0); and 50 mM CAPS (pH 11.0). After the temperature optimum was determined, specific activities were estimated by reacting 4 μl of translation reaction with 10 mg/ml laminarin in 50 mM phosphate buffer, pH 6.0, for 10 min. The concentration of enzyme in the translation reactions was determined using SDS-PAGE and Gel Dock (Bio-Rad). Enzymes that did not show activity after 10 min were incubated up to 24 h and then assayed for products.

Crystallization—All purified SacteLam55A preparations were screened for crystallization using a Mosquito nanoliter liquid handling robot (TTP LabTech, Cambridge, MA). Crystals of SacteLam55A and SacteLam55A E502A were observed in the IndexHT (Hampton Research, Aliso Viejo, CA) and PACT premier HT (Molecular Dimensions, Altamonte Springs, FL) screens, respectively, and were further optimized. Crystals used to determine the SacteLam55A structure were grown using hanging drop vapor diffusion by mixing 1 μl of the protein solution (20 mg/ml) with 1 μl of 27% PEG 3350, 200 mM MgCl₂, and 100 mM BTP, pH 7.0, at 20 °C. Crystals used to determine the substrate-bound structures were grown by mixing 1 μl of SacteLam55A E502A with the following: (glucose bound) 1 μl of 20% PEG 3350, 125 mM sodium formate, 75 mM glucose, and 100 mM BTP, pH 6.5; (laminaritriose bound) 1 μl of 20% PEG 3350, 75 mM sodium formate, 25 mM laminaritriose, and 100 mM BTP, pH 6.0; (laminaritetraose bound) 1 μl of 22% PEG 3350, 50 mM sodium formate, 25 mM laminaritetraose, and 100 mM BTP, pH 6.5; and (laminarihexaose bound) 1 μl of 20% PEG 3350, 200 mM sodium acetate, 25 mM laminarihexaose, and 100 mM BTP, pH 6.0. Crystals were directly frozen in liquid N₂ after the addition of 15% (v/v) 1,2-ethanediol to the solutions described above.

Structure Determination—All diffraction data were collected at the Life Sciences Collaborative Access Team 21-ID-G Beamline at the Advanced Photon Source of Argonne National Laboratory (Argonne, IL). Diffraction data were processed, integrated, and scaled using HKL2000 (42). The selenomethionine-labeled SacteLam55A structure was initially phased to 1.9 Å using a single wavelength (0.97857 Å) with PHENIX AutoSol (43). Hyss (44) was used to determine the selenium substructure, and the results from Hyss indicated there were eight heavy atoms sites per asymmetric unit (four sites per monomer), which matched the expected number of selenomethionines from the amino acid sequence. The experimentally determined electron density map was of excellent quality and was used as input for PHENIX AutoBuild (45). The SacteLam55A E502A substrate bound structures were solved by molecular replacement with Phaser (46) using the SacteLam55A structure as a starting model. Because of the high resolution, the atomic displacement factors of SacteLam55A and SacteLam55A E502A laminaritriose structures were refined anisotropically. All SacteLam55A structures were completed with alternating rounds of model building in Coot (47) and refinement in PHENIX (48). Surface areas were calculated by PISA (49).

To examine the dynamic motion of several residues surrounding the active site, the glucose-bound and all of the substrate-bound SacteLam55A E502A structures were subjected to an ensemble refinement (50). We excluded the selenomethionine-labeled structure of the unbound enzyme from the ensemble refinement because of the presence of the anomalous signal from incorporated selenium. All model building and refinement steps were monitored using an R_free value based on 5.0% of the independent reflections. Model quality was accessed with Molprobity (51) before deposition to the PDB. Structural images were generated with the PyMOL Molecular Graphics System (version 1.5.0.4 Schrödinger, LLC).

RESULTS

Structure of SacteLam55A—The data collection, refinement, and model statistics are presented in Table 1. The structure of selenomethionine-labeled SacteLam55A was solved using single wavelength anomalous dispersion. The initial phased and density-modified electron density map was of excellent quality, and individual residues could be easily identified. The structure contained two SacteLam55A molecules in the asymmetric unit and belonged to the P2₁2₁2₁ space group. The structure was refined to a resolution of 1.51 Å, had an R_work (R_free) of 11.7% (16.1%), and had interpretable electron density that corresponded to residues 58–605 and 57–605 for monomers A and B, respectively. Overall, the two monomers were nearly identical and had a Ca root mean square deviation (RMSD) of 0.163 Å for residues 58–605. Unless noted, figures and descriptions are of the A monomer.

In addition to SacteLam55A, we determined the structure of SacteLam55A E502A, an inactive form of the enzyme, in complex with glucose, laminaritriose (L3), laminaritetraose (L4) and laminarihexaose (L6) to the resolutions and final R_cryst (R_free) values shown in Table 1. All substrate and product-bound structures were solved by molecular replacement using SacteLam55A (PDB code 4PEW) as an initial model. Unlike SacteLam55A, the SacteLam55A E502A structures belonged to the P2₁, space group. The L3- and L4-bound structures had one monomer per asymmetric unit, whereas the glucose and L6-bound structures had two monomers per asymmetric unit. Similar to SacteLam55A, the two monomers observed in the SacteLam55A E502A structures with either L6 or glucose bound were nearly identical and had Ca RMSDs of 0.071 and 0.108 Å, respectively. Whereas both monomers of the L6 structure had clear interpretable electron density for the substrate, density corresponding to glucose was found only in the B monomer of the SacteLam55A E502A structure. Regardless of the crystal form or unit cell parameters, all SacteLam55A monomers were equivalent and had an average Ca RMSD of 0.154 Å.
TABLE 1
Summary of crystal parameters, data collection, and refinement statistics for SacteLam55A and SacteLam55A E502A bound to product and substrates

The values in parentheses are for the highest resolution shell.

| Crystal parameters |
|-------------------|
| Space group       | P2₁,2,1         |
| Unit cell parameters (Å) | p3.97, b = 92.57, c = 198.3 |
| Molecules in asymmetric unit | 2 |

| Data collection statistics |
|---------------------------|
| Wavelength (Å)            | 0.97857          |
| Resolution range (Å)      | 37.93–1.51 (1.54–1.51) |
| Completeness (%)           | 90.5 (79.2)      |
| Rmerge²                   | 0.058 (0.306)    |
| Redundancy                | 11.5 (10.2)      |
| Mean I/σ(I)               | 51.88 (6.97)     |

| Refinement and model statistics |
|--------------------------------|
| Resolution range (Å)            | 37.93–1.51       |
| No. of reflections (work/test)  | 253,649/13,405   |
| R cryst                       | 0.117 (0.138)    |
| RMSD bond (Å)                  | 0.012            |
| RMSD angle (°)                 | 1.343            |
| Average B-factor (Å²)          | 16.1             |
| No. of protein atoms           | 8476             |
| No. of waters                  | 1427             |

| Ramachandran plot (%)          |
|--------------------------------|
| Favorable                      | 96.53            |
| Allowed                        | 2.9              |
| Disallowed                     | 0.18             |

| PDB code |
|----------|
| 4PEW     |
| 4PEX     |
| 4PEY     |
| 4PEZ     |
| 4PF0     |

* Rmerge = ∑|I(h) – <I(h)>| / ∑|I(h)|, where I(h) is the intensity of an individual measurement of the reflection, and <I(h)> is the mean intensity of the reflection.
* R cryst = ∑|Fcalc|–|Fobs| / ∑|Fcalc|, where Fcalc and Fobs are the observed and calculated structure-factor amplitudes, respectively.
* Rfree was calculated as R cryst using ~5% of randomly selected unique reflections that were omitted from the structure refinement.

Overall Structure—SacteLam55A is composed of two right-handed β-helical domains connected via an extended 49-residue linker (residues 258–306) that includes two antiparallel β-strands that distinguish SacteLam55A from the larger PcLam55A (752 residues). A hallmark of the right-handed β-helical fold is repetition of a three β-strand/turn motif, termed a coil (52). In SacteLam55A, there are seven and ten coils in the N-terminal and C-terminal domains, respectively (Fig. 1A). The N-terminal (residues 58–257, blue ribbon) and C-terminal (residues 307–605, red ribbon) domains pack against one another, burying ~1909 Å². The torsion angle between the β-helical domains is ~25°, which is similar to the domain arrangement in PcLam55A (20). A hallmark of the right-handed β-helical fold is repetition of a three β-strand/turn motif, termed a coil (52). In SacteLam55A, there are seven and ten coils in the N-terminal and C-terminal domains, respectively (Fig. 1A). Although the β-strands that comprise the N-terminal and C-terminal domains are similar, the length and amino acid composition of the loops are highly divergent. Additionally, the SacteLam55A N-terminal domain is truncated by ~100 residues when compared with the C-terminal domain. The smaller size (548 residues in PDB code 4PF0) and structural asymmetry distinguishes SacteLam55A from the larger PcLam55A (752 residues included in PDB code 3EQO), where both domains contain 12.
Structure and Functional Characterization of SacteLam55A

![Diagram](image)

**FIGURE 2. Structure of substrate-bound SacteLam55A E502A.** The substrate-binding cleft located at the interface between the N-terminal and C-terminal domains contains at least six sugar binding subsites (−1, +1, +2, +3, +4, and +5). Residues from the C-terminal domain form extensive interactions with the glucosyl groups bound in the −1 and +1 subsites. The position of the predicted catalytic acid (Glu-502, yellow sticks) was obtained from alignment of the SacteLam55A structure (PDB code 4PWE) with SacteLam55A E502A bound to L6 (PDB code 4PF0). Phe-153, Trp-444, and Trp-446 (purple sticks) form a “aromatic block” that caps one end of the substrate-binding cleft. Two residues (Phe-143 and Trp-144, orange sticks) occupy space that overlaps with a putative substrate binding cleft proposed for PcLam55A. The solvent-exposed C6 hydroxyl groups of the bound L6 are shown as green sticks. All C6 hydroxyl groups are fully exposed to solvent.

coils and are composed of a similar numbers of residues (20). The length of loops that connect the individual coils of the β-helical domains also contribute to the increased size of PcLam55A. These differences are shown in Fig. 1B.

**Substrate-binding Cleft—**The substrate-binding cleft of SacteLam55A is located at the interface between the N-terminal and C-terminal β-helical domains and does not contain any contribution from the linker joining them (Fig. 1A). The L6-bound structure revealed that six sugar-binding subsites comprise SacteLam55A substrate-binding cleft (Fig. 2). We have labeled these subsites from the reducing to nonreducing end (−1, +1, +2, +3, +4, and +5) using the previously developed formalism for GHs (53), where the hydroxylized β-1,3-glycosidic bond is located between the −1 and +1 subsites (Fig. 2). Similar positions are occupied in the L3- and L4-bound structures. In SacteLam55A E502A bound to glucose (PDB code 4PEX), the protein contacts to the ligand bound in the −1 site are identical to the contacts observed in the oligosaccharide-bound structures. Similar contacts are also observed in PcLam55A (PDB code 3EQO) bound to the product analog gluconolactone (20).

The asymmetry caused by the truncation of the N-terminal domain compared with the C-terminal domain and the torsion angle (~25°) between the two domains causes the substrate-binding cleft to adopt a curved conformation that follows the external contours of the N-terminal domain. This curvature matches the conformation of β-1,3-linked glucans (Fig. 2), which has been noted for GH16 ZgLamA (54) and excludes the linear conformation of β-1,4-linked polymers such as cellulose. Unlike GH16, GH17, and GH64 (54–56), where the substrate-binding clefts are open at both ends allowing endo-glucanase activity, the binding cleft of SacteLam55A is capped at one end by Phe-153, Trp-444, and Trp-446 (Fig. 2, purple sticks). This “aromatic block” closes the substrate-binding cleft and so restricts access to one direction along the domain boundary. Similar residues are observed in this position in aligned PcLam55A (20) and are conserved throughout the rest of the GH55 family (Figs. 2 and 3, purple sticks and surface). Thus steric constraints shown by the structures of substrate-bound enzyme and the distribution of end products (described below) show that SacteLam55A releases glucose from a β-1,3 linked oligosaccharide, making it an exo-β-1,3-glucanase (EC 3.2.1.58).

Most of the residues that interact with bound substrates are provided by the C-terminal domain. Additionally, the majority of these interactions are localized to the −1 subsite, where the terminal glucosyl group of substrate is bound (Fig. 4). In sum, every oxygen atom in the glucosyl group bound in the −1 subsite except O6 has at least one favorable hydrogen bonding interaction within 3 Å. Moreover, all of the residues that form the −1 subsite are highly conserved in both bacterial and fungal GH55s, suggesting that a similar binding mode will be present in all. Indeed, the residues from PcLam55A that form part of the active site by contacting the bound product analog gluconolactone (PDB code 3EQO; Ref. 20) are all conserved and occupy overlapping positions with the residues observed to interact with glucose in the SacteLam55A E502A structure.

The structures of SacteLam55A E502A bound to substrates provide additional information on the interactions with the enzyme. Interactions with sugars at the +1 and +2 subsites are limited when compared with the −1 subsite. For example, the glucosyl group at the +1 subsite is coordinated by Thr-149 and Gln-217 from the N-terminal domain and is in close proximity to Tyr-543 from the C-terminal domain. Although both Thr-149 and Gln-217 are strictly conserved in bacterial GH55 mem-
bers, Tyr-543 is not as well conserved in the bacterial GH55 members and can be tyrosine, phenylalanine, or asparagine. In addition to interacting with the +1 subsite, Gln-217 hydrogen bonds to the glucose moiety at the +2 subsite. The only other interaction with the +2 subsite is through the formation of a stacking interaction between bound substrate and Trp-196, which is conserved in bacterial GH55s but not in fungal GH55s (Fig. 2).

Two other striking features of the SacteLam55A substrate-binding cleft are the relatively few interactions with bound substrate at the +3, +4, and +5 subsites and its solvent accessibility. Unlike the −1, +1, and +2 subsites, which are partially occluded from solvent, the +3, +4, and +5 subsites are exposed to solvent (Fig. 2). There are no hydrogen bonding or aromatic residues that are close enough to interact with either the +3 or +4 subsites and the only interaction with the +5 subsite comes from a stacking interaction with Tyr-194 (Fig. 2), which is poorly conserved throughout the GH55 family.

SacteLam55A has no pocket adjacent to the −1 subsite comparable with the suggested β-1,6 branched sugar binding pocket in PcLam55 (20). Instead, Phe-143 and Trp-144 of SacteLam55A (Fig. 3, orange spheres) are adjacent to the −1 subsite and adopt a conformation that fills the space proposed for this putative branched sugar binding site. All C6 hydroxyl groups of bound L6, including the glucose moiety in the −1 subsite, are fully exposed to solvent and point away from the substrate-binding cleft (Fig. 2, green atoms). Indeed, Gln-150 (orange sticks) is the closest residue to the C6 hydroxyl group (~3.5 Å; short dashed line in Fig. 4). This accessibility accounts for the reactivity of the enzyme with laminarin, as the infrequent β-1,6 branching can be accommodated along the entire length of the substrate-binding cleft. Interestingly, Phe-143 and Trp-144 are not well conserved among the bacterial GH55 enzymes (Fig. 3), suggesting that structural variations are possible near the −1 subsite.

Dynamics of Substrate Binding—A comparison of all SacteLam55A structures revealed that the quality of the electron density for Tyr-194 and Trp-196 depended on the presence and length of bound substrate (Fig. 5, left panels). These differences in electron density suggested that Tyr-194, Trp-196, and other surrounding residues were mobile and might play a role in substrate binding. To investigate the dynamic nature of the aromatic residues in the substrate-binding cleft, the structures of SacteLam55A E502A with glucose, L3, L4, and L6 bound were refined using ensemble refinement (50). Ensemble refinements of crystals structures have been shown to extract protein dynamics from “static” crystal structures (57). In all cases, the SacteLam55A E502A ensemble structures had lower Rwork (Rfree) than the static structures (Table 2). The loop that contained Tyr-194 and Trp-196, which formed part of the substrate-binding cleft, displayed substrate-dependent flexibility (Fig. 5, right panels) that was not observed elsewhere in SacteLam55A.

These comparisons show that, in the absence of substrate, the side chains of Tyr-194 and Trp-196 lacked well defined electron density consistent with their ability to adopt numerous conformations (Fig. 5A). Furthermore, comparison of the elec-
tron density of Tyr-194 and Trp-196 in the A and B monomers of the glucose-bound structure (where glucose was either absent or present, respectively) showed that the presence of glucose did not improve the electron density around these two residues. Thus the presence of substrate at the +1 subsite was not sufficient to order either Tyr-194 or Trp-196.

In the L3-bound structure, where the –1, +1, and +2 subsites were occupied, Trp-196 interacted with substrate at the +2 subsite. The interaction at the +2 subsite helped to stabilize Trp-196, and this interaction led to a reduction in the number of discrete positions needed to model its electron density (Fig. 5B). Despite the lack of interactions between Tyr-194 and L3, a reduced number of positions were needed to account for this electron density also. A similar pattern was observed in the L4-bound structure (not shown). In the L6-bound structure, which had a fully occupied substrate-binding cleft that included interactions with both Trp-196 and Tyr-194, there was no appreciable change in the distribution of positions observed for Tyr-194 and Trp-196 when compared with the L3- and L4-bound structures (Fig. 5C). These results indicate that occupation of the –1, +1, and +2 sites is sufficient to orient these distal residues and help position the substrate for catalysis.

**Active Site Residues—**Inverting GHs use acidic and basic residues to activate another nucleophile to cleave glycosidic bonds (58, 59). In the following, the position of the predicted catalytic acid (Glu-502, yellow sticks) was obtained from alignment of SacteLam55A (PDB code 4PWE) with SacteLam55A E502A bound to L6 (PDB code 4PF0). There were three carboxylate-containing residues within 5 Å of the /H110011 subsite (Asp-449, Glu-480, and Glu-502; Figs. 4 and 6). These residues are strictly conserved throughout the GH55 family, and the corresponding residues adopt a similar conformation in PcLam55A (Ref. 20; see Fig. 6). Although all three of these residues were capable of interacting with a glucosyl group bound in the +1 subsite, only Glu-502 was capable of interacting with the anomeric oxygen (2.3 Å), indicating that it was likely the catalytic acid. Furthermore, although Tyr-505 was close to the anomeric oxygen (3.3 Å), it was better positioned to provide a hydrogen bond (2.7 Å) to O4 of glucose in the /H110012 position and to the carboxylate group of Glu-502. Upon considering this, the role of Tyr-505 may be to orient Glu-502 and the anomeric oxygen for better reaction. Neither Asp-449 nor Glu-480 could achieve a position relative to Glu-502 similar to that observed in other inverting GH families (58–60). Thus we explored other options for the nucleophile and identified a water (Fig. 6) positioned ~3.4 Å from the anomeric oxygen.

**FIGURE 5.** Conformational flexibility along the SacteLam55A substrate-binding cleft. A comparison of 2Fo – Fc electron density maps contoured at 1.5 σ (left column) and results of ensemble refinement of SacteLam55A (right column) in complex with glucose (A), laminaritriose (B), and laminarihexaose (C). In both columns, the SacteLam55A structures are shown using stick representation. Tyr-194 and Trp-196 (green sticks) interact with the glucosyl groups at the +5 and +3 subsites, respectively. Conformation flexibility is decreased as substrate occupies more of the SacteLam55A binding cleft.

**TABLE 2**

| Glucose | Laminaritriose | Laminarihexaose |
|---------|----------------|-----------------|
| Refinement parameters |
| Relaxation time (ps) | 0.6 | 0.8 | 0.6 | 0.8 |
| pTLS2 (%) | 0.8 | 0.8 | 0.8 | 0.8 |
| Conformers (#) | 25 | 25 | 25 | 25 |
| Refinement and model statistics |
| Resolution range (Å) | 24.27–1.75 | 24.23–1.50 | 28.34–1.90 | 30.32–1.75 |
| Rcryst | 0.112 (0.138) | 0.101 (0.1075) | 0.115 (0.112) | 0.124 (0.150) |
| Rfree | 0.168 (0.218) | 0.124 (0.1792) | 0.151 (0.186) | 0.163 (0.202) |
| ΔRcryst | -0.03 | -0.017 | -0.023 | -0.029 |
| ΔRfree | -0.025 | -0.017 | -0.018 | -0.019 |
| Mean RMSD per structure |
| Bonds (Å) | 0.0168 | 0.0163 | 0.0176 | 0.0181 |
| Angles (°) | 1.887 | 1.923 | 1.828 | 1.956 |
| Dihedral (°) | 16.467 | 15.239 | 15.610 | 16.060 |
| PDB code | 4TYV | 4TZI | 4TZ3 | 4TZ5 |

*pTLS is the percent of atoms that are included in the torsion libration screw (TLS) fitting procedure.*
from C1 of the glucosyl moiety in the −1 subsite as a likely candidate (58, 59). The side chains of Gln-174 and Ser-198 (two rotamers shown) and the backbone carbonyl oxygen of Thr-149, which are conserved in both bacterial and fungal GH55, coordinated a highly ordered water molecule in the active site (Fig. 6 and Ref. 20), whereas Glu-480 (also conserved) provided a stabilizing hydrogen bond to the side chain of Gln-174, Ser-198, and Thr-149 (orange sticks) to activate a water for attack on the anomeric carbon. The leaving group anionic oxygen is protonated by the catalytic acid Glu-502 (yellow sticks), which is stabilized by interaction with Tyr-505 (cyan sticks).

**Catalytic Reaction**—Among the polysaccharides tested, pure laminari-oligosaccharides and the natural product laminarin were the only substrates that were hydrolyzed by SacteLam55A. The end products from an exhaustive hydrolysis of pure laminari-oligosaccharides L5, L4, or L3 were a mixture of glucose and laminaribiose (L2) in a ratio consistent with removal of glucose groups until the remaining oligosaccharide was L2, which did not react. Thus L5 gave 3 mol of glucose and 1 mol of L2, L4 gave 2 mol of glucose and 1 mol of L2, whereas L3 gave 1 mol of glucose and 1 mol of L2. SacteLam55A had a maximal specific activity at 50 °C and pH 6 and maintained 70% or more of the maximal activity from 35 to 65 °C and from pH 6.0 to 9.0. Other kinetic parameters for SacteLam55A that reacted with soluble laminarin are shown in Table 3.

Guided by the structure, we tested active site residues that might participate in catalysis by generation of mutations at nine different residues in SacteLam55A. SacteLam55A with mutations at Glu-480 (E480A and E480Q) or Glu-502 (E502A and E502Q) was inactive. Mutations at Gln-174 (Q174A and Q174N), Asp-449 (D449A and D449N), or Tyr-505 (Y505A) had 102–104-fold reductions in $k_{cat}/K_m$ (Table 3), which arose from a combination of changes in both catalytic rate and binding affinity. This likely reflects subtle changes in positioning of the substrate and ability to activate the water nucleophile caused by these mutations. SacteLam55A S198A had $k_{cat}/K_m$ of 700 ± 200, which is ~13-fold lower than SacteLam55A, while retaining a similar apparent $K_m$ value (~0.6 mg/ml versus ~0.9 mg/ml), implicating its role in the proposed proton relay system as opposed to substrate binding.

**Functional Survey of the GH55 Family**—To date, the majority of studies on GH55 have been on fungal enzymes (19, 20, 22–27, 61). With the exception of a single report on the cloning and sequence analysis of GluA from *Arthrobacter* sp. NHB-1 (21), little is known about the many numerous bacterial members of this family. To investigate the natural diversity found within GH55, we generated a phylogenetic tree of all nonredundant GH55 sequences. The tree shows that the nonredundant GH55s (177 total, 102 bacterial, and 75 fungal) form two distinct clades (Fig. 7), with truncation of the N-terminal domain accounting for the separation of the bacterial and fungal clades. Additionally, the bacterial clade could be further subdivided into five subclades that are composed of either mostly *Actinobacteria* (subclades 1, 2, and 4) or *Firmicutes*, *Clostridia*, *Proteobacteria*, and others (subclades 3 and 5; Fig. 7 and Table 4). The sequences within bacterial subclades 1–4 had a Blast bit score of <200 ($E$ value $<10^{-49}$), respectively, indicating that they are closely related. Sequences in the fungal subclade had a Blast bit score of <50 ($E$ value $>10^{-9}$), indicating greater divergence, whereas bacterial subclade 5, a small out-group, shared little sequence identity with other bacterial GH55 enzymes (bit score >30, $E$ value >0.001). Although the active site residues described above were also found in subclade 5, the predicted lengths and positions of loops and gaps needed to generate the alignment accounted for the overall divergence of subclade 5 from SacteLam55A.

To probe the reactivity in the GH55 family, we identified 56 genes for synthesis (32% total coverage of the 177 nonredundant GH55), and 51 of these were successfully synthesized (29% total coverage, 45 bacterial GH55 and 6 fungal GH55). All 51 synthesized genes were successfully expressed as soluble proteins using cell-free translation (Fig. 7 and Table 4). The translated proteins were assayed using a panel of substrates (43), and 33 of the bacterial and 1 fungal GH55 were found to be active (19% functional coverage). Specifically, however, laminarin was the only substrate hydrolyzed by these 34 enzymes from 10

### Table 3

| Enzyme      | $V_{max}$ | $K_m$ | $k_{cat}/K_m$ | Relative $k_{cat}/K_m$ |
|-------------|-----------|-------|---------------|------------------------|
| SacteLam55A | 128 ± 4   | 0.93 ± 0.14 | 138 ± 3.9 | 148.1 ± 26.5 |
| E480A       | ND        | ND    | ND            | ND                     |
| E480Q       | ND        | ND    | ND            | ND                     |
| E502A       | ND        | ND    | ND            | ND                     |
| E502Q       | ND        | ND    | ND            | ND                     |
| Q174A       | 1.7 ± 0.2 | 16.9 ± 2.3 | 1.8 ± 0.2 | 0.11 ± 0.02 |
| Q174N       | 4.0 ± 0.2 | 8.1 ± 0.6 | 4.3 ± 0.2 | 0.53 ± 0.06 |
| S198A       | 6.1 ± 0.3 | 0.6 ± 0.1 | 6.6 ± 0.3 | 11.6 ± 3.8 |
| D449A       | 0.6 ± 0.1 | 5.3 ± 0.9 | 0.6 ± 0.1 | 0.12 ± 0.03 |
| D449N       | 0.8 ± 0.1 | 1.8 ± 0.3 | 0.8 ± 0.1 | 0.45 ± 0.09 |
| Y505A       | 6.2 ± 0.2 | 6.8 ± 0.4 | 6.7 ± 0.2 | 0.99 ± 0.06 |

* ND, no activity of enzyme was detected above the level of 0.001 μmol of reducing sugar produced after 24 h of reaction.
Structure and Functional Characterization of SacteLam55A

Figure 7. Phylogenetic tree of the GH55 family. GH55 was separated into five bacterial subclades (subclades 1–5 numbered around circle; light brown or no background) and a fungal subclade (green background). Sequences selected for gene synthesis, cell-free expression, and enzyme assays are indicated by red nodes in the tree. Specific activity at pH 6 and 50 °C and the pH optimum and the highest temperature where >70% of maximum activity was observed for each enzyme in a 2-h assay are mapped as concentric rings moving outward from the center of the tree. Published values from fungal GH55 are also included. The positions in the tree of P. chrysosporium Lam55A and SacteLam55A, for which structures are available, are marked with red stars.

Discussion—Inverting GHs use acidic and basic residues that are separated by ~10 Å to activate water to serve as the nucleophile in the hydrolysis reaction (58, 60, 63, 64). The role of the catalytic acid is to protonate the anomeric oxygen and thus stabilize the leaving hydroxyl group, whereas the catalytic base activates water for nucleophilic attack on the anomeric carbon (Fig. 6). The substrate-bound SacteLam55A structures show that Glu-502 is best positioned to interact with the anomeric oxygen at the −1 subsite as the general acid. The conservation of this residue throughout the entire GH55 family and the loss of all activity in SacteLam55A bearing the E502A and E502Q mutations support this functional assignment.

Neither Asp-449 nor Glu-480 (two other conserved carboxylates) adopt positions relative to Glu-502 that are similar to the active sites in other inverting glucanases (58, 60). Furthermore, it is unlikely that Asp-449 can function as a base to activate water because it is over 6 Å from the anomeric oxygen, has no water molecules in close proximity, and provides stabilizing interactions with O4 of the glucosyl group bound in the −1 subsite (Fig. 6) and with Trp-446, which is part of the “aromatic block.” Mutation of Asp-449 to either Asn or Ala decreased $k_{cat}/K_{m}$ by ~10$^3$ and ~10$^4$, respectively, with a larger proportional decrease in $k_{cat}$ than $K_{m}$. Therefore we propose Asp-449 has a primary role in positioning the substrate with consequences on the catalytic rate.

Glu-480 is positioned on the same side of the active site as Glu-502, and SacteLam55A mutants that lacked Glu-480 (E480A and E480Q) were completely inactive; thus Glu-480 has an important role in catalysis, and one possibility is to function as a general base in the catalytic mechanism by interaction with Gln-174, Ser-198, and the active site water proposed to be the nucleophile. Another possible function is to assist in placement of the substrate. Although Glu-480 is too distant to interact with the proposed water nucleophile directly, it does interact with both O2 and O3 and so assists in precise positioning of substrate in the −1 subsite. Glu-480 also participates in the proton relay network including the side chains of Gln-174 and Ser-198 and the carbonyl group of Thr-14. SacteLam55A S198A had ~8% residual activity, indicating that Ser-198 plays an important, but not essential, role in catalysis. Moreover, the Q174A mutation gave more severe defects in catalysis than S198A but nevertheless retained measurable activity.

Proton relay networks have been proposed for the inverting GH6 and GH97 families (63, 65). For example, the catalytic base (Asp-226) of exocellulase Cel6B from Thermobifida fusca activates water indirectly through a second water that is coordinated by a highly conserved Ser residue (63, 64, 66). Glu-480 of SacteLam55A appears to play a similar role as Asp-226 of Cel6B in the indirect activation of water. Additionally, in the mechanism proposed for SusB from Bacteroides thetaiotaomicron, a GH97 α-glucosidase, two Glu residues (Glu-439 and Glu-508) work together to activate the water nucleophile (65).

Protein Dynamics in the Substrate Binding Cleft—Comparisons of SacteLam55A structures showed that the quality of electron density for Tyr-194 and Trp-196 was dependent on the presence of substrate at the +2 subsite (Fig. 5). The interactions of these residues observed in the L3, L4, and L6 structures likely help retain the oligosaccharide for bond cleavage. After bond...
Table 4

| Clade | Source of organism | Protein | Average ± S.D. | Temperature\(^a\) | pH\(^b\) |
|-------|-------------------|---------|----------------|-------------------|---------|
|       |                   | µg/ml   | µmol sugar/mg enzyme/min | Optimum | Range | Optimum | Range |
| 1●    | ADW03597          | 152     | 26 ± 3.8          | 50          | 30–50 | 6.0      | 4.0–7.0 |
| 1     | AG57089           | 184     | 1.7 ± 0.2         | 50          | 50    | 5.0      | 4.0–7.0 |
| 1     | Sac3lam55A        | 100     | 42.0 ± 1.2        | 50          | 30–60 | 6.0      | 4.0–8.0 |
| 1     | BAG19231          | 139     | 35.9 ± 3.7        | 40          | 30–60 | 6.0      | 4.0–8.0 |
| 1     | AGK79825          | 136     | 35.2 ± 3.5        | 30          | 30–60 | 6.0      | 4.0–8.0 |
| 1     | ADW70102          | 145     | 11.0 ± 0.8        | 40          | 30–60 | 6.0      | 4.0–7.0 |
| 1     | BAG61520          | 210     | 21.9 ± 1.9        | 50          | 40–60 | 5.0      | 4.0–8.0 |
| 1     | AGK75398          | 293     | 17.3 ± 1.4        | 50          | 30–60 | 6.0      | 5.0–8.0 |
| 1     | AEY93509          | 191     | 56.1 ± 1.6        | 50          | 30–60 | 6.0      | 4.0–8.0 |
| 1     | CCA35917          | 147     | 49.1 ± 6.5        | 40          | 30–50 | 5.0      | 4.0–8.0 |
| 1     | CBZ21676          | 215     | 33.7 ± 2.4        | 40          | 40–60 | 5.0      | 4.0–8.0 |
| 1     | ADJ12502          | 189     | 37.3 ± 2.9        | 40          | 30–60 | 6.0      | 4.0–8.0 |
| 1     | ACZ49991          | 624     | 0.3 ± 0.1         | 50          | 40–50 | 6.0      | 4.0–7.0 |
| 1     | ACY14033          | 272     | 10.2 ± 0.8        | 40          | 30–50 | 7.0      | 4.0–8.0 |
| 1     | BAL78999          | 121     | 52.5 ± 0.5        | 60          | 30–60 | 7.0      | 4.0–8.0 |
| 1     | AEA62613          | 172     | 52.2 ± 1.7        | 40          | 30–40 | 7.0      | 4.0–8.0 |

\(^a\) The optima and effective range for temperature and pH defined as retention of at least 70% of maximum activity in a standard assay, are also shown.

\(^b\) ND indicates no enzyme activity was detected above the level of 0.001 µmol of reducing sugar produced after 24 h of reaction.

\(^c\) Activity observed above the level of 0.001 µmol of reducing sugar produced after 24 h.

\(^d\) Cloning and sequence analysis of GluA from Arthrobacter sp. NHB-1 have been previously reported (21).

\(^e\) Biochemical and structural studies of PcLam55A purified from the natural host were reported elsewhere (20).

Cleavage, glucose can be released through the open access to solvent that is adjacent to the −1 subsite, leaving it available to rebind substrate.

Aromatic residues contribute to the processivity of glycosyl hydrolases by providing dynamic interactions with a bound polysaccharide chain. GH55 lacks a tunnel as in GH6 (67) and GH7 (68, 69) and also lacks the extensive track of aromatic residues seen in processive GH18 chitinases (70). Although speculative, the dynamic motion of Tyr-194 and Trp-196 suggested by the ensemble refinements could allow the remaining bound oligosaccharide to processively slide along the substrate-binding cleft to place a new sugar into the −1 subsite without being released from the active site cleft. If this were so, the flexibility of Tyr-194 and Trp-196 could play an important role in the GH55 reaction beyond simple sequestration of L3 and longer laminarioligosaccharides.

**Secondary Binding Site**—During the final stages of refining the L3-, L4-, and L6-bound SacteLam55A structures, clear and interpretable electron density corresponding to a bound oligosaccharide was identified on the N-terminal domain (Fig. 8). An
Structure and Functional Characterization of SacteLam55A

Additional GH families have been shown to contain SBS, and insoluble xylan substrates and may play a role in localizing to the active site as the SBS of SacteLam55A (71). Ludwiczek et al. (71) showed that the SBS of BcX binds both soluble and insoluble xylan substrates and may play a role in localization. Additional GH families have been shown to contain SBS, and it has been proposed that these auxiliary binding sites affect substrate targeting, processivity, and activity (72). It is possible that the SBS of SacteLam55A plays a role in localizing SacteLam55A to the surface of insoluble laminarins in a manner similar to a carbohydrate-binding module (73).

The residues that compose the secondary binding site are conserved in subclade 1 (primarily Streptomyces), of which SacteLam55A is a member. Otherwise, the residues that compose the secondary binding site are not conserved within the GH55 family. For example, PcLam55A has a helix that overlays in space with part of the subclade 1 secondary binding site. Interestingly, however, PcLam55A Tyr-135, which is located at the start of this helix, was suggested to play a role in substrate binding (20).

Other Members of GH55—We used biochemical assays to provide a broader functional annotation of the GH55 family (Fig. 7 and Table 4). Laminarin was the only substrate that gave a positive assay response from the enzymes tested. The enzyme from Streptomyces hygroscopicus subsp. jinggangensis TL01 (GenBank™ locus AEY93509) had the highest specific activity of all enzymes tested for reaction with soluble laminarin at pH 6 and 50 °C.

Subclade 1 and 2 enzymes, primarily from Actinomycetales, had similar ranges of optima for pH, temperature. The clad 3 enzymes, from a broader distribution of bacterial genera and also showed more diversity in specific activity and pH optimum, and their temperature optima were generally lower than observed for the subclade 1 enzymes. In subclade 4, also primarily from Actinomycetales, the enzymes from soil organisms Arthrobacter phenanthrenivorans Sphe3 (ADX71276) (74), Deinococcus peraridilitoris DSM19664 (AFZ68118), and Micromonospora sp. L5 (ADU06434) (75) had the highest temperature optima and largest range of pH stabilities.

Six fungal GH55s were also expressed and assayed, and only PcLam55A showed activity with laminarin. Because there are eight disulfide bonds in PcLam55A (20), incorrect formation of disulfide bonds during cell-free translation may have contributed to the low number of active fungal GH55 detected.

Conclusion—This study of SacteLam55A provides an in-depth characterization of the structure and function of the GH55 family. By use of mutagenesis and kinetic analyses, it provides new insights into the catalytic residues and also the residues that form an extended substrate-binding cleft. The combination of gene synthesis, cell-free translation and assays using a diagnostic panel of substrates across the entire GH55 represents, to our knowledge, the most complete functional mapping of an entire GH family available to date. Biochemical assays show that laminarioligosaccharides longer than two glucose units long react exclusively by an exo-glucanase reaction, and crystal structures of the enzyme substrate complexes have provided a molecular level understanding of its origin.

REFERENCES

1. Ramadhar, T. R., Beemelmanns, C., Currie, C. R., and Clardy, J. (2014) Bacterial symbionts in agricultural systems provide a strategic source for antibiotic discovery. J. Antibiot. 67, 53–58
2. Aylward, F. O., Suen, G., Biedermann, P. H., Adams, A. S., Scott, J. J., Malfatti, S. A., Glavina del Rio, T., Tringe, S. G., Poulsen, M., Raffa, K. F., Klepzig, K. D., and Currie, C. R. (2014) Convergent bacterial microbiotas
Structure and Functional Characterization of SacteLam55A

45. Terwilliger, T. C., Grosse-Kunstleve, R. W., Afonine, P. V., Moriarty, N. W., Zwart, P. H., Hung, L.-W., Read, R. J., and Adams, P. D. (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. Acta Crystallogr. D Biol. Crystallogr. 64, 61–69

46. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674

47. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132

48. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221

49. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystallographic state. J. Mol. Biol. 372, 774–797

50. Burnley, B. T., Afonine, P. V., Adams, P. D., and Gros, P. (2012) Modelling dynamics in protein crystal structures by ensemble refinement. Elife 1, e00311

51. Burnley, B. T., Afonine, P. V., Adams, P. D., and Gros, P. (2012) Modelling dynamics in protein crystal structures by ensemble refinement. Elife 1, e00311

52. Jenkins, J., and Pickersgill, R. (2001) The architecture of parallel b-helices and related folds. Prog. Biophys. Mol. Biol. 77, 111–175

53. Davies, G. J., Wilson, K. S., and Henrissat, B. (1997) Nomenclature for carbohydrate-binding modules. FEBS Lett. 403, 274–279

54. Davies, G. J., Wilson, K. S., and Henrissat, B. (1997) Nomenclature for carbohydrate-binding modules. FEBS Lett. 403, 274–279

55. Wojtkowiak, A., Witek, K., Hennig, J., and Jaskolski, M. (2012) Structures of an active-site mutant of a plant 1,3-glucanase in complex with oligosaccharides. Acta Crystallogr. D Biol. Crystallogr. 68, 21–28

56. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132

57. Levin, E. J., Kondrashov, D. A., Wesenberg, G. E., and Phillips, G. N., Jr. (2014) The 2.78 Å crystal structure of an 1,3-glucanase from Zobellia galactanivorans. Acta Crystallogr. D Biol. Crystallogr. 70, 383–393

58. Levin, E. J., Kondrashov, D. A., Wesenberg, G. E., and Phillips, G. N., Jr. (2014) The 2.78 Å crystal structure of an 1,3-glucanase from Zobellia galactanivorans. Acta Crystallogr. D Biol. Crystallogr. 70, 383–393

59. Rye, C. S., Wilson, K. S., and Henrissat, B. (2000) Glycosidase mechanisms. Eur. J. Biochem. 267, 3435–3444

60. Damude, H. G., Withers, S. G., Kilburn, D. G., Miller, R. C., Jr., and Warren, R. A. (1995) Site-directed mutation of the putative catalytic residues of endoglucanase CenA from Cellulomonas fimii. Biochemistry 34, 2220–2224

61. Kawai, R., Igarashi, K., and Samejima, M. (2006) Gene cloning and heterologous expression of glycoside hydrolase family 55 β-1,3-glucanase from the basidiomycete Phanerochaete chrysosporium. Biotechnol. Lett. 28, 365–371

62. Kim, D. J., Baek, J. M., Uribe, P., Kenerley, C. M., and Cook, D. R. (2002) Cloning and characterization of multiple glycosyl hydrolase genes from Thermotoga maritima. J. Bacteriol. 184, 2121–2129

63. Koivula, A., Ruohonen, L., Wohlfiart, G., Reinikainen, T., Teeri, T. T., Piens, K., Claeyssens, M., Weber, M., Vasella, A., Becker, D., Sinnott, M. L., Zou, J. Y., Kleywegt, G. J., Szardenings, M., Stähler, J., and Jones, T. A. (2002) The active site of cellobiohydrolase Cel6A from Trichoderma reesi: the roles of aspartic acids D221 and D175. J. Am. Chem. Soc. 124, 10015–10024

64. Vuong, T. V., and Wilson, D. B. (2010) Glycoside hydrolases: catalytic base/nucleophile diversity. Biotechnol. Bioeng. 107, 195–205

65. Vuong, T. V., and Wilson, D. B. (2010) Glycoside hydrolases: catalytic base/nucleophile diversity. Biotechnol. Bioeng. 107, 195–205

66. Vuong, T. V., and Wilson, D. B. (2009) The absence of an identifiable single catalytic base residue in Thermobifida fusca exocellulase Cel6B. FEBS Lett. 579, 3837–3845

67. Vuong, T. V., and Wilson, D. B. (2009) The absence of an identifiable single catalytic base residue in Thermobifida fusca exocellulase Cel6B. FEBS Lett. 579, 3837–3845

68. Payne, C. M., Bomble, Y. J., Taylor, C. B., McCabe, C., Himmel, M. E., Crowley, M. F., and Beckham, G. T. (2011) Multiple functions of aromatic-carbohydrate interactions in a processive cellulase examined with molecular simulation. J. Biol. Chem. 286, 41028–41035

69. Nakamura, A., Tsukada, T., Auer, S., Furuta, T., Wada, M., Koivula, I., Igarashi, K., and Samejima, M. (2013) The tryptophan residue at the active site tunnel entrance of Trichoderma reesei cellobiohydrolase Cel7A is important for initiation of degradation of crystalline cellulose. J. Biol. Chem. 288, 13503–13510

70. Taylor, C. B., Payne, C. M., Himmel, M. E., Crowley, M. F., McCabe, C., and Beckham, G. T. (2013) Binding site dynamics and aromatic-carbohydrate interactions in processive and non-processive family 7 glycoside hydrolases. J. Phys. Chem. B 117, 4924–4933

71. Zakariassen, H., Aam, B. B., Horn, S. J., Vårum, K. M., Sørlie, M., and Eijssink, V. G. (2009) Aromatic residues in the catalytic center of chitinase A from Serratia marcescens affect processivity, enzyme activity, and biomass converting efficiency. J. Biol. Chem. 284, 10610–10617

72. Ludwikiewicz, M. L., Heller, M., Kattner, T., and McIntosh, L. P. (2007) A secondary xylan-binding site enhances the catalytic activity of a single-domain family 11 glycoside hydrolase. J. Mol. Biol. 373, 337–354

73. Cockburn, D., Wilkens, C., Ruzanski, C., Andersen, S., Nielsen, J. W., Smith, A. M., Field, R. A., Willemoes, M., Abou Hachem, M., and Svensson, B. (2014) Analysis of surface binding sites (SBSs) in carbohydrate active enzymes with focus on glycoside hydrolase families 13 and 77-a mini-review. Biologia 69, 705–712

74. Boraston, A. B., Bolam, D. N., Gilbert, H. J., and Davies, G.J. (2004) Carbohydrate-binding modules: fine-tuning polysaccharide recognition. Biochem. J. 382, 769–781

75. Kalimann, A., Kavakiotis, K., Perisynakis, A., Spröer, C., Pukall, R., Drai, N., and Rügendorf, U. (2009) Arthrobacter phenanthrenivorans sp. nov., to accommodate the phenanthrene-degrading bacterium Arthrobacter sp. strain Sphe3. Int. J. Syst. Evol. Microbiol. 59, 275–279

76. Igarashi, Y., Trujillo, M. E., Martínez-Molina, E., Yanase, S., Miyanaga, S., Obara, T., Sakurai, H., Saito, F., Fujita, T., and Furumai, T. (2007) Antitumor anthraquinones from an endophytic actinomycete Micromonospora lupini sp. nov. Bioorg. Med. Chem. Lett. 17, 3702–3705