Structure
Article

SusG: A Unique Cell-Membrane-Associated α-Amylase from a Prominent Human Gut Symbiont Targets Complex Starch Molecules

Nicole M. Koropatkin[1,2] and Thomas J. Smith[1,*]
[1]Donald Danforth Plant Science Center, St. Louis, MO 63132, USA
[2]Present address: Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, MI 48109, USA
*Correspondence: tsmith@danforthcenter.org
DOI 10.1016/j.str.2009.12.010

SUMMARY

SusG is an α-amylase and part of a large protein complex on the outer surface of the bacterial cell and plays a major role in carbohydrate acquisition by the animal gut microbiota. Presented here, the atomic structure of SusG has an unusual extended, bilobed structure composed of amylase at one end and an unprecedented internal carbohydrate-binding motif at the other. Structural studies further demonstrate that the carbohydrate-binding motif binds maltooligosaccharide distal to, and on the opposite side of, the amylase catalytic site. SusG has an additional starch-binding site on the amylase domain immediately adjacent to the active cleft. Mutagenesis analysis demonstrates that these two additional starch-binding sites appear to play a role in catabolism of insoluble starch. However, elimination of these sites has only a limited effect, suggesting that they may have a more important role in product exchange with other Sus components.

INTRODUCTION

The trillions of microbes inhabiting the human distal gut have a profound effect on human health. The indigenous microbial flora (microbiota), which outnumbers human cells by several orders of magnitude (Hooper and Gordon, 2001), shields the intestinal tract from pathogen colonization and promotes maturation and proliferation of gut cells (Mazmanian et al., 2005). Systemically, the microbiota stimulates the development of the immune system and may offer protection from allergic inflammatory responses such as asthma (Noverr and Huffnagle, 2004; Penders et al., 2007; Umetsu et al., 2002). The microbiota functions as a metabolic organ with enzymatic properties that enhance or supercede our own, such as the ability to degrade resistant dietary or host-derived glycans that transit the distal gut (Bjursell et al., 2006; Martens et al., 2008; Sonnenburg et al., 2005; Xu et al., 2003). Through this symbiotic activity, the microbiota also supplies nutrients to the animal host. For example, short-chain fatty acids provided from the bacterial fermentation of glycans can account for as much as 10% of the daily caloric intake for individuals on a Western-style, carbohydrate-rich, diet (Backhed et al., 2005). Thus, a thorough understanding of the metabolic capabilities of the microbiota will provide significant insight into our own biochemical makeup and may lead to better strategies for manipulating human nutrition and the treatment of colon-related diseases.

Colonic *Bacteroides* species account for nearly 45% of all bacterial species in the human gut microbiota (Ley et al., 2006) and harvest a vast array of dietary and host-derived glycans via outermembrane protein complexes that capture, degrade, and import polysaccharides (Bjursell et al., 2006; Martens et al., 2008; Sonnenburg et al., 2005; Xu et al., 2003). The genes encoding these proteins are clustered together in similarly patterned polysaccharide utilization loci (PULs) and include one or more glycolytic enzymes as well as homologs of the proteins SusC and SusD involved in glycan recognition and import. The *s*tarch *u*tilization *s*ystem (Sus) of the prominent human gut symbiont *Bacteroides thetaiotaomicron* was the first such PUL to be described (Anderson and Salyers, 1989a, 1989b). The Sus system of *B. thetaiotaomicron* is composed of eight genes, *susRABCDEFG* (Figure 1). SusR is a transcriptional regulator that turns on the expression of the other seven *sus* genes in response to maltooligosaccharides, amylose, amylopectin, and pullulan (D'Elia and Salyers, 1996b). SusCDEFG are localized to the cell surface and likely form a complex that processes and imports starch (Anderson and Salyers, 1989b; Cho and Salyers, 2001; Shipman et al., 1999, 2000; Tancula et al., 1992). SusDEFG are lipoproteins tethered to the outer surface of the cell, whereas SusC is predicted to be a TonB-dependent, β barrel porin. Unlike other TonB-dependent porins characterized to date, SusC cannot bind ligand alone and requires the starch-binding protein SusD for starch import (Cho and Salyers, 2001). We recently determined the atomic structure of SusD and demonstrated that its binding to starch molecules is driven by recognition of the overall three-dimensional shape of the ligand rather than by individual moieties (Koropatkin et al., 2008). Therefore, SusD likely plays a critical role in targeting polymeric starch to the Sus complex and may facilitate movement of linear oligosaccharides to the SusC porin. SusA and SusB are a periplasmic neopullulanase and an α-glucosidase, respectively, that presumably break down smaller maltooligosaccharides (D'Elia and Salyers, 1996a; Kitamura et al., 2008).

SusG is the α-amylase expressed concomitantly with SusCDEF on the outer surface of the cell and is absolutely required for growth on starch. α-amylases are members of glycoside

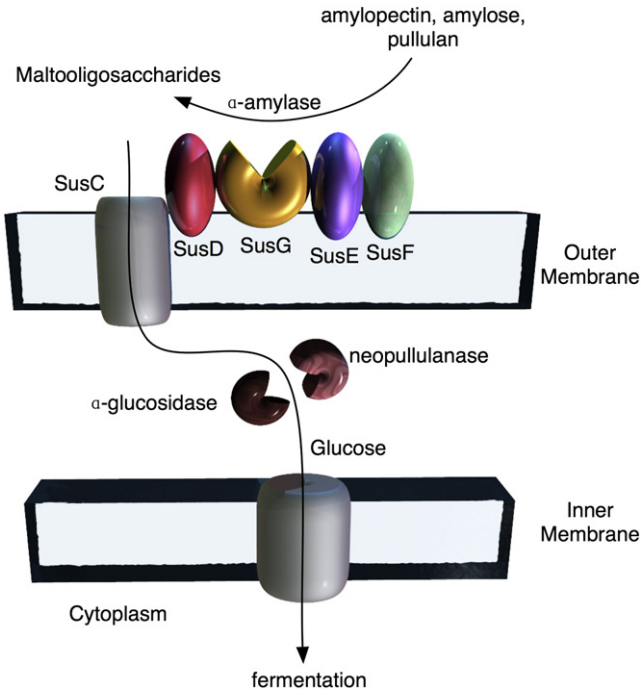

Figure 1. The Starch Utilization System of *Bacteroides thetaiotaomicron*

Cartoon representation of *Sus* operon protein products (Cho and Salyers, 2001; D'Elia and Salyers, 1996a; Shipman et al., 1999, 2000). The stoichiometry of the various proteins in the Sus complex is not known. SusD is a starch-binding protein of known structure (Koropatkin et al., 2008).

hydrolase family 13 (GH13), one of largest families of carbohydrate-active enzymes (http://afmb.cnrs-mrs.fr/cazy/index.html). Members of this family include amylases, α-glucosidases, (neo)-pullulanases, and cyclodextrin glucosyltransferases, many of which have been extensively studied structurally and biochemically. These enzymes share a highly conserved $(\alpha/\beta)_8$ barrel core structure and an enzymatic mechanism featuring a double-displacement, general acid/base catalytic scheme that retains stereochemistry at the anomeric carbon.

SusG is essential for the growth of *B. thetaiotaomicron* on starch, despite the presence of four other (non-Sus-associated) predicted amylases in its genome. In previous work, *susG* deletion mutants could still bind starch at the cell surface, but could not grow on amylopectin or pullulan (Shipman et al., 1999). Although the Sus genes are not required for growth on maltose or maltotriose, a *ΔsusABCDEFG* mutant complemented with SusG cannot grow on amylose or amylopectin even though such complementation restored extracellular starch-degrading activity (Shipman et al., 1999). Therefore, SusG may have evolved to work as part of a carbohydrate-processing/import complex rather than just as an outer-membrane amylase. Initial investigation of the SusG amino acid sequence revealed an internal stretch of amino acids (residues 190–360) with no identifiable sequence homology to the many well-characterized GH13 enzymes or carbohydrate-binding modules (CBMs) (Altschul et al., 1997).

Since the discovery of the Sus complex in *B. thetaiotaomicron*, 87 similar Sus-like PULs have been identified in this species, comprising 18% of the genome. Whole-genome transcriptional profiling suggests that each of the Sus-like PULs targets different glycans with only minor redundancy (Martens et al., 2008; Sonnenburg et al., 2005). An additional 269 Sus-like PULs have been identified in four other human gut isolates and many others have been identified in nongut environmental Bacteroidetes, suggesting that the Sus-like complexes represent a paradigm for glycan uptake in these bacteria (Xu et al., 2007). However, despite the predominance of these complexes in Bacteroidetes, little is understood about the individual protein components of any particular complex, or how these proteins work together to import glycans. To that end, the structure and biochemical properties of another Sus component, SusG, are presented here. The protein has an extended bilobed architecture with a novel CBM at one end and the amylase at the other. From structural and biochemical analysis, it seems likely that this unusual domain organization is designed not only for digestion of large, insoluble starch molecules but also for the retention of oligosaccharides by the Sus complex for passage into the cell. Therefore, it seems possible that substrate specificity in these nutrient acquisition systems is not only governed by the details of carbohydrate-protein interactions but also in the manner that these protein modules are assembled.

RESULTS

SusG Apo Structure

The SeMet-substituted structure of apo SusG (residues 24–692) was determined using single-wavelength anomalous dispersion (SAD) to 2.2 Å (R_{work} = 19.6%, R_{free} = 23.0%). The final model contained two molecules of SusG (residues 43–692 and 44–692 were observed) in the asymmetric unit in addition to a number of small ligands: two Mg^{2+}, two Ca^{2+}, 13 ethylene glycol, 5 acetate, 1 PEG, and 467 water molecules. Native PAGE analysis suggested that SusG is a monomer in solution (data not shown).

SusG is composed of A, B, and C domains that share structural features with other amylases (Figure 2A). The A domain (residues 43–152 and 364–607) has an eight-stranded α/β barrel that contains the catalytic site, with the B domain (residues 153–215 and 336–363) inserted between β3 and α3 of the A domain. The B domain consists of two two-stranded antiparallel β sheets, two α helices, and three 3_{10} helices that pack against the A domain and contribute to the size and accessibility of the active site. The C domain (residues 608–692) folds into an eight-stranded β sandwich, and is a common feature of many GH13 family enzymes. SusG displays an unusual elongated shape, ~120 Å in length, due to the insertion of a CBM (residues 216–335) that protrudes from the B domain. This CBM, hereafter referred to as CBM58, displays a β sandwich fold with immunoglobulin-like topology, composed of one five-stranded antiparallel β sheet opposing a four-stranded antiparallel β sheet. CBM58 makes no direct contact with the ABC domains, and is linked to the core amylase fold by two short linkers that span the 12 Å between the B domain and the CBM58. The linker sequences, SDETAA (residues 212–217) and DSQQI (residues 334–338), are not inherently flexible, and the B factors of these atoms are consistent with neighboring protein atoms. The linker strands do not directly interact with each other, the core

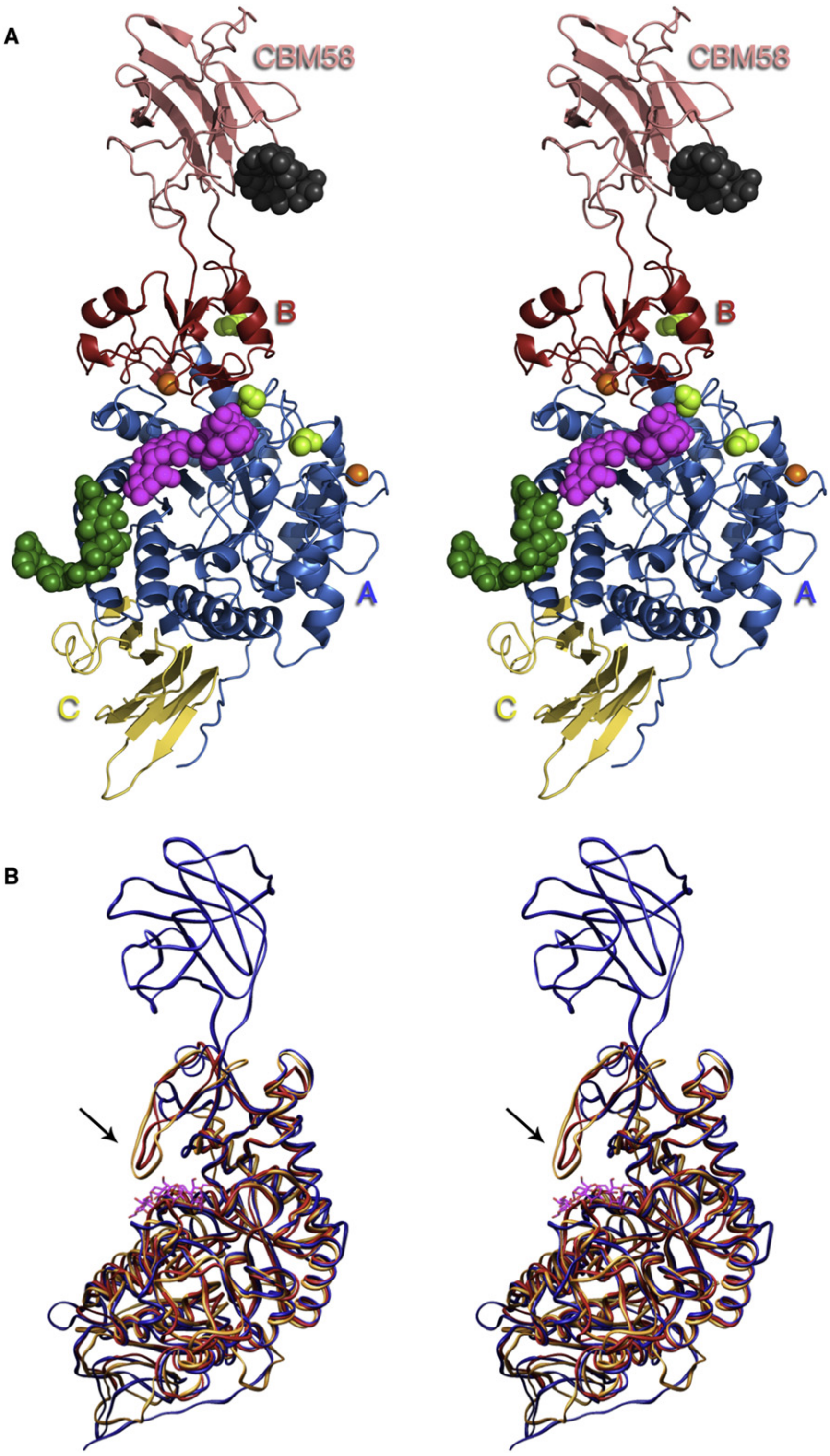

Figure 2. The Structure of SusG

(A) Ribbon diagram of SusG, colored by domain: the A domain (residues 43–152 and 364–607) is in blue, the B-domain (residues 153–215 and 336−363) is in red, the starch-binding domain (CBM58, residues 216–335) is in pink, and the C-domain (residues 608–692) is in yellow. The metal ions are displayed as orange spheres, and likely ethylene glycol molecules are in light green. The locations of bound maltoheptaose molecules are represented by mauve, green, and gray spheres to the active site, the secondary starch-binding site, and CBM58, respectively.

(B) Overlay of SusG (blue) with the *Halothermothrix orenii* α-amylase (PDB ID code 1WZA; yellow), and the *Thermotoga maritima* 4-α-glucanotransferase (PDB ID code 1LWJ). The arrow highlights a loop that occludes the active site of the α-amylase and glucanotransferase, missing in SusG, that may account for differences in substrate specificity between these homologs.

domains, or the CBM58, but have a few potential water-mediated hydrogen bonds with each other. When residues 365–692 of chains A and B are superimposed (root-mean-square deviation [rmsd] 0.3 Å) there is some displacement of the CBM58s, with a 3.3 Å deviation in the Cα atoms at the distal end. The location of the CBM58, both internal to the amino acid sequence and remote from the core domains, has not been observed in any other GH13 structure, although some bacterial amylopullulanases have predicted starch-binding CBM20 domains internal to the polypeptide sequence (Machovic et al., 2005).

The apo SusG structure has two metal ions, modeled as Ca^{2+} and Mg^{2+} based upon the distance and geometry of the coordinating atoms (Figure 2A). Both of these ions bind in locations typically occupied by Ca^{2+} in other amylase structures, and contribute to structural integrity (Abe et al., 2004; Hondoh et al., 2003; Robert et al., 2005; Roujeinikova et al., 2002; Sivakumar et al., 2006). The Ca^{2+} ion is located between the A and B domains, ~12 Å from the catalytic site, and is coordinated by two water molecules, the main-chain O of H392 and I393, and the side chains of D352 and N153, with an average coordinating distance of ~2.4 Å. The Mg^{2+} ion is coordinated by one water molecule, the main-chain O of Y79, and the side chains of D73, D75, D77, and D81, with average distances of 2.0–2.3 Å. These residues lie within a surface loop that connects β1 and α1 in the catalytic A domain, remote from the active site.

Excluding CBM58, SusG shares the most sequence and structural similarity with *Halothermothrix orenii* α-amylase A (Protein Data Bank [PDB] ID code 1WZA; rmsd of ~1.8 Å for 447 Cα atoms with a 34% sequence identity) (Sivakumar et al., 2006) and *Thermotoga maritima* 4-α-glucanotransferase (PDB ID code 1LWJ; rmsd of ~2.0 Å for 426 Cα atoms and a 28% sequence identity) (Roujeinikova et al., 2002), as determined by the Dali server

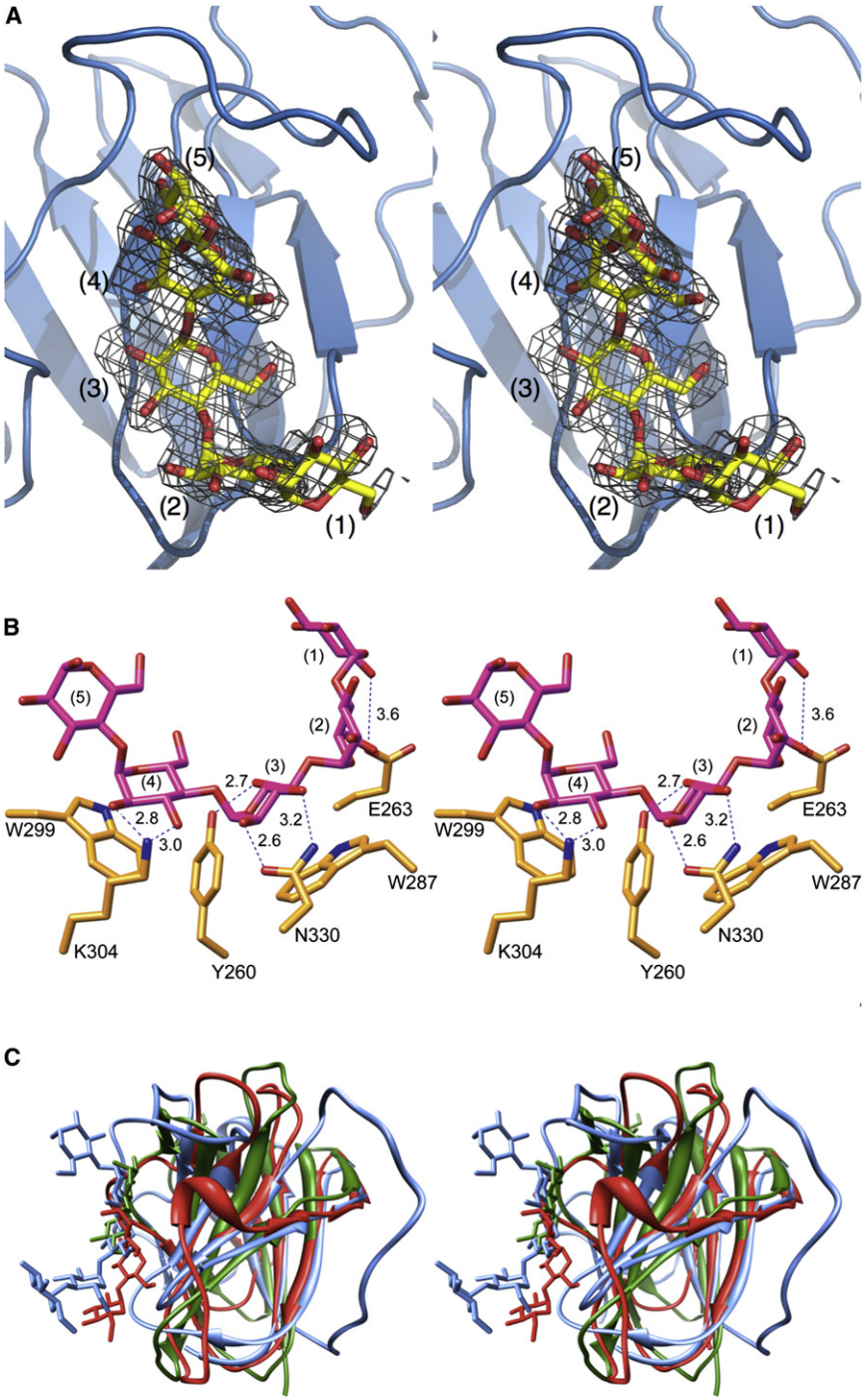

Figure 3. Maltooligosaccharide Bound to the Novel Carbohydrate-Binding Domain, CBM58, of SusG

(A) Electron density from an omit map at the carbohydrate binding to the CBM58 domain. The electron density is contoured at 3σ.

(B) Stereo view of maltopentaose bound at the starch-binding domain with the potential hydrogen bonds denoted by dashed lines. Glucose residues in the oligosaccharide are numbered from the nonreducing end. For clarity, the view in (A) is looking into the carbohydrate-binding region, whereas in (B) the view is from the side.

(C) Alignment of CBM58 (from SusD), CBM26 (from the amylase of *Bacillus halodurans*; PDB ID code 2C3H), and CBM41 (from the pullulanase of *Thermotoga maritima*, PulA; PDB ID code 2J73) in blue, red, and green, respectively. The bound oligosaccharides are shown as stick figures in colors corresponding to the ribbon diagrams.

(Holm and Sander, 1995). In both instances, the most divergent part of the structures is within the B domain, adjacent to the active site (Figure 2B). The B domain loop created by residues 163–175 in *H. orenii* amylase A and 123–136 in *T. maritima* glucanotransferase lines one side of the entrance to the active site, partially restricting substrate access. In contrast, residues 183–207 in the B domain of SusG form a helix-turn-3_{10} helix that points away from the catalytic site, creating a much wider cleft for substrates to enter, and may contribute to substrate specificity as discussed below.

Based upon amino acid sequence, the domain composed of residues 216–335 represents a new starch-binding CBM family, designated CBM58 by the Carbohydrate-Active EnZymes database (CAZy; http://www.cazy.org). The topology of the β strands is quite different compared to other CBM families, with one five-stranded β sheet composed of β5-β6-β1-β2-β10 opposing a four-stranded β sheet composed of β4-β3-β7-β8, and a 14 residue elongated loop that connects β2 and β3. Two additional parallel β strands, β9 and β11, form a small β sheet. The sheet β5-β6-β1-β2-β10 is flat whereas the opposite face features three protruding loops, residues 262–264 connecting β3 and β4, residues 286–288 connecting β6 and β7, and residues 295–301 connecting β7 and β8, that create an oligosaccharide-binding depression over the sheet. Within this binding pocket, Y260, W287, and W299 form the starch-binding site. This new CBM shares the most structural similarity with the starch-binding proteins CBM26 of *Bacillus halodurans* maltohexaose-forming amylase (PDB ID code 2C3H) and CBM41 of *T. maritima* pullulanase PulA (PDB ID code 2J73) (Boraston et al., 2006; Lammerts van Bueren and Boraston, 2007). Despite a different pattern of connectivity, the β strands of CBM58 are positioned similarly to CBM26 (PDB ID code 2C3H; rmsd of 1.39 Å for 45 Cα, 18% sequence identity) and CBM41 (PDB ID code 2J73; rmsd of 1.3 Å for 43 Cα, 15% sequence identity), with the starch-binding sites on the same face of the β sandwich (Figure 3C). These structural alignments reveal a conserved mode of starch binding, with a Trp-Trp/Tyr pair that creates a shallow pocket for binding helical α1,4-glucan.

SusG-D498N/Maltoheptaose Structure

Catalytic residues of SusG were initially identified from a sequence alignment with the *H. orenii* AmyA and SusG was inactivated with a D498N mutation, as confirmed using PNP-maltohexaose and amylose as substrates. This D498N mutant was crystallized in the presence of maltoheptaose and the structure was determined to a resolution of 2.3 Å (R_{cryst} = 18.6%, R_{free} = 21.7%). These crystals were isomorphous to the apo crystals despite the fact that 100 mM $LiSO_4$ and 0.5 mM $CaCl_2$ replaced the magnesium acetate. Under these conditions, Ca^{2+} replaced the Mg^{2+} ion observed in the apo structure, inferred by the increase in the average coordination distance to 2.3–2.6 Å. In this mutant form of SusG, maltooligosaccharide was observed at three sites: the catalytic cleft, CBM58, and a surface starch-binding site adjacent to the active site (Figure 2A). The structure of this D498N mutant can be superimposed onto the wild-type (WT) SusG with an rmsd of <0.45 Å, and no conformational changes were observed as a result of maltooligosaccharide binding. In the following sections, glucose residues are numbered from the nonreducing end of the maltoheptaose. At both the surface site and the CBM58, the ϕ (O5-C1-O4′-C4′), ψ (C1-O4′-C4′-C5′) angles of maltooligosaccharide approximate those typically found in double-helical amylose (ϕ = 91.8°, ψ = −153.2°; ϕ = 85.7°, ψ = −145.3°; ϕ = 91.8°, ψ = −151.3°) (Imberty et al., 1988).

The CBM58-Binding Site

Five glucose residues of maltoheptaose are well ordered at the CBM58, cradled by the loops that form the binding pocket on one face of the β sandwich (Figure 3). Glc3 and Glc4 of the bound maltoheptaose stack against two tryptophans in the binding pocket: W299 and W287. Adding to the hydrophobic character of this area, L290 lies between these two tryptophans. The side chain of E263 is located 3.6 and 2.8 Å from the O-6 of Glc1 and Glc2, respectively. The O-2 and O-3 of Glc3 are hydrogen bonded with the side-chain $O^{\delta1}$ and $N^{\delta2}$ of N330, at 2.6 and 3.2 Å, respectively. Similarly, the N^{ζ} of K304 is positioned 3.0 Å from the O-3 and 2.8 Å from the O-2 of Glc4. The O-6 of Glc3 points toward the phenolic oxygen of Y260 (2.7 Å). At the reducing end of the maltoheptaose, Glc5 does not form any stacking or hydrogen-bonding interactions with the protein. The pattern of starch binding at the CBM58, characterized by an arc of hydrophobic residues with additional hydrogen bonding to the 2′ and 3′ hydroxyl groups of adjacent glucose residues, is a generally conserved feature of many starch-binding CBMs (Boraston et al., 2006). In addition, this binding pattern is observed in SusD (Koropatkin et al., 2008) and in barley and pancreatic α-amylases that bind raw starch on the surface of the catalytic domain (Qian et al., 1995; Robert et al., 2005).

The Active Site

Maltoheptaose assumes a curved shape similar to β-cyclodextrin in the active site with ϕ,ψ angles of 46° and −150° between the −1 and +1 subsite residues Glc4 and Glc5. In general, the glucose residues in the active site are positioned by hydrogen-bonding interactions through their O-2 and O-3 atoms with few hydrophobic stacking interactions (Figure 4). The most extensive protein-glycan interactions are observed with Glc3, Glc4, and Glc5, corresponding to subsites −2, −1, and 1. The architecture of the active site is similar to many GH13 enzymes, and this homology infers that D388, E431, and D498 have key roles in a general acid-base double-displacement mechanism (Qian et al., 2001). At the nonreducing end, Glc1 makes potential hydrogen bonds between O-2 and the side chains of D545 (2.7 Å) and K541 (3.3 Å), and O-3 with K541 (3.4 Å). This residue is disordered in the other copy of SusG in the crystallographic asymmetric unit. The O-3 of Glc2 hydrogen bonds with the backbone N and the side chain of D545. Glc3 stacks against the hydrophobic face of H112, with O-2 forming likely hydrogen bonds with side-chain N atoms of R549, and its O-3 atom forming likely hydrogen bonds with the side chains of D545 and R549. The phenyl side chains of F345 and F350 lend additional hydrophobic character along the O-5, C-6 face of Glc3 and Glc4. Glc4 stacks against the phenolic side chain of Y114, and is positioned by an extensive network of hydrogen bonds. O-2 of Glc4 is proximal to the side-chain atoms of N498, R386, and E431, whereas the O-3 atom interacts with the side chains H497 and N498. O-6 of Glc4 interacts with the $N^{\varepsilon2}$ of H154, and O-5 with the side chain of D388. D388 is positioned 3 Å from the C1 of Glc4, supporting its role as the nucleophilic base that forms the β-glucosyl enzyme intermediate. E431 interacts with O-4 of Glc5, likely supplying a proton to the leaving α-glucan chain and then activating a water molecule for hydrolysis of the enzyme intermediate. D498, mutated to Asn in this structure, plays a critical role in catalysis, perhaps by maintaining the pKa of the general acid and/or by stabilizing the positive charge of the transition state, believed to have carbonium ion character (McCarter and Withers, 1994, 1996; Qian et al., 2001; Strokopytov et al., 1995; Uitdehaag et al., 1999). The portion of maltoheptaose representing the leaving α-glucan chain, Glc5–Glc7, has fewer contacts with the protein. The O-2 and O-3 atoms of Glc5 form hydrogen bonds with the side chains of H392 and E431. The O-2 of Glc6 interacts with the backbone O of L433, whereas the O-3 interacts with the N^{ζ} of K391. Glc7 at the reducing end of maltoheptaose is ~4 Å from the aromatic face of Y456 but is not optimally positioned for aromatic stacking interactions.

The Surface Starch-Binding Site

Directly adjacent to the reducing end of the active site maltoheptaose is an additional surface starch-binding site (Figure 5). The O1 atom of Glc7 from the oligosaccharide bound to the active site is 4.5 Å from the O-2 atom of Glc5 of the surface-site-bound ligand, resulting in close proximity of the reducing ends of each maltooligosaccharide. This orientation makes it unlikely that a continuous segment of α-glucan spans both the active site and the surface site during catalysis. In both subunits of SusG, six of the seven glucose residues of maltoheptaose could be modeled at the surface site. As described for CBM58, an arc of aromatic amino acids creates a hydrophobic surface for binding α-glucan, with additional hydrogen-bonding interactions with the O-2 and O-3 atoms provided by polar side chains. Glc4 and Glc3 stack against W460 and Y469, while T466 and several water molecules occupy the small space between W460 and Y469. This is in contrast to CBM58, in which L290 and Y260 lend additional hydrophobic character to the binding cleft between the two tryptophans. The O-2 and O-3 atoms of Glc3 interact with the side chain of D473, whereas O-2 and O-3 of

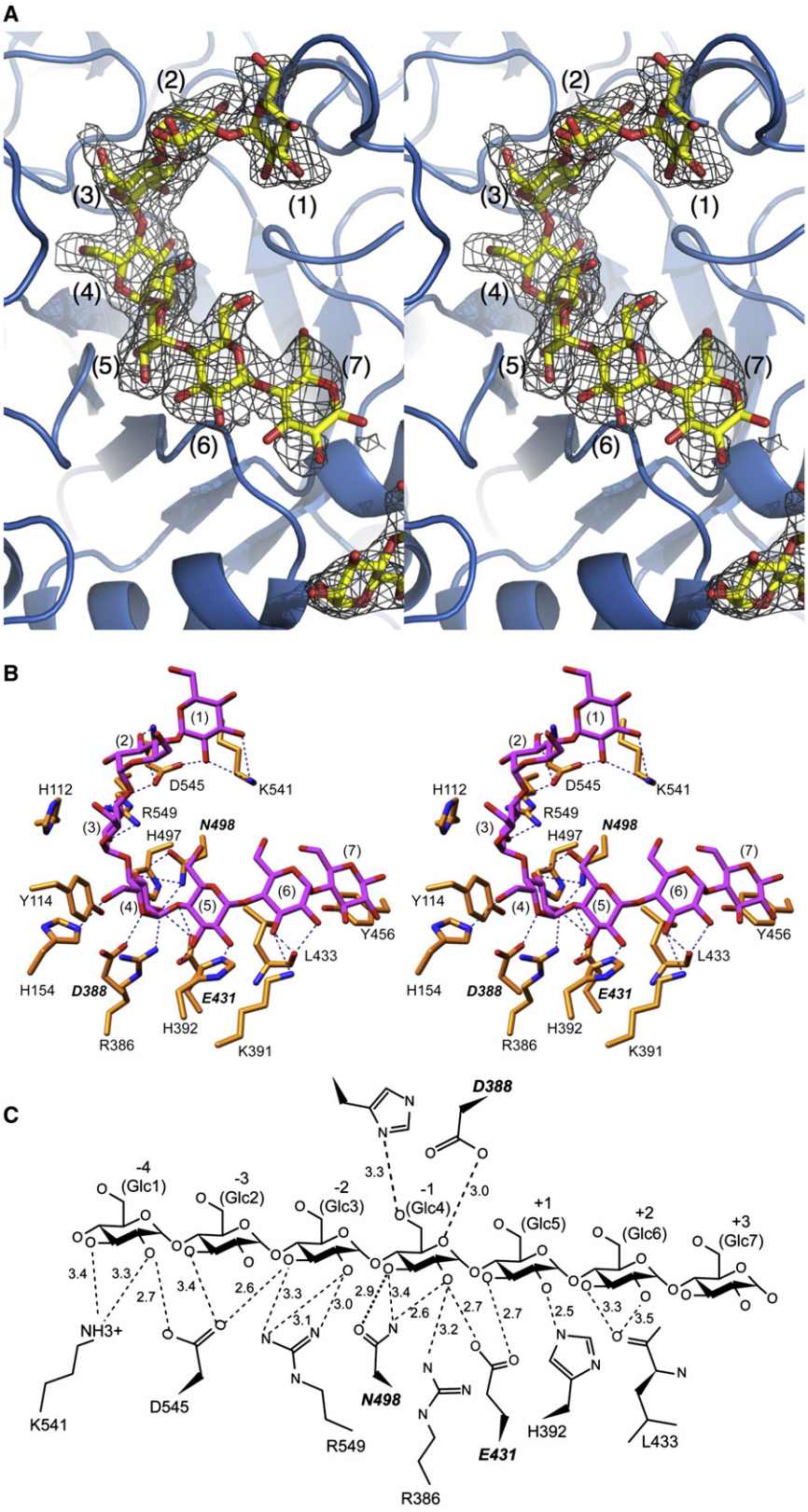

Figure 4. Maltooligosaccharide Bound to the Active Site of SusG

(A) Electron density from an omit map at the SusG active site of the SusG-D498N mutant cocrystallized with maltoheptaose. The electron density is contoured at 3σ and the stick model of the bound oligosaccharide is colored according to atom type.

(B) Stereo view of oligosaccharide bound at the starch-binding domain, with the potential hydrogen bonds denoted by dashed lines.

(C) Schematic of enzyme-substrate interactions in the active site. Glucose residues in the oligosaccharide are numbered from the nonreducing end, and enzyme subsites −4 through +3 are labeled for clarity.

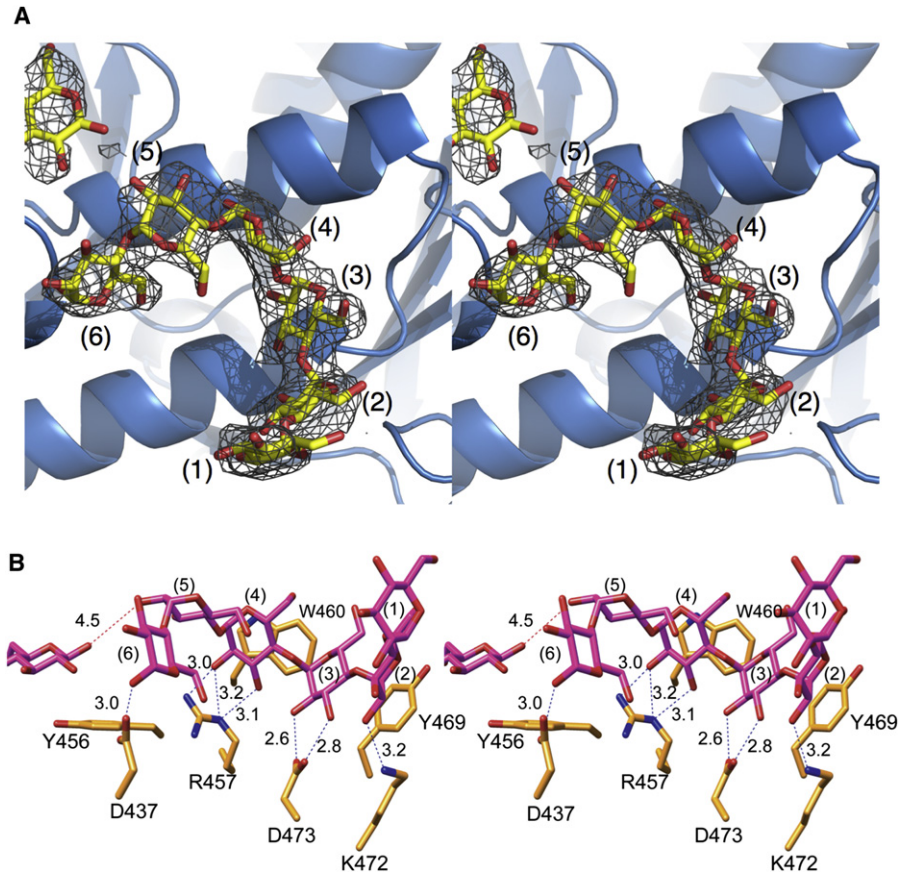

Figure 5. Maltoheptaose Bound to the Secondary Oligosaccharide-Binding Site Adjacent to the Active Site

(A) Electron density from an omit map of the SusG-maltoheptaose complex at the secondary carbohydrate-binding site immediately adjacent to the active site. Note that the glucose ring in the upper left corner is from the oligosaccharide bound to the active site. The electron density is contoured at 3σ.

(B) Stereoview of oligosaccharide bound at this secondary binding site, with the potential hydrogen bonds denoted by dashed lines. Glucose residues in the oligosaccharide are numbered from the nonreducing end. For clarity, the view in (A) is looking into the carbohydrate-binding region, whereas in (B) the view is from the side.

Glc4 bind to the guanidinium group of R457. The O-2 atom of Glc2 hydrogen bonds with the N^{ζ} of K472, and the reducing end O1 of Glc6 binds to the side chain of D437. The close proximity of this additional binding site and the catalytic site may restrict the length of α-glucans that can bind to either site during the catalytic cycle. The architectural similarity of the active site of SusG compared with other α-amylases, including the orientation of maltoheptaose, suggests that the α-glucan chain is cleaved at the reducing end, making the surface site proximal to the leaving α-glucan chain.

Acarbose Binding to SusG

Acarbose is used to treat type 2 diabetes and, in some countries, prediabetes, by blocking starch degradation via inhibition of α-amylase (Sharma and Garber, 2009). To further detail the active site of SusG, the structure of selenomethionine-substituted WT SusG complexed with the amylase inhibitor, acarbose, was determined to a resolution of 2.5 Å (R_{cryst} = 20.2%, R_{free} = 24.9%). Acarbose is a pseudotetrasaccharide composed of the disaccharide acarviosine with an α-1,4 linkage to maltose (Figure 6A). The nonhydrolyzable acarviosine moiety contains an unsaturated cyclitol ring N-linked to 4,6-dideoxy-4-amino-D-glucose, and is thought to act as a transition-state analog for the amylase reaction. Whereas intact acarbose was observed at both the CBM- and surface-binding sites, a β-glucosyl enzyme intermediate was captured in the active site (Figure 6) with D388 covalently linked to the pseudotrisaccharide acarviosine-glucose, with maltose occupying subsites +1/+2. To our knowledge, this is the first GH13 enzyme structure bound to a covalent intermediate that is not derived by the use of a fluorinated substrate or a site-directed mutant. However, this has been previously observed with acarbose in the case of *Thermus thermophilus* amylomaltase, a related enzyme belonging to glycoside hydrolase family 77 (Barends et al., 2007).

It is not clear how the intermediate was trapped in the active site. Acarbose was added to the protein prior to crystallization and occasionally stored for more than 2 weeks prior to crystallization, which may have allowed for significant degradation of the pseudotetrasaccharide. Crystals were grown over a period of ~2 weeks at room temperature, prepared for data collection by transferring to increasing concentrations of cryoprotectant containing fresh acarbose, and flash-frozen within 1 min of the

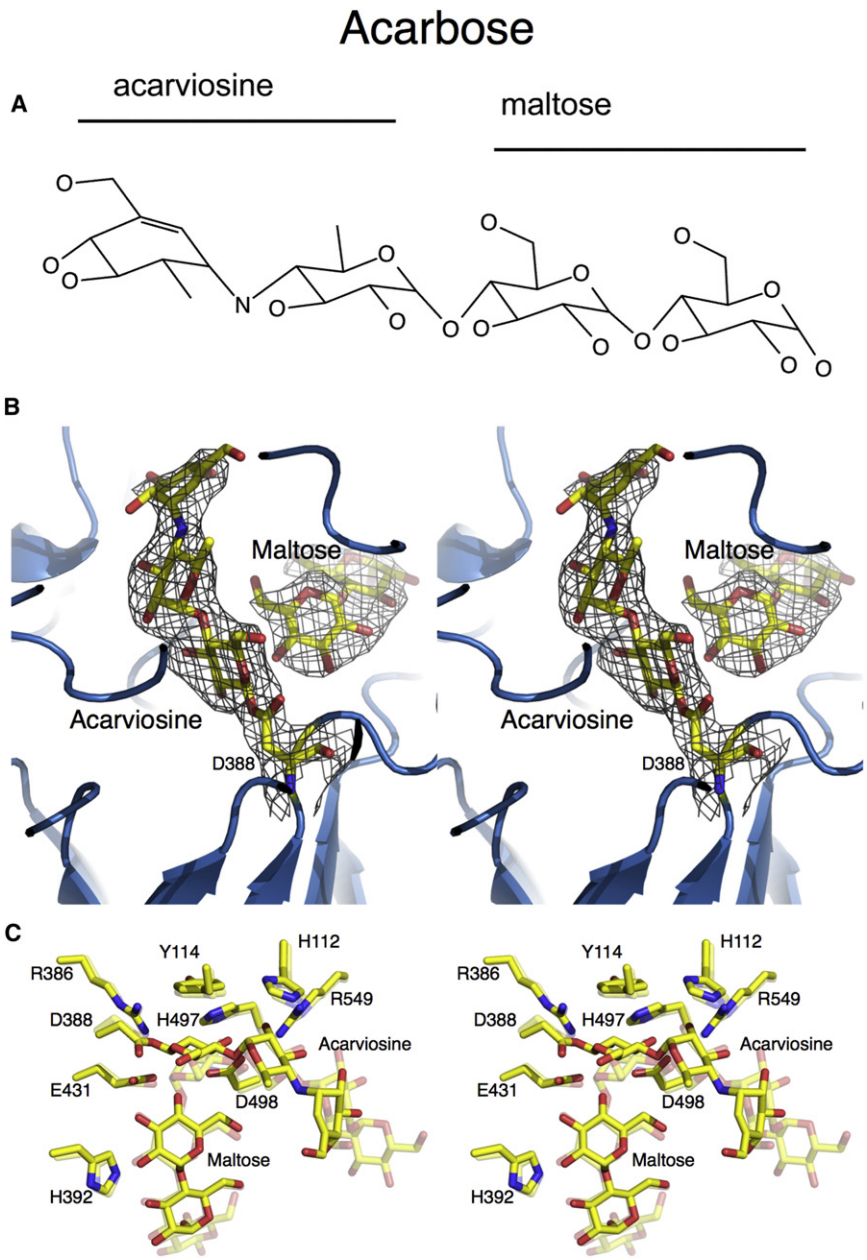

Figure 6. Enzyme-Linked Intermediate of the SusG Reaction with Acarbose

(A) Chemical structure of acarbose, a pseudotetrasaccharide inhibitor of α-amylase and α-glucosidase.

(B) Electron density of acarviosine-glucose covalently linked to the active site, D388, and maltose from the corresponding omit map contoured at 3σ.

(C) Overlay of the active site from the WT-SusG structure with acarbose (solid sticks) with that of the SusG-D498N structure complexed with maltoheptaose (transparent model).

final transfer. Slow release of the pseudotrisaccharide, combined with the displacement of water by cryoprotectant, may have trapped the covalent intermediate. The presence of maltose, rather than glucose, in subsites +1/+2 is surprising. Neither the acarbose stock solution nor the SusG and acarbose reaction have appreciable quantities of maltose as assessed by thin-layer chromatography (TLC). It is possible that SusG cleaves maltose, rather than glucose, from the reducing end of acarbose as a minor reaction, and that maltose observed in this structure is simply the result of such a previous catalytic event. Many GH13 enzymes perform transglycosylation and subsequent cleavage of acarbose that can result in the generation and breakdown of a pentasaccharide that would yield acarviosine-glucose and maltose (e.g., human pancreatic α-amylase; Li et al., 2005). However, if SusG performs transglycosylation it must be an infrequent event, because the products of such a side reaction were not detected by TLC (see Figure 9).

Despite slight perturbations, an identical pattern of hydrogen bonding is observed in the active sites of the WT SusG complexed with acarbose and the SusG-D498N mutant bound to maltoheptaose. The nucleophilic base D388 shifts slightly upon attack on acarbose C-1, resulting in the inversion of stereochemistry at the anomeric carbon and the β-glucosyl intermediate. The side chain of E431 is positioned ~3.1 Å from the nonreducing O-4 end of the leaving maltose group, whereas this distance is 3.4 Å in the maltoheptaose-bound structure. Similarly, the side chain of D498 and the anomeric C-1 of the pseudotrisaccharide are both ~3.1 Å from the nonreducing O-4 of maltose. Both the bound glucosyl-enzyme intermediate and maltose molecules are well ordered, suggesting that nearly all of the active sites are occupied by these ligands in this configuration.

At the CBM, acarbose binds in a similar manner as the oligosaccharide in the SusG-D498N structure and occupies the same subsites as Glc2–Glc5 of maltopentaose (Figure 7). By virtue of the unsaturated bond in acarbose, the O-6 atom of the cyclitol moiety is slightly shifted, and the side chain of E263 is directed away from the binding site. At the surface site, acarbose is bound in the same subsites occupied by Glc2–Glc5 of maltohexaose in the SusG-D498N structure (Figure 8). Glucose at the reducing end of acarbose is shifted less than 1 Å in the direction of the active site compared to Glc5 of maltopentaose.

Enzymatic Activity

A qualitative analysis of the enzymatic and starch-binding properties of SusG was performed to better understand the possible roles of these various carbohydrate-binding domains. Polysaccharide affinity gel electrophoresis indicated that SusG is flexible in its carbohydrate selectivity because it binds to and degrades

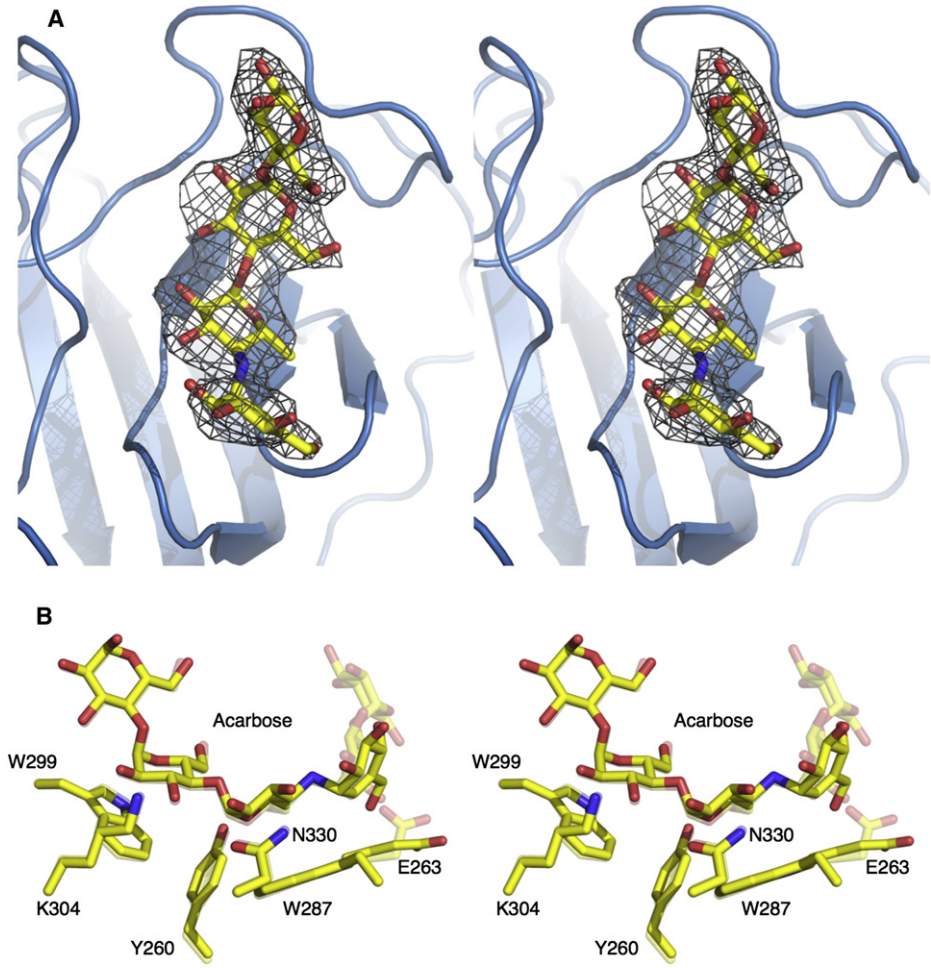

Figure 7. Acarbose Bound to the CBM58 Domain

(A) The electron density of acarbose bound to the CBM58 domain. For this figure, the omit map was contoured at 3σ.

(B) Comparison of acarbose binding (solid sticks) with the SusG-D498N structure cocrystallized with maltoheptaose (transparent model).

pullulan, amylopectin, and amylose. To assess the relative substrate preference of SusG, the amount of reducing sugar liberated over time was assessed using soluble starch (potato), amylopectin (maize), pullulan, α-cyclodextrin, and β-cyclodextrin as substrates (Table 1). From these data, it is clear that SusG prefers soluble starch (defined as 100% activity) over pure amylopectin (51.5%), which contains α-1,6 branch points about every 25 glucose residues. In addition, it is able to process pullulan (46.8%), a property not universally possessed by α-amylase enzymes. Interestingly, SusG is also able to degrade both β-cyclodextrin and α-cyclodextrin, albeit at rates approximately 8% and 2% of the rate of soluble starch. The ability to degrade the cyclodextrins was surprising, as these tend to act as nonhydrolyzable inhibitors for most amylose/amylopectin-preferring amylases.

The reaction products of WT SusG on maltooligosaccharides, amylose, amylopectin, pullulan, dextran, cyclodextrins, and acarbose were analyzed by TLC (Figure S1). All of the substrates tested were significantly degraded with the exception of maltose, acarbose, and dextran. Typical byproducts were glucose and maltose for maltose oligosaccharides (G3–G7), α- and β-cyclodextrin, amylose, and amylopectin. SusG degraded pullulan exclusively to panose (Glc-α1,6-Glc-α1,4-Glc), with strict specificity for the α1,4-glycosidic bonds. This is consistent with the finding that SusG has no detectable activity toward dextran, an α-1,6-linked polymer of glucose (Figure 9).

In order to understand how CBM58 and the surface-binding site affect the starch-binding properties of the enzyme, several mutant versions of SusG were created. A mutant of WT SusG lacking CBM58, named ΔCBM58, was created by deleting residues 210–339 and inserting the five residue loop GSPTG, similar to that observed in the *H. orenii* amylase A, a close structural homolog of SusG that does not have CBM58. A second mutant of WT SusG, ΔSURF, was created by mutating the surface site (W460A/Y469A/D473V) to prevent starch binding on the α-amylase domain. The ΔCBM58, ΔSURF, and WT-SusG enzymes were assayed for activity using *p*-nitrophenyl-maltopentaose (PNP-G5) and were found to have nearly identical catalytic turnover rates. This suggests that neither mutation significantly perturbs the intrinsic catalytic rate of the α-amylase (Table 2).

The enzymes were then tested for their ability to degrade soluble starch, amylopectin, and pullulan. Reactions contained

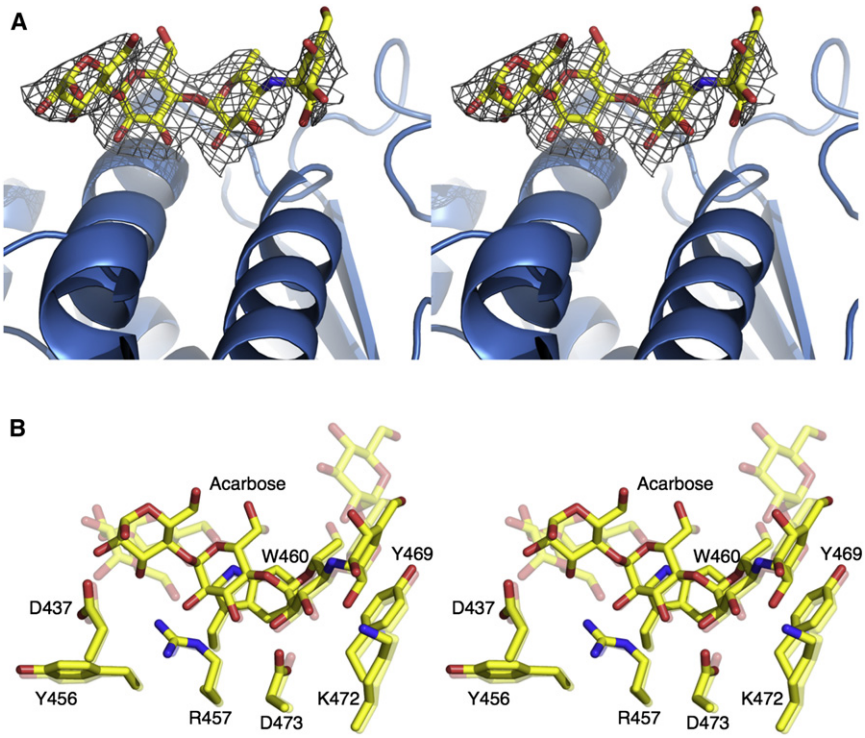

Figure 8. Acarbose Bound to the Secondary Oligosaccharide-Binding Site Adjacent to the Active Site

(A) Shown here is the omit electron density, contoured at 3σ, of acarbose bound to the surface carbohydrate-binding site immediately adjacent to the active site.

(B) Comparison of acarbose binding (solid sticks) with that of the SusG-D498N structure cocrystallized with maltoheptaose (transparent model).

equal amounts (50 nM) of enzyme, and the amount of reducing sugar liberated was quantified in a DNSA-based assay. For each substrate, the activity of WT SusG was defined as 100%, and the ΔCBM58 and ΔSURF mutant enzymes were compared to the wild-type. The enzymes had relatively similar levels of activity on pullulan, with the ΔCBM58 mutant being ~30% more active than WT, and the ΔSURF mutant ~20% less active than WT. From the starch preference assay, pullulan is not the preferred substrate for SusG, and the absence of the starch-binding sites do not greatly affect enzymatic activity. The solubility and inherent flexibility of pullulan (Leathers, 2003) may allow it access to the active site without the help of the starch-binding sites. More profound effects were observed for both soluble starch and amylopectin. The ΔCBM58 mutant had significantly higher activity, with 162% and 292% that of WT SusG on soluble starch and amylopectin, respectively. If the purpose of CBM58 is solely to provide the enzyme accessibility to insoluble substrates, then it is not surprising that the removal of the CBM does not decrease catalytic efficacy. However, it was unexpected to see such an increase in activity on soluble starch and amylopectin when CBM58 was removed. Isothermal titration calorimetry (ITC) with isolated CBM58 indicates a K_d of ~10 μM for maltoheptaose and β-cyclodextrin, and ~72 μM for 6^3-α-D-glucosyl-maltotriose-maltotriose, an oligosaccharide of pullulan. The ΔSURF mutant enzyme displays only a slight decrease in activity on soluble starch and amylopectin, suggesting that, like CBM58, the surface site is dispensable for soluble starch degradation. However, unlike the ΔCBM mutant, removal of the surface site does not increase activity.

Thus far, none of the biochemical analyses have identified an essential role for the starch-binding domain on soluble forms of starch. Therefore, the ability of these mutants to bind and hydrolyze insoluble cornstarch were measured (Table 2). Both the ΔCBM58 and ΔSURF mutant enzymes had lower efficacy with cornstarch as a substrate compared to WT SusG, suggesting that both sites play a role in processing of an insoluble α-glucan. SusG-D498N, ΔCBM58-D498N, and CBM58 were tested for their ability to bind insoluble cornstarch (Figure 10). The CBM58 domain and the SusG-D498N mutant displayed similar relative affinities for insoluble cornstarch. The ΔCBM58-D498N mutant is very deficient in insoluble starch binding, and a precise K_d could not be determined. This suggests that SusG binding to substrate is very much dependent upon the CBM58 domain. The fact that CBM58 alone had the highest affinity for cornstarch may be due to more facile interactions between such a small protein and a large starch molecule compared with the full-length enzyme.

Table 1. SusG Activity on Various Oligosaccharide Substrates

	Activity of WT SusG (%)
Soluble starch	100 ± 4.2
Amylopectin	52 ± 8.0
Pullulan	47 ± 2.6
β-cyclodextrin	7.8 ± 2.0
α-cyclodextrin	1.8 ± 2.6

DISCUSSION

The α-amylase SusG has an atypical bilobed structure, with the core amylase A, B, and C domains at one end and a starch-specific CBM at the other end that is formed by a large insertion in the α-amylase domain. This starch-binding domain represents a new class of CBMs, now designated CBM58 by the CAZy database (Cantarel et al., 2009). Like other starch-specific CBMs, the α-glucan-binding platform is composed of two aromatic side chains that create an arc complementing the natural helical twist of amylose, with additional hydrogen bonding to the O-2 and O-3 atoms of adjacent glucose residues (Boraston et al., 2006; Machovic and Janecek, 2006). The position of CBM58 in SusG is highly irregular for two reasons. First, the domain is not at the N or C terminus of the protein like most CBMs found in the

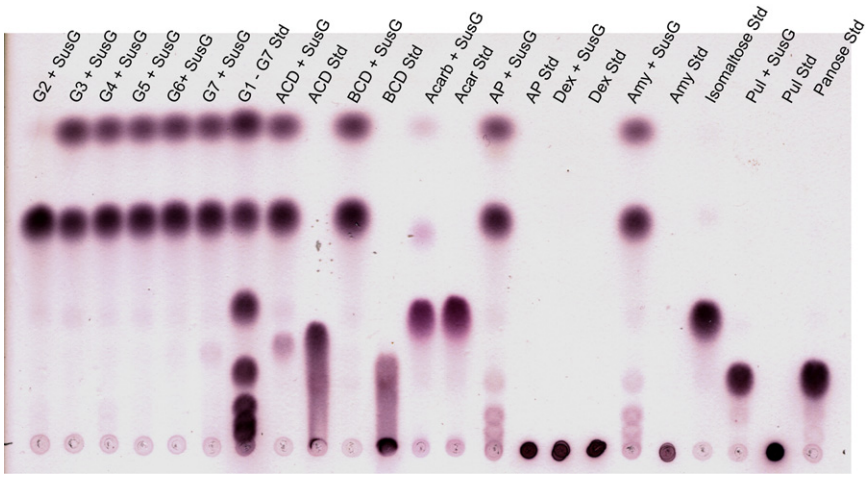

Figure 9. Thin-Layer Chromatography of Starch Hydrolysis Byproducts and Standards

All reactions contained 5 mg/ml of maltooligosaccharide or starch polysaccharide, 22 μg/ml SusG, 15–20 mM HEPES (pH 7.0), and 75–100 mM NaCl. Reaction products were sampled after a 2 hr incubation at 37°C. G1–G7, maltooligosaccharides; ACD, α-cyclodextrin; BCD, β-cyclodextrin; Acarb, acarbose; AP, amylopectin; Dex, dextran; Amy, amylose; Pul, pullulan.

α-amylase family (Machovic and Janecek, 2006; Machovic et al., 2005). Second, CBM58 of SusG does not make any hydrogen-bond contacts to any part of the core catalytic domain. Whereas some GH13 enzymes have an N-terminal CBM that is somewhat extended from the core of the structure, these are typically involved in dimerization and contribute to the shape and specificity of the active site of the neighboring molecule (Fritzsche et al., 2003; Hondoh et al., 2003; Kamitori et al., 1999; Lee et al., 2002). The extension of CBM58 away from the rest of the SusG structure is reminiscent of the multidomain endoglucanases (i.e., cellulases) Cel9G of *Clostridium cellulolyticum* (Mandelman et al., 2003) and CelE4 of *Thermomonospora fusca* (Sakon et al., 1997), which display a C-terminal cellulose-binding CBM3 connected to the catalytic domain via a 15–18 residue linker. In these cases, the linker maintains an extended conformation via an extensive hydrogen-bond network between the linker and the two domains. The CBM3 found in the Cel9G and CelE4 endoglucanases is believed to disrupt the hydrogen-bonding network of crystalline cellulose and guide the cellulose chains to the catalytic site, because the flat face of CBM3 is in-line with the active site on the catalytic domain. With SusG, the starch-binding site on CBM58 is 45 Å away from, and at 90° to, the active site, making it difficult to envision a similar mechanism.

An overlay of the catalytic domains of the two noncrystallographically related copies of SusG in the maltoheptaose-bound D498N structure showed a displacement of the CBM by only a few Ångstroms. The small amount of movement observed and the relatively short length (~12 Å) of the strands linking CBM58 to the B domain make it difficult to envision it passing maltooligosaccharide to the catalytic domain. Perhaps this flexibility is important for interacting with, or channeling substrates to, other members of the Sus complex.

Generally speaking, CBMs of glycoside hydrolases promote adsorption to an insoluble substrate, disrupt the structure of the polysaccharide, and improve catalytic efficacy by concentrating the enzyme on the glycan (Boraston et al., 2004). In members of the GH13 family, mutation of a CBM can dramatically decrease the ability of the enzyme to utilize raw or granular starch (Penninga et al., 1996; Sumitani et al., 2000; Tan et al., 2008), and the addition of a CBM can enhance the ability to digest raw starch (Juge et al., 2006; Latorre-Garcia et al., 2005). CBM58 clearly enhances the ability of SusG to bind insoluble cornstarch, because the ΔCBM-D498N mutant has at least 10-fold weaker binding compared to SusG-D498N (Figure 10). Residual binding of ΔCBM-D498N is likely due to binding at the surface site on the catalytic domain. For optimal degradation of insoluble cornstarch, both CBM58 and the surface site are required because the ΔCBM and ΔSURF mutants display similarly reduced levels of activity compared to WT SusG. In contrast, CBM58 seems to actually hinder degradation of soluble starch and amylopectin (Table 2). This disparity between soluble and insoluble starch digestion was surprising. Perhaps the function of CBM58 in SusG is not simply to concentrate the enzyme on the substrate, a role typically assigned to CBMs. CBM58 may have an essential starch-binding role that affects the apparent rate of catalysis yet is not directly related to enzyme activity. In vivo, CBM58 may help sequester α-glucan substrates or products to the cell surface for passage to other members of the Sus complex or into the SusC porin. In vitro, tight binding of CBM58 to starch may aid insoluble starch degradation by enhancing the local concentration of the substrate but inhibit the processivity of the reaction on a soluble substrate that can access the active site without the aid of a CBM.

Many amylases, such as barley α-amylase (Kadziola et al., 1998; Robert et al., 2005; Sogaard et al., 1993), yeast glucoamylase (Sevcik et al., 2006), salivary amylase (Ragunath et al., 2008), and pancreatic amylase (Payan and Qian, 2003; Qian et al., 1995), do not have a starch-binding CBM, but instead have one or more secondary starch-binding sites on the catalytic domain for raw starch utilization. In addition to CBM58, SusG also has a secondary starch-binding site (the "surface site") on

Table 2. Role of Carbohydrate-Binding Regions on Oligosaccharide Processing

	v_i (min^{-1}), 2 mM PNP-G5	v_i (min^{-1}), 0.2 mM PNP-G5	Pullulan (%)	Soluble Starch (%)	Amylopectin (%)	Insoluble Cornstarch (%)
WT	100 ± 1.3	100 ± 1.6	100 ± 8.1	100 ± 12.7	100 ± 8.1	100 ± 5.7
ΔCBM	112 ± 0.6	103 ± 1.4	137 ± 7.1	162 ± 5.9	292 ± 9.6	29 ± 1.4
ΔSURF	102 ± 6.2	90 ± 2.4	79 ± 9.8	82 ± 3.8	91 ± 3.7	44 ± 2.8

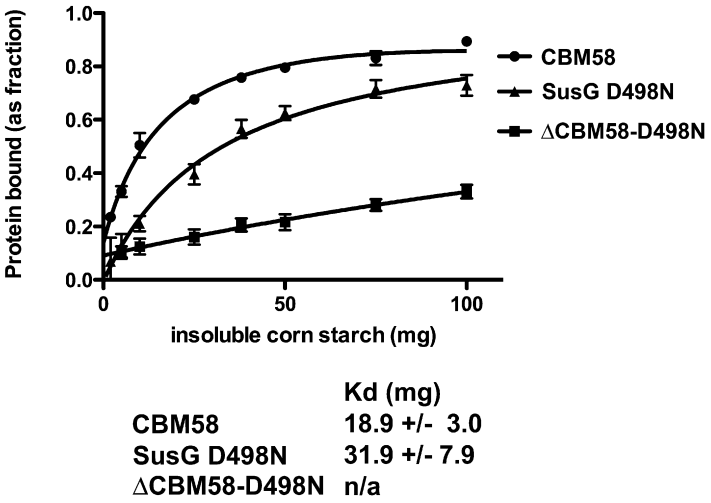

Figure 10. Protein Binding to Insoluble Cornstarch

Fraction of protein bound versus mg of cornstarch was plotted and fit using the one-site, total binding model (*B*max = 1) in the program Prism. A K_d was not calculated for the ΔCBM58-D498N mutant because the degree of saturation was too low to accurately extrapolate.

the catalytic domain directly adjacent to the active site (Figure 2). However, in SusG, the surface-binding site is only ~4.5 Å to the active site, whereas these sites in other amylases are typically separated by distances of 15 Å or more. Further, in SusG, the reducing ends of the bound oligosaccharides are pointed toward each other, making it unlikely that a single α-glucan chain spans both sites. Although the presence of the surface site in SusG aids the hydrolysis of insoluble starch (Table 2), the orientation of the maltooligosaccharide at the surface site may hint at a distinct role for this site. At the surface site, the hydrophobic platform created by W460 and Y469 is more or less perpendicular with the protein surface, and maltoheptaose binds along the length of the protein with many additional interactions between the reducing end of the sugar and the protein. Such an orientation precludes a longer amylose helix from binding, as the pitch of the helix would be directed into the protein instead of along the surface. This is in contrast to barley α-amylase, in which two tryptophan residues at the raw starch-binding site lie parallel to the surface of the protein, allowing larger amylose helices to bind. These differences suggest that perhaps SusG uses this site to sequester reaction products for subsequent import via the other Sus proteins.

In summary, it is clear that SusG has complex interactions with large polysaccharides; however, there is an apparent disconnect between oligosaccharide binding and in vitro catalytic turnover that may be pointing to the in vivo functions of these sites. Without the CBM58 domain, SusG binds very weakly to insoluble starch and, in fact, CBM58 binds better alone to starch than when linked to the rest of the enzyme. In contrast, the removal of CBM58 improves catalytic efficiency using soluble oligosaccharides (up to 3-fold in the case of amylopectin) with only 55% loss in activity against insoluble starch. Similarly, the surface starch-binding domain on the amylase does not seem to play a large role in digestion of soluble carbohydrates, but seems to be more important in processing insoluble starch. However, neither additional carbohydrate-binding region appears to be essential for processing these substrates. It has been previously demonstrated that a *ΔsusABCDEFG* mutant complemented with SusG cannot grow on amylose or amylopectin (Shipman et al., 1999), yet can grow on maltose and maltotriose. If SusG acts mainly to degrade α-glucan polymers to maltooligosaccharides as small as maltose, then it is not clear why this complementation mutant cannot utilize the larger substrates and import the products via other maltotriose/maltose uptake systems. It seems probable that the binding modules on SusG have a far greater role in starch sequestering and import within the Sus complex than starch degradation alone. Future experiments examining the binding and catalytic properties of the Sus proteins in vivo are needed to build a working model as to how the Sus complex functions as a whole.

EXPERIMENTAL PROCEDURES

Heterologous Protein Expression

The *susG* gene (residues 24–692) was amplified by PCR from genomic DNA prepared from *B. thetaiotaomicron* ATCC 29148 (also known as VPI-5482). The amplicon was cloned into pET28rTEV, where the thrombin cleavage site of pET-28a (Novagen) has been modified to a tobacco etch virus (TEV) protease cleavage site. From this wild-type (WT) *susG*-pET28rTEV construct, several mutant versions of *susG* were created. The construct *ΔCBM*-pET28-rTEV was created by PCR amplification of the DNA encoding residues 24–209 and 336–692. A BamHI restriction site was inserted within a linker region encoding the amino acid sequence G-S-P-T-G to bridge both parts of the gene. The two pieces of the *susG* gene were cut with BamHI and ligated together, followed by PCR amplification of the ligation product to obtain a continuous *susG* gene without the CBM (residues 210–335). The *ΔCBM* mutant gene was cloned into pET28rTEV for expression. Additional site-directed mutants of SusG included the catalytically inactive mutants *SusG*-D498N and *ΔCBM*-D498N, as well as *ΔSURF* (W460A/Y469A/D473V), in which the surface starch-binding site was removed. These mutants were constructed in the corresponding *susG*-pET28rTEV or *ΔCBM*-pET28rTEV expression plasmid using the QuikChange II Multi Site-Directed Mutagenesis kit (Stratagene) according to the manufacturer's instructions. The CBM of SusG, residues 215–337, was cloned and expressed independently as well, in a similar manner as described for the SusG enzymes. All primers utilized for cloning and mutagenesis are listed in Table S1.

Native SusG, mutants of SusG, and CBM58 alone were expressed similarly with an rTEV-cleavable N-terminal 6-His tag. The pET28rTEV plasmids were transformed into Rosetta (DE3) pLysS (Novagen) for protein expression. Cells were grown in TB medium at 37°C with shaking (225 rpm) until they reached an OD of ~0.4, at which time the temperature was adjusted to 22°C. At an OD of ~0.8, cells were treated with 0.2 mM ITPG to induce protein expression and allowed to grow 16 hr at 22°C. Cells were subsequently harvested by centrifugation, frozen in liquid nitrogen, and stored at −80°C. Selenomethionine-substituted protein was produced via the methionine inhibitory pathway (Van Duyne et al., 1993) as previously described (Koropatkin et al., 2007).

Purification of Native and Selenomethionine-Substituted SusG

All SusG proteins were purified using a 5 ml Hi-Trap metal-affinity cartridge (GE Healthcare) according to the manufacturer's instructions. The cell lysate was applied to the column in His buffer (25 mM NaH_2PO_4, 500 mM NaCl, 20 mM imidazole [pH 7.4]) and SusG was eluted with an imidazole (20–300 mM) gradient. The His tag was removed by incubation with rTEV (1:100 molar ratio relative to protein) at room temperature for 16 hr. The cleaved protein was then dialyzed against His buffer and the His-tagged rTEV and undigested target

protein were removed via affinity chromatography. Purified SusG was dialyzed against 20 mM HEPES, 100 mM NaCl (pH 7.0) prior to crystallization.

Crystallization and Data Collection

All crystals were grown using the hanging-drop vapor-diffusion method, and Hampton Screen kits (Hampton Research) were used to determine initial conditions. Large single crystals of selenomethionine-substituted (SeMet) WT SusG were grown at room temperature using ~11 mg/ml of SusG and mother liquor that contained 18% polyethylene glycol (PEG) 4000, 200 mM magnesium acetate, and 100 mM HEPES (pH 7.5). The SeMet WT-SusG apo crystals were of the tetragonal space group $P4_1$ with unit cell dimensions of $a = b = 128.038$ Å, $c = 129.774$ Å. Crystals were serially transferred to a final cryoprotectant solution composed of 20% ethylene glycol in the above mother liquor, and the crystals were flash-frozen in liquid nitrogen. Crystals of SeMet WT SusG complexed with acarbose were grown at room temperature using a mother liquor containing 22% PEG 4000, 50 mM $LiSO_4$, and 100 mM HEPES (pH 7.5). Prior to crystallization, WT SusG (16 mg/ml) was premixed with 10 mM acarbose, 0.5 mM $CaCl_2$ and stored at 4°C from several days to about 2 weeks prior to mixing 1:1 with mother liquor for crystallization trials at room temperature. The SeMet WT-SusG/acarbose crystals were also of the space group $P4_1$, with unit cell dimensions of $a = b = 127.709$ Å, $c = 127.987$ Å. For data collection, crystals were serially transferred into a final cryoprotectant containing 24% PEG 4000, 200 mM NaCl, 50 mM $LiSO_4$, 10 mM acarbose, 100 mM HEPES (pH 7.5), and 19% ethylene glycol, and flash-frozen in liquid nitrogen.

Native SusG-D498N crystals complexed with maltoheptaose were grown at room temperature using a 9.3 mg/ml protein solution containing 10 mM maltoheptaose and 0.5 mM $CaCl_2$, and then diluted 1:1 with mother liquor containing 19%–20% PEG 4000, 100 mM $LiSO_4$, and 100 mM HEPES (pH 7.5). For data collection, crystals were serially transferred into cryoprotectant containing 20% PEG 4000, 50 mM NaCl, 100 mM $LiSO_4$, 10 mM maltoheptaose, and 15% ethylene glycol, and flash-frozen in liquid nitrogen prior to data collection.

Diffraction maxima for all three data sets were collected on a 3 × 3 tiled SBC3 CCD detector at the Structural Biology Center 19-ID beamline (Advanced Photon Source, Argonne National Laboratory, Argonne, IL, USA). X-ray data were processed with HKL3000 and scaled with SCALEPACK (Otwinowski and Minor, 1997). Data collection statistics are reported in Table S2.

X-Ray Structure Determination

The structure of SusG was solved using SAD phasing from the X-ray data collected from the SeMet WT-SusG apo crystals. The programs SHELXD and SHELXE (Sheldrick, 2008) were used to determine the initial positions of the selenomethionines and estimate the initial phases from the peak wavelength data set, followed by refinement of the heavy-atom parameters using MLPHARE from the CCP4 suite of programs (CCP4, 1994). Solvent flattening was performed using DM (Cowtan, 1994; Terwilliger, 2000), and ARP/wARP (Morris et al., 2003) was used for initial model building. Alternate cycles of manual model building in O (Jones et al., 1991) with maximum-likelihood refinement with CNS (Brunger et al., 1998) was then used to build and refine the 2.2 Å selenomethionine-substituted SusG structure. The structures of SusG complexed with acarbose and maltoheptaose were determined by molecular replacement using the program AMoRe (Navaza, 1994) from the CCP4 suite of programs (CCP4, 1994) with the apo WT-SusG structure as a search model. Alternate cycles of manual model building in O and refinement using CNS were combined to complete the models, and the geometry was analyzed using PROCHECK (Laskowski et al., 1993). In all three models, residue 634 was on the border of the allowed/disallowed region of the Ramachandran analysis. Initial coordinates and geometric constraints for the oligosaccharides were downloaded from the HIC-Up server (http://xray.bmc.uu.se/hicup). Relevant refinement statistics are presented in Table S2.

Thin-Layer Chromatography

The reaction byproducts of WT SusG with various starch substrates were analyzed by thin-layer chromatography. Each reaction contained 22 μg/ml of SusG in 20 mM HEPES (pH 7.0), 100 mM NaCl and 5 mg/ml of one of the following: amylose, amylopectin, pullulan, α-cyclodextrin, β-cyclodextrin, dextran, maltose, maltotriose, maltotetraose, maltopentaose, maltohexaose, maltoheptaose, or acarbose. After 5 hr at 37°C, 3 μl of each reaction was blotted onto a 20 cm × 20 cm, 500 μm thick Partisil PK6F silica gel 60 Å plate (Whatman). The spots were dried by incubation in an 80°C oven for 5 min, and then another 3 μl of each reaction was added to the correct spot. The plate was dried a second time and then transferred to a solvent chamber containing a 3:1:1 mixture of isopropanol:ethyl acetate:water. Two irrigations were performed, drying the plate between washes. Controls of each sugar (2–5 mg/ml) were also blotted on the plate for comparison and determination of the reaction products, as well as isomaltose and panose, two potential products of dextran and pullulan hydrolysis.

Starch Specificity Assay

The ability of WT SusG to degrade various forms of starch was tested by monitoring the reducing sugars released over time using a dinitrosalicylic acid (DNSA) -based assay (Bernfeld, 1955). Each 1 ml reaction contained 450 μl of buffer A (20 mM HEPES, 100 mM NaCl [pH 7.0]), 500 μl of 10 mg/ml of pullulan (Sigma P4516), soluble starch (from potato; Sigma S2004), amylopectin (from maize; Sigma 10120), α-cyclodextrin (Fluka 28705), or β-cyclodextrin (Sigma C4767) in buffer A, and 50 μl of 0.43 μg/ml WT SusG. Reactions were performed in triplicate at 37°C. The amount of reducing sugar was assayed at 0 and 10 min by mixing equal amounts of the reaction mixture with the DNSA reagent (1% DNSA, 0.2% phenol, 1% NaOH, 0.05% sodium sulfite, 0.004% glucose), followed by heating at 100°C for 15 min. Samples were incubated on ice for 5 min, equilibrated to room temperature, and the absorbance at 575 nm was measured. The amount of reducing sugar liberated was determined via a standard curve using maltose, and included the assayed starch substrate to account for intrinsic maltooligosaccharides. These results are summarized in Table 1.

SusG Mutant Activity on Various Substrates

The relative activities of WT SusG, *ΔCBM*, and *ΔSURF* were examined in a series of assays using equimolar concentrations of the proteins. Protein concentrations were determined using the Pierce BCA protein assay kit (Thermo Fisher Scientific) according to the manufacturer's instructions. The enzymes were tested for activity on the starch analog *p*-nitrophenyl-maltopentaose (PNP-G5) in a continuous spectrophotometric assay monitoring the increase in A_{420nm} over time. Each reaction contained 10 μM enzyme, 20 mM HEPES (pH 7.0), 1.0 mM $CaCl_2$, 100 mM NaCl, and 0.2 or 2.0 mM PNP-M5. After confirming that the three SusG enzymes had nearly identical initial velocities with PNP-M5 as a substrate, the enzymes were tested for their ability to degrade pullulan (Sigma P4516), soluble starch (partially hydrolyzed potato starch; Sigma S2004), and amylopectin (from maize; Sigma 10120) using the DNSA-based reducing sugar assay. Soluble starch and amylopectin solutions were dissolved in buffer A by brief heating, then cooled to room temperature. Reactions, performed in triplicate, were initiated by the addition of 1.8 ml of 5 mg/ml polysaccharide solution in buffer A to 200 μl of 0.5 μM enzyme. At 0 and 25 min incubation at 22°C, three 250 μl aliquots of each reaction were mixed with 250 μl of DNSA reagent and processed as previously described. The SusG enzymes were also assayed for their ability to degrade insoluble cornstarch (Sigma S4126). These reactions contained 2 ml of a 50 mg/ml slurry of buffer-washed cornstarch in buffer A. Reactions were initiated by the addition of 0.5 ml of 25 μM enzyme, and incubated at 37°C with vigorous agitation. At 0, 10, 20, 30, and 40 min, 310 μl of the reaction was removed and centrifuged for 1 min to pellet the starch. Two hundred fifty microliters of the supernatant was added to 250 μl of the DNSA reagent, and the assay was developed as previously described.

Insoluble Starch Binding

The SusG-D498N, *ΔCBM*-D498N, and CBM58 proteins were assayed for their ability to adsorb to insoluble cornstarch. Cornstarch was prepared by washing several times with an excess of double-distilled water, followed by buffer A. A 100 mg/ml suspension of cornstarch in buffer A was dispensed into 1.5 ml microfuge tubes for aliquots containing 2, 5, 10, 25, 38, 50, 75, and 100 mg of cornstarch. The polysaccharide was pelleted by centrifugation and all supernatant was carefully removed. To each aliquot, 0.45 ml of 1 mg/ml of SusG-D498N, *ΔCBM-D498*, CBM58, or bovine serum albumin (BSA) was added. The reactions were agitated at 22°C for 2 hr, and then centrifuged to pellet the cornstarch and bound protein. Two 25 μl aliquots of each sample were removed and the amount of protein was measured using the BCA assay.

A relative dilution factor was calculated based upon the amount of free BSA recovered from each starch aliquot to correct for any dilution due to the wet starch. The amount of bound protein as a fraction of the total protein versus mg of cornstarch is plotted in Figure 6. These data were analyzed with the program Prism (http://www.graphpad.com) by nonlinear regression analysis using the total binding equation

$$Y = \frac{B\max^{*}X}{K_D + X} + NS^{*}X + Background,$$

where Y is the fraction of protein bound, X is amount of cornstarch, K_D is the dissociation constant, Bmax is the maximum amount of bound protein, NS is the slope of nonspecific binding (Y/X), and Background is the nonspecific binding observed in the absence of ligand. Because the maximum amount of protein could not exceed 100%, Bmax was constrained to 1.0 during curve fitting.

Isothermal Titration Calorimetry

ITC measurements were carried out using a MicroCal VP-ITC titration calorimeter (MicroCal). CBM58 was dialyzed overnight against a solution containing 20 mM HEPES (pH 7.0) and 100 mM NaCl, and oligosaccharide solutions were prepared using dialysis buffer. Protein (0.57 mM) was placed in the reaction cell and the reference cell was filled with deionized water. After the temperature was equilibrated to 25°C, a first injection was performed using 2 μl, followed by 39 successive 6 μl injections of 5 mM maltoheptaose, β-cyclodextrin, or 6^3-α-D-glucosyl-maltotriose-maltotriose. The solution was stirred at 460 rpm while the resulting heat of reaction was measured. Baseline measurements were made using an identical injection regime in the absence of protein. The data were analyzed by fitting a one-site model with the MicroCal Origin software package.

SUPPLEMENTAL INFORMATION

Supplemental Information includes two tables and can be found with this article online at doi:10.1016/j.str.2009.12.010.

ACKNOWLEDGMENTS

This work was supported by a grant from the Missouri Life Sciences Trust Fund and funds from the Donald Danforth Plant Science Center.

Received: October 22, 2009
Revised: December 17, 2009
Accepted: December 18, 2009
Published: February 9, 2010

REFERENCES

Abe, A., Tonozuka, T., Sakano, Y., and Kamitori, S. (2004). Complex structures of *Thermoactinomyces vulgaris* R-47 α-amylase 1 with malto-oligosaccharides demonstrate the role of domain N acting as a starch-binding domain. J. Mol. Biol. *335*, 811–822.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402.

Anderson, K.L., and Salyers, A.A. (1989a). Biochemical evidence that starch breakdown by *Bacteroides thetaiotaomicron* involves outer membrane starch-binding sites and periplasmic starch-degrading enzymes. J. Bacteriol. *171*, 3192–3198.

Anderson, K.L., and Salyers, A.A. (1989b). Genetic evidence that outer membrane binding of starch is required for starch utilization by *Bacteroides thetaiotaomicron*. J. Bacteriol. *171*, 3199–3204.

Backhed, F., Ley, R.E., Sonnenburg, J.L., Peterson, D.A., and Gordon, J.I. (2005). Host-bacterial mutualism in the human intestine. Science *307*, 1915–1920.

Barends, T.R., Bultema, J.B., Kaper, T., van der Maarel, M.J., Dijkhuizen, L., and Dijkstra, B.W. (2007). Three-way stabilization of the covalent intermediate in amylomaltase, an α-amylase-like transglycosylase. J. Biol. Chem. *282*, 17242–17249.

Bernfeld, P. (1955). Amylases, α and β. Methods Enzymol. *1*, 149–158.

Bjursell, M.K., Martens, E.C., and Gordon, J.I. (2006). Functional genomic and metabolic studies of the adaptations of a prominent adult human gut symbiont, *Bacteroides thetaiotaomicron*, to the suckling period. J. Biol. Chem. *281*, 36269–36279.

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.

Boraston, A.B., Healey, M., Klassen, J., Ficko-Blean, E., Lammerts van Bueren, A., and Law, V. (2006). A structural and functional analysis of α-glucan recognition by family 25 and 26 carbohydrate-binding modules reveals a conserved mode of starch recognition. J. Biol. Chem. *281*, 587–598.

Brunger, A.T., Adams, P.D., Clore, G.M., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.-S., Kuszewski, J., Nilges, N., Pannu, N.S., Read, R.J., et al. (1998). Crystallography & NMR system: a new software system for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. Nucleic Acids Res. *37*, D233–D238.

CCP4 (Collaborative Computational Project, Number 4) (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

Cho, K.H., and Salyers, A.A. (2001). Biochemical analysis of interactions between outer membrane proteins that contribute to starch utilization by *Bacteroides thetaiotaomicron*. J. Bacteriol. *183*, 7224–7230.

Cowtan, K. (1994). DM: an automated procedure for phase improvement by density modification. Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography *31*, 34–38.

D'Elia, J.N., and Salyers, A.A. (1996a). Contribution of a neopullulanase, a pullulanase, and an α-glucosidase to growth of *Bacteroides thetaiotaomicron* on starch. J. Bacteriol. *178*, 7173–7179.

D'Elia, J.N., and Salyers, A.A. (1996b). Effect of regulatory protein levels on utilization of starch by *Bacteroides thetaiotaomicron*. J. Bacteriol. *178*, 7180–7186.

Fritzsche, H.B., Schwede, T., and Schulz, G.E. (2003). Covalent and three-dimensional structure of the cyclodextrinase from *Flavobacterium* sp. no. 92. Eur. J. Biochem. *270*, 2332–2341.

Holm, L., and Sander, C. (1995). Dali: a network tool for protein structure comparison. Trends Biochem. Sci. *20*, 478–480.

Hondoh, H., Kuriki, T., and Matsuura, Y. (2003). Three-dimensional structure and substrate binding of *Bacillus stearothermophilus* neopullulanase. J. Mol. Biol. *326*, 177–188.

Hooper, L.V., and Gordon, J.I. (2001). Commensal host-bacterial relationships in the gut. Science *292*, 1115–1118.

Imberty, A., Chanzy, H., Perez, S., Buleon, A., and Tran, V. (1988). The double-helical nature of the crystalline part of A-starch. J. Mol. Biol. *201*, 365–378.

Jones, T.A., Zou, J.-Y., and Cowan, S.W. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

Juge, N., Nohr, J., Le Gal-Coeffet, M.F., Kramhoft, B., Furniss, C.S., Planchot, V., Archer, D.B., Williamson, G., and Svensson, B. (2006). The activity of barley α-amylase on starch granules is enhanced by fusion of a starch binding domain from *Aspergillus niger* glucoamylase. Biochim. Biophys. Acta *1764*, 275–284.

Kadziola, A., Sogaard, M., Svensson, B., and Haser, R. (1998). Molecular structure of a barley α-amylase-inhibitor complex: implications for starch binding and catalysis. J. Mol. Biol. *278*, 205–217.

Kamitori, S., Kondo, S., Okuyama, K., Yokota, T., Shimura, Y., Tonozuka, T., and Sakano, Y. (1999). Crystal structure of *Thermoactinomyces vulgaris* R-47 α-amylase II (TVAII) hydrolyzing cyclodextrins and pullulan at 2.6 Å resolution. J. Mol. Biol. *287*, 907–921.

Kitamura, M., Okuyama, M., Tanzawa, F., Mori, H., Kitago, Y., Watanabe, N., Kimura, A., Tanaka, I., and Yao, M. (2008). Structural and functional analysis of a glycoside hydrolase family 97 enzyme from *Bacteroides thetaiotaomicron*. J. Biol. Chem. *283*, 36328–36337.

Koropatkin, N.M., Koppenaal, D.W., Pakrasi, H.B., and Smith, T.J. (2007). The structure of a cyanobacterial bicarbonate transport protein, CmpA. J. Biol. Chem. *282*, 2606–2614.

Koropatkin, N.M., Martens, E.C., Gordon, J.I., and Smith, T.J. (2008). Starch catabolism by a prominent human gut symbiont is directed by the recognition of amylose helices. Structure *16*, 1105–1115.

Lammerts van Bueren, A., and Boraston, A.B. (2007). The structural basis of α-glucan recognition by a family 41 carbohydrate-binding module from *Thermotoga maritima*. J. Mol. Biol. *365*, 555–560.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the sterochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.

Latorre-Garcia, L., Adam, A.C., Manzanares, P., and Polaina, J. (2005). Improving the amylolytic activity of *Saccharomyces cerevisiae* glucoamylase by the addition of a starch binding domain. J. Biotechnol. *118*, 167–176.

Leathers, T.D. (2003). Biotechnological production and applications of pullulan. Appl. Microbiol. Biotechnol. *62*, 468–473.

Lee, H.S., Kim, M.S., Cho, H.S., Kim, J.I., Kim, T.J., Choi, J.H., Park, C., Oh, B.H., and Park, K.H. (2002). Cyclomaltodextrinase, neopullulanase, and maltogenic amylase are nearly indistinguishable from each other. J. Biol. Chem. *277*, 21891–21897.

Ley, R.E., Peterson, D.A., and Gordon, J.I. (2006). Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell *124*, 837–848.

Li, C., Begum, A., Numao, S., Park, K.H., Withers, S.G., and Brayer, G.D. (2005). Acarbose rearrangement mechanism implied by the kinetic and structural analysis of human pancreatic α-amylase in complex with analogues and their elongated counterparts. Biochemistry *44*, 3347–3357.

Machovic, M., and Janecek, S. (2006). Starch-binding domains in the post-genome era. Cell. Mol. Life Sci. *63*, 2710–2724.

Machovic, M., Svensson, B., MacGregor, E.A., and Janecek, S. (2005). A new clan of CBM families based on bioinformatics of starch-binding domains from families CBM20 and CBM21. FEBS J. *272*, 5497–5513.

Mandelman, D., Belaich, A., Belaich, J.P., Aghajari, N., Driguez, H., and Haser, R. (2003). X-ray crystal structure of the multidomain endoglucanase Cel9G from *Clostridium cellulolyticum* complexed with natural and synthetic cello-oligosaccharides. J. Bacteriol. *185*, 4127–4135.

Martens, E.C., Chiang, H.C., and Gordon, J.I. (2008). Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. Cell Host Microbe *4*, 447–457.

Mazmanian, S.K., Liu, C.H., Tzianabos, A.O., and Kasper, D.L. (2005). An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. Cell *122*, 107–118.

McCarter, J.D., and Withers, S.G. (1994). Mechanisms of enzymatic glycoside hydrolysis. Curr. Opin. Struct. Biol. *4*, 885–892.

McCarter, J.D., and Withers, S.G. (1996). Unequivocal identification of Asp-214 as the catalytic nucleophile of *Saccharomyces cerevisiae* α-glucosidase using 5-fluoro glycosyl fluorides. J. Biol. Chem. *271*, 6889–6894.

Morris, R.J., Perrakis, A., and Lamzin, V.S. (2003). ARP/wARP and automatic interpretation of protein electron density maps. Methods Enzymol. *374*, 229–244.

Navaza, J. (1994). AMoRe: an automated package for molecular replacement. Acta Crystallogr. A *50*, 157–163.

Noverr, M.C., and Huffnagle, G.B. (2004). Does the microbiota regulate immune responses outside the gut? Trends Microbiol. *12*, 562–568.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. In Methods in Enzymology, C.W.J. Carter and R.M. Sweet, eds. (New York: Academic Press), pp. 307–326.

Payan, F., and Qian, M. (2003). Crystal structure of the pig pancreatic α-amylase complexed with malto-oligosaccharides. J. Protein Chem. *22*, 275–284.

Penders, J., Stobberingh, E.E., van den Brandt, P.A., and Thijs, C. (2007). The role of the intestinal microbiota in the development of atopic disorders. Allergy *62*, 1223–1236.

Penninga, D., van der Veen, B.A., Knegtel, R.M., van Hijum, S.A., Rozeboom, H.J., Kalk, K.H., Dijkstra, B.W., and Dijkhuizen, L. (1996). The raw starch binding domain of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251. J. Biol. Chem. *271*, 32777–32784.

Qian, M., Haser, R., and Payan, F. (1995). Carbohydrate binding sites in a pancreatic α-amylase-substrate complex, derived from X-ray structure analysis at 2.1 Å resolution. Protein Sci. *4*, 747–755.

Qian, M., Nahoum, V., Bonicel, J., Bischoff, H., Henrissat, B., and Payan, F. (2001). Enzyme-catalyzed condensation reaction in a mammalian α-amylase. High-resolution structural analysis of an enzyme-inhibitor complex. Biochemistry *40*, 7700–7709.

Ragunath, C., Manuel, S.G., Venkataraman, V., Sait, H.B., Kasinathan, C., and Ramasubbu, N. (2008). Probing the role of aromatic residues at the secondary saccharide-binding sites of human salivary α-amylase in substrate hydrolysis and bacterial binding. J. Mol. Biol. *384*, 1232–1248.

Robert, X., Haser, R., Mori, H., Svensson, B., and Aghajari, N. (2005). Oligosaccharide binding to barley α-amylase 1. J. Biol. Chem. *280*, 32968–32978.

Roujeinikova, A., Raasch, C., Sedelnikova, S., Liebl, W., and Rice, D.W. (2002). Crystal structure of *Thermotoga maritima* 4-α-glucanotransferase and its acarbose complex: implications for substrate specificity and catalysis. J. Mol. Biol. *321*, 149–162.

Sakon, J., Irwin, D., Wilson, D.B., and Karplus, P.A. (1997). Structure and mechanism of endo/exocellulase E4 from *Thermomonospora fusca*. Nat. Struct. Biol. *4*, 810–818.

Sevcik, J., Hostinova, E., Solovicova, A., Gasperik, J., Dauter, Z., and Wilson, K.S. (2006). Structure of the complex of a yeast glucoamylase with acarbose reveals the presence of a raw starch binding site on the catalytic domain. FEBS J. *273*, 2161–2171.

Sharma, M.D., and Garber, A.J. (2009). What is the best treatment for prediabetes? Curr. Diab. Rep. *9*, 335–341.

Sheldrick, G.M. (2008). A short history of SHELX. Acta Crystallogr. A *64*, 112–122.

Shipman, J.A., Cho, K.H., Siegel, H.A., and Salyers, A.A. (1999). Physiological characterization of SusG, an outer membrane protein essential for starch utilization by *Bacteroides thetaiotaomicron*. J. Bacteriol. *181*, 7206–7211.

Shipman, J.A., Berleman, J.E., and Salyers, A.A. (2000). Characterization of four outer membrane proteins involved in binding starch to the cell surface of *Bacteroides thetaiotaomicron*. J. Bacteriol. *182*, 5365–5372.

Sivakumar, N., Li, N., Tang, J.W., Patel, B.K., and Swaminathan, K. (2006). Crystal structure of AmyA lacks acidic surface and provide insights into protein stability at poly-extreme condition. FEBS Lett. *580*, 2646–2652.

Sogaard, M., Kadziola, A., Haser, R., and Svensson, B. (1993). Site-directed mutagenesis of histidine 93, aspartic acid 180, glutamic acid 205, histidine 290, and aspartic acid 291 at the active site and tryptophan 279 at the raw starch binding site in barley α-amylase 1. J. Biol. Chem. *268*, 22480–22484.

Sonnenburg, J.L., Xu, J., Leip, D.D., Chen, C.H., Westover, B.P., Weatherford, J., Buhler, J.D., and Gordon, J.I. (2005). Glycan foraging in vivo by an intestine-adapted bacterial symbiont. Science *307*, 1955–1959.

Strokopytov, B., Penninga, D., Rozeboom, H.J., Kalk, K.H., Dijkhuizen, L., and Dijkstra, B.W. (1995). X-ray structure of cyclodextrin glycosyltransferase complexed with acarbose. Implications for the catalytic mechanism of glycosidases. Biochemistry *34*, 2234–2240.

Sumitani, J., Tottori, T., Kawaguchi, T., and Arai, M. (2000). New type of starch-binding domain: the direct repeat motif in the C-terminal region of *Bacillus* sp. no. 195 α-amylase contributes to starch binding and raw starch degrading. Biochem. J. *350*, 477–484.

Tan, T.C., Mijts, B.N., Swaminathan, K., Patel, B.K., and Divne, C. (2008). Crystal structure of the polyextremophilic α-amylase AmyB from *Halothermothrix orenii*: details of a productive enzyme-substrate complex and an N domain with a role in binding raw starch. J. Mol. Biol. *378*, 852–870.

Tancula, E., Feldhaus, M.J., Bedzyk, L.A., and Salyers, A.A. (1992). Location and characterization of genes involved in binding of starch to the surface of *Bacteroides thetaiotaomicron*. J. Bacteriol. *174*, 5609–5616.

Terwilliger, T.C. (2000). Maximum-likelihood density modification. Acta Crystallogr. D Biol. Crystallogr. *56*, 965–972.

Uitdehaag, J.C., Mosi, R., Kalk, K.H., van der Veen, B.A., Dijkhuizen, L., Withers, S.G., and Dijkstra, B.W. (1999). X-ray structures along the reaction pathway of cyclodextrin glycosyltransferase elucidate catalysis in the α-amylase family. Nat. Struct. Biol. *6*, 432–436.

Umetsu, D.T., McIntire, J.J., Akbari, O., Macaubas, C., and DeKruyff, R.H. (2002). Asthma: an epidemic of dysregulated immunity. Nat. Immunol. *3*, 715–720.

Van Duyne, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L., and Clardy, J. (1993). Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. J. Mol. Biol. *229*, 105–124.

Xu, J., Bjursell, M.K., Himrod, J., Deng, S., Carmichael, L.K., Chiang, H.C., Hooper, L.V., and Gordon, J.I. (2003). A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis. Science *299*, 2074–2076.

Xu, J., Mahowald, M.A., Ley, R.E., Lozupone, C.A., Hamady, M., Martens, E.C., Henrissat, B., Coutinho, P.M., Minx, P., Latreille, P., et al. (2007). Evolution of symbiotic bacteria in the distal human intestine. PLoS Biol. *5*, e156.

