AMERICAN SOCIETY FOR MICROBIOLOGY | mBio®

RESEARCH ARTICLE

Molecular Dissection of Xyloglucan Recognition in a Prominent Human Gut Symbiont

Alexandra S. Tauzin,[a]* Kurt J. Kwiatkowski,[b] Nicole I. Orlovsky,[b] Christopher J. Smith,[b] A. Louise Creagh,[c] Charles A. Haynes,[c] Zdzislaw Wawrzak,[d] Harry Brumer,[a] Nicole M. Koropatkin[b]

Michael Smith Laboratories and Department of Chemistry, University of British Columbia, Vancouver, British Columbia, Canada[a]; Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, Michigan, USA[b]; Michael Smith Laboratories and Department of Chemical and Biological Engineering, University of British Columbia, Vancouver, British Columbia, Canada[c]; The Synchrotron Research Center, Life Science Collaborative Access Team, Northwestern University, Argonne, Illinois, USA[d]

* Present address: Alexandra Tauzin, Université de Toulouse, INSA, UPS, INP, LISBP, Toulouse, France; INRA, UMR792 Ingénierie des Systèmes Biologiques et des Procédés, Toulouse, France; CNRS, UMR5504, Toulouse, France.

ABSTRACT **Polysaccharide utilization loci (PUL) within the genomes of resident human gut *Bacteroidetes* are central to the metabolism of the otherwise indigestible complex carbohydrates known as "dietary fiber." However, functional characterization of PUL lags significantly behind sequencing efforts, which limits physiological understanding of the human-bacterial symbiosis. In particular, the molecular basis of complex polysaccharide recognition, an essential prerequisite to hydrolysis by cell surface glycosidases and subsequent metabolism, is generally poorly understood. Here, we present the biochemical, structural, and reverse genetic characterization of two unique cell surface glycan-binding proteins (SGBPs) encoded by a xyloglucan utilization locus (XyGUL) from *Bacteroides ovatus*, which are integral to growth on this key dietary vegetable polysaccharide. Biochemical analysis reveals that these outer membrane-anchored proteins are in fact exquisitely specific for the highly branched xyloglucan (XyG) polysaccharide. The crystal structure of SGBP-A, a SusD homolog, with a bound XyG tetradecasaccharide reveals an extended carbohydrate-binding platform that primarily relies on recognition of the β-glucan backbone. The unique, tetra-modular structure of SGBP-B is comprised of tandem Ig-like folds, with XyG binding mediated at the distal C-terminal domain. Despite displaying similar affinities for XyG, reverse-genetic analysis reveals that SGBP-B is only required for the efficient capture of smaller oligosaccharides, whereas the presence of SGBP-A is more critical than its carbohydrate-binding ability for growth on XyG. Together, these data demonstrate that SGBP-A and SGBP-B play complementary, specialized roles in carbohydrate capture by *B. ovatus* and elaborate a model of how vegetable xyloglucans are accessed by the *Bacteroidetes*.**

IMPORTANCE **The *Bacteroidetes* are dominant bacteria in the human gut that are responsible for the digestion of the complex polysaccharides that constitute "dietary fiber." Although this symbiotic relationship has been appreciated for decades, little is currently known about how *Bacteroidetes* seek out and bind plant cell wall polysaccharides as a necessary first step in their metabolism. Here, we provide the first biochemical, crystallographic, and genetic insight into how two surface glycan-binding proteins from the complex *Bacteroides ovatus* xyloglucan utilization locus (XyGUL) enable recognition and uptake of this ubiquitous vegetable polysaccharide. Our combined analysis illuminates new fundamental aspects of complex polysaccharide recognition, cleavage, and import at the *Bacteroidetes* cell surface that may facilitate the development of prebiotics to target this phylum of gut bacteria.**

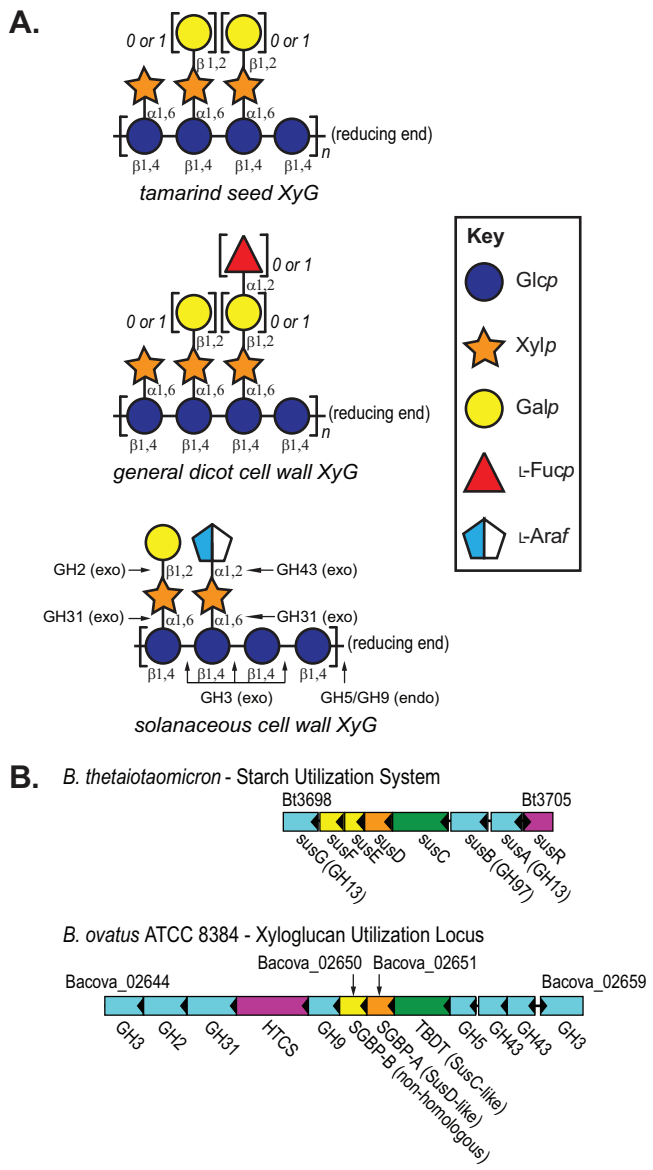
Received 7 December 2015 **Accepted** 23 March 2016 **Published** 26 April 2016

Citation Tauzin AS, Kwiatkowski KJ, Orlovsky NI, Smith CJ, Creagh AL, Haynes CA, Wawrzak Z, Brumer H, Koropatkin NM. 2016. Molecular dissection of xyloglucan recognition in a prominent human gut symbiont. mBio 7(2):e02134-15. doi:10.1128/mBio.02134-15.

Invited Editor Justin Sonnenburg, Stanford University School of Medicine **Editor** Margaret J. McFall-Ngai, University of Hawaii

Address correspondence to Harry Brumer, brumer@msl.ubc.ca, or Nicole M. Koropatkin, nkoropat@umich.edu.

The human gut microbiota influences the course of human development and health, playing key roles in immune stimulation (1), intestinal cell proliferation (2), and metabolic balance (3, 4). This microbial community is largely bacterial, with the *Bacteroidetes*, *Firmicutes*, and *Actinobacteria* comprising the dominant phyla (5, 6). The ability to acquire energy from carbohydrates of dietary or host origin is central to the adaptation of human gut bacterial species to their niche. More importantly, this makes diet a tractable way to manipulate the abundance and metabolic output of the microbiota toward improved human health. However, there is a paucity of data regarding how the vast array of complex carbohydrate structures are selectively recognized and imported by members of the microbiota, a critical process that enables these organisms to thrive in the competitive gut environment. The human gut bacteria *Bacteroidetes* share a profound capacity for dietary glycan degradation, with many species containing >250 predicted carbohydrate-active enzymes (CAZymes), compared to 50 to 100 within many *Firmicutes* and only 17 in the human genome devoted toward carbohydrate utilization (7). A remarkable feature of the *Bacteroidetes* is the packaging of genes for carbohydrate



catabolism into discrete polysaccharide utilization loci (PUL) (8), which are transcriptionally regulated by specific substrate signatures (9–11). The archetypal PUL-encoded system is the starch utilization system (Sus) (Fig. 1B) of *Bacteroides thetaiotaomicron* (12, 13). The Sus includes a lipid-anchored, outer membrane endo-amylase, SusG (14, 15); a TonB-dependent transporter (TBDT), SusC, which imports oligosaccharides with the help of an associated starch-binding protein, SusD (13, 16, 17); two additional carbohydrate-binding lipoproteins, SusE and SusF (18); and two periplasmic exo-glucosidases, SusA and SusB, which generate glucose for transport into the cytoplasm (19). The importance of PUL as a successful evolutionary strategy is underscored by the observation that *Bacteroidetes* such as *B. thetaiotaomicron* and *Bacteroides ovatus* devote ~18% of their genomes to these systems (11). Moving beyond seminal genomic and transcriptomic analyses, the current state-of-the-art PUL characterization involves combined reverse-genetic, biochemical, and structural studies to illuminate the molecular details of PUL function (20–23).

We recently reported the detailed molecular characterization of a PUL that confers the ability of the human gut commensal *B. ovatus* ATCC 8483 to grow on a prominent family of plant cell wall glycans, the xyloglucans (XyG) (23). XyG variants (Fig. 1A) constitute up to 25% of the dry weight of common vegetables (24, 25). Analogous to the Sus locus, the xyloglucan utilization locus (XyGUL) encodes a cohort of carbohydrate-binding, -hydrolyzing, and -importing proteins (Fig. 1B and C). The number of glycoside hydrolases (GHs) encoded by the XyGUL is, however, more expansive than that by the Sus locus (Fig. 1B), which reflects the greater complexity of glycosidic linkages found in XyG vis-à-vis starch (10). Whereas our previous study focused on the characterization of the linkage specificity of these GHs (23), a key outstanding question regarding this locus is how XyG recognition is mediated at the cell surface.

In the archetypal starch utilization system of *B. thetaiotaomicron*, starch binding to the cell surface is mediated at eight distinct starch-binding sites distributed among four surface glycan-binding proteins (SGBPs): two within the amylase SusG, one within SusD, two within SusE, and three within SusF (15, 16, 18). The functional redundancy of many of these sites is high: whereas SusD is essential for growth on starch, combined mutations of the SusE, SusF, and SusG binding sites are required to impair growth on the polysaccharide (16, 26). *Bacteroidetes* PUL ubiquitously encode homologs of SusC and SusD, as well as proteins whose genes are immediately downstream of *susD*, akin to *susE/F*, and these are typically annotated as "putative lipoproteins" (8, 27). The genes coding for these proteins, sometimes referred to as "*susE/F* positioned," display products with a wide variation in amino acid sequence and which have little or no homology to

FIG 1 Xyloglucan and the *Bacteroides ovatus* xyloglucan utilization locus (XyGUL). (A) Representative structures of common xyloglucans (60) using the Consortium for Functional Glycomics Symbol Nomenclature (http://www.functionalglycomics.org/static/consortium/Nomenclature.shtml). Cleavage sites for *Bo*XyGUL glycosidases (GHs) are indicated for solanaceous xyloglucan. (B) *Bt*Sus and *Bo*XyGUL. (C) Localization of *Bo*XyGUL-encoded proteins in cellular membranes and concerted modes of action in the degradation of xyloglucans to monosaccharides. The location of SGBP-A/B is presented in this work; the location of GH5 has been empirically determined, and the enzymes have been placed based upon their predicted cellular location (23).

other PUL-encoded proteins or known carbohydrate-binding proteins (28, 29). As the Sus SGBPs remain the only structurally characterized cohort to date, we therefore wondered whether such glycan binding and function are extended to other PUL that target more complex and heterogeneous polysaccharides, such as XyG.

We describe here the detailed functional and structural characterization of the noncatalytic SGBPs encoded by Bacova_02651 and Bacova_02650 of the XyGUL, here referred to as SGBP-A and SGBP-B, to elucidate their molecular roles in carbohydrate acquisition *in vivo*. Combined biochemical, structural, and reverse-genetic approaches clearly illuminate the distinct, yet complementary, functions that these two proteins play in XyG recognition as it impacts the physiology of *B. ovatus*. These data extend our current understanding of the Sus-like glycan uptake paradigm within the *Bacteroidetes* and reveals how the complex dietary polysaccharide xyloglucan is recognized at the cell surface.

RESULTS AND DISCUSSION

SGBP-A and SGBP-B are cell-surface-localized, xyloglucan-specific binding proteins. SGBP-A, encoded by the XyGUL locus tag Bacova_02651 (Fig. 1B), shares 26% amino acid sequence identity (40% similarity) with its homolog, *B. thetaiotaomicron* SusD (16), and similar homology with the SusD-like proteins encoded within syntenic XyGUL identified in our earlier work (23). In contrast, SGBP-B, encoded by locus tag Bacova_02650, displays little sequence similarity to the products of similarly positioned genes in syntenic XyGUL nor to any other gene product among the diversity of *Bacteroidetes* PUL (18, 27). Whereas sequence similarity among SusC/SusD homolog pairs often serves as a hallmark for PUL identification, the sequence similarities of downstream genes encoding SGBPs are generally too low to allow reliable bioinformatic classification of their products into protein families, let alone prediction of function (30). Hence, there is a critical need for the elucidation of detailed structure-function relationships among PUL SGBPs, in light of the manifold glycan structures in nature.

Immunofluorescence of formaldehyde-fixed, nonpermeabilized cells grown in minimal medium with XyG as the sole carbon source to induce XyGUL expression (23), reveals that both SGBP-A and SGBP-B are presented on the cell surface by N-terminal lipidation, as predicted by signal peptide analysis with SignalP (Fig. 2). Here, the SGBPs very likely work in concert with the cell-surface-localized endo-xyloglucanase *B. ovatus* GH5 (*Bo*GH5) (23) to recruit and cleave XyG for subsequent periplasmic import via the SusC-like TBDT of the XyGUL (Fig. 1B and C).

In our initial study focused on the functional characterization of the glycoside hydrolases of the XyGUL, we reported preliminary affinity PAGE and isothermal titration calorimetry (ITC) data indicating that both SGBP-A and SGBP-B are competent xyloglucan-binding proteins (affinity constant [K_a] values of 3.74×10^5 M^{-1} and 4.98×10^4 M^{-1}, respectively [23]). Additional affinity PAGE analysis (Fig. 3) demonstrates that SGBP-A also has moderate affinity for the artificial soluble cellulose derivative hydroxyethyl cellulose [HEC; a $\beta(1 \rightarrow 4)$-glucan] and limited affinity for mixed-linkage $\beta(1\rightarrow3)/\beta(1\rightarrow4)$-glucan (MLG) and glucomannan (GM; mixed glucosyl and mannosyl backbone), which together indicate general binding to polysaccharide backbone residues and major contributions from side-chain recognition. In contrast, SGBP-B bound to HEC more weakly than SGBP-A and did not bind to MLG or GM. Neither SGBP recognized galactomannan (GGM), starch, carboxymethylcellulose, or mucin (see Fig. S1 in the supplemental material). Together, these results highlight the high specificities of SGBP-A and SGBP-B for XyG, which is concordant with their association with XyG-specific GHs in the XyGUL, as well as transcriptomic analysis indicating that *B. ovatus* has discrete PUL for MLG, GM, and GGM (11). Notably, the absence of carbohydrate-binding modules in the GHs encoded by the XyGUL (23) implies that noncatalytic recognition of xyloglucan is mediated entirely by SGBP-A and -B.

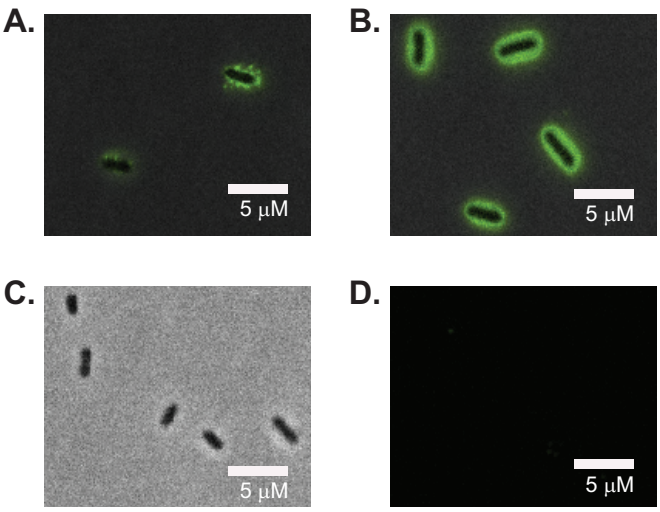


FIG 2 SGBP-A and SGBP-B visualized by immunofluorescence. Formalin-fixed, nonpermeabilized *B. ovatus* cells were grown in minimal medium plus XyG, probed with custom rabbit antibodies to SGBP-A or SGBP-B, and then stained with Alexa Fluor 488 goat anti-rabbit IgG. (A) Overlay of bright-field and FITC images of *B. ovatus* cells labeled with anti-SGBP-A. (B) Overlay of bright-field and FITC images of *B. ovatus* cells labeled with anti-SGBP-B. (C) Bright-field image of ΔSGBP-B cells labeled with anti-SGBP-B antibodies. (D) FITC images of ΔSGBP-B cells labeled with anti-SGBP-B antibodies. Cells lacking SGBP-A (ΔSGBP-A) do not grow on XyG and therefore could not be tested in parallel.

The vanguard endo-xyloglucanase of the XyGUL, *Bo*GH5, preferentially cleaves the polysaccharide at unbranched glucosyl residues to generate xylogluco-oligosaccharides (XyGOs) comprising a Glc_4 backbone with variable side-chain galactosylation ($XyGO_1$) (Fig. 1A; $n = 1$) as the limit of digestion products *in vitro* (23); controlled digestion and fractionation by size exclusion chromatography allow the production of higher-order oligosaccharides (e.g., $XyGO_2$) (Fig. 1A; $n = 2$). ITC demonstrates that SGBP-A binds to XyG polysaccharide and $XyGO_2$ (based on a Glc_8 backbone) with essentially equal affinities, while no binding of $XyGO_1$ (Glc_4 backbone) was detectable (Table 1; see Fig. S2 and S3 in the supplemental material). Similarly, SGBP-B also bound to XyG and $XyGO_2$ with approximately equal affinities, although in both cases, K_a values were nearly 10-fold lower than those for SGBP-A. Also in contrast to SGBP-A, SGBP-B also bound to $XyGO_1$, yet the affinity for this minimal repeating unit was poor, with a K_a value of ca. 1 order of magnitude lower than for XyG and $XyGO_2$. Together, these data clearly suggest that polysaccharide binding of both SGBPs is fulfilled by a dimer of the minimal repeat, corresponding to $XyGO_2$ (cf. Fig. 1A). The observation by affinity PAGE that these proteins specifically recognize XyG is further substantiated by their lack of binding for the undecorated oligosaccharide cellotetraose (Table 1; see Fig. S3). Furthermore,

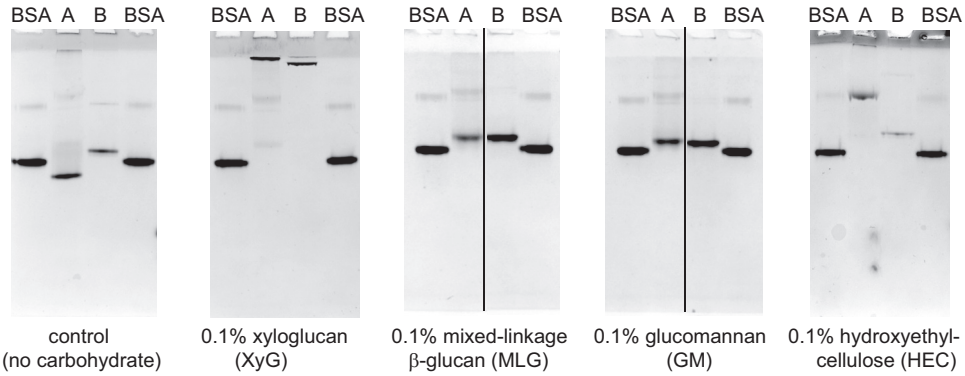


FIG 3 SGBP-A and SGBP-B preferentially bind xyloglucan. Affinity electrophoresis (10% acrylamide) of SGBP-A and SGBP-B with BSA as a control protein. All samples were loaded on the same gel next to the BSA controls; thin black lines indicate where intervening lanes were removed from the final image for both space and clarity. The percentage of polysaccharide incorporated into each native gel is displayed.

SGBP-A binds cellohexaose with ~770-fold weaker affinity than XyG, while SGBP-B displays no detectable binding to this linear hexasaccharide. To provide molecular-level insight into how the XyGUL SGBPs equip *B. ovatus* to specifically harvest XyG from the gut environment, we performed X-ray crystallography analysis of both SGBP-A and SGPB-B in oligosaccharide-complex forms.

SGBP-A is a SusD homolog with an extensive glycan-binding platform. As anticipated by sequence similarity, the high-resolution tertiary structure of *apo*-SGBP-A (1.36 Å, R_{work} = 14.7%, R_{free} = 17.4%, residues 28 to 546) (Table 2) displays the canonical "SusD-like" protein fold dominated by four tetratricopeptide repeat (TPR) motifs that cradle the rest of the structure (Fig. 4A) (31). Specifically, SGBP-A overlays *B. thetaiotaomicron* SusD (*Bt*SusD) with a root mean square deviation (RMSD) value of 2.2 Å for 363 Cα pairs, which is notable given the 26% amino acid identity (40% similarity) between these homologs (Fig. 4C). Cocrystallization of SGBP-A with $XyGO_2$ generated a substrate complex structure (2.3 Å, R_{work} = 21.8%, R_{free} = 24.8%, residues 36 to 546) (Fig. 4A and B; Table 2) that revealed the distinct binding-site architecture of the XyG binding protein. The SGBP-A:$XyGO_2$ complex superimposes closely with the *apo* structure (RMSD of 0.6 Å) and demonstrates that no major conformational change occurs upon substrate binding; small deviations in the orientation of several surface loops are likely the result of differential crystal packing. It is particularly notable that although the location of the ligand-binding site is conserved between SGBP-A and SusD, that of SGBP-A displays an ~29-Å-long aromatic platform to accommodate the extended, linear XyG chain (see reference 32 for a review of XyG secondary structure), versus the shorter, ~18-Å-long, site within SusD that complements the helical conformation of amylose (16) (Fig. 4C and D).

Seven of the eight backbone glucosyl residues of $XyGO_2$ could be convincingly modeled in the ligand electron density, and only two $\alpha(1\rightarrow6)$-linked xylosyl residues were observed (Fig. 4B; cf. Fig. 1). Indeed, the electron density for the ligand suggests some disorder, which may arise from multiple oligosaccharide orientations along the binding site. Three aromatic residues—W82, W283, W306—comprise the flat platform that stacks along the naturally twisted β-glucan backbone (Fig. 4E). The functional importance of this platform is underscored by the observation that the W82A W283A W306A mutant of SGBP-A, designated SGBP-A*, is completely devoid of XyG affinity (Table 3; see Fig. S4 in the supplemental material). Dissection of the individual contribution of these residues reveals that the W82A mutant displays a significant 4.9-fold decrease in the K_a value for XyG, while the W306A substitution completely abolishes XyG binding. Contrasting with the clear importance of these hydrophobic interactions, there are remarkably few hydrogen-bonding interactions with the ligand, which are provided by R65, N83, and S308, which are proximal to Glc5 and Glc3. Most surprising in light of the saccharide-binding data, however, was a lack of extensive recognition of the XyG side chains; only Y84 appeared to provide a hydrophobic interface for a xylosyl residue (Xyl1).

SGBP-B has a multimodular structure with a single, C-terminal glycan-binding domain. The crystal structure of full-

TABLE 1 Summary of thermodynamic parameters for wild-type SGBP-A and SGBP-B obtained by isothermal titration calorimetry at 25°C[a]

	K_a (M^{-1})		ΔG (kcal · mol^{-1})		ΔH (kcal · mol^{-1})		$T\Delta S$ (kcal · mol^{-1})	
Carbohydrate	SGBP-A	SGBP-B	SGBP-A	SGBP-B	SGBP-A	SGBP-B	SGBP-A	SGBP-B
XyG[b]	$(4.4 \pm 0.1) \times 10^5$	$(5.7 \pm 0.2) \times 10^4$	−7.7	−6.5	−14 ± 3	−14 ± 2	−6.5	−7.6
$XyGO_2$[c]	3.0×10^5	2.0×10^4	−7.5	−5.9	−17.2	−17.6	−9.7	−11.7
$XyGO_1$	NB[d]	$(2.4 \pm 0.1) \times 10^3$	NB	−4.6	NB	−4.4 ± 0.2	NB	0.2
Cellohexaose	568.0 ± 291.0	NB	−3.8	NB	−16 ± 8	NB	−12.7	NB
Cellotetraose	NB	NB	NB	NB	NB	NB	NB	NB

[a] Shown are average values ± standard errors from two independent titrations, unless otherwise indicated.
[b] Binding thermodynamics for XyG based on the concentration of the binding unit, $XyGO_2$.
[c] Values from a single titration.
[d] NB, no binding observed.

TABLE 2 X-ray data collection and refinement statistics

Parameter	Value(s) for[a]: SGBP-A *apo*	SGBP-A/$XyGO_2$	SGBP-B/$XyGO_2$	SGBP-B (CD)/$XyGO_2$
PDB ID no.	5E75	5E76	5E7G	5E7H
Resolution (Å)	21.48–1.36 (1.409–1.36)	56.13–2.3 (2.382–2.3)	39.19–2.37 (2.455–2.37)	30.69–1.57 (1.626–1.570)
Space group	$P2_1$	I422	R32	$P6_122$
Unit cell dimensions, *a*, *b*, *c* (Å)	52.8, 81.4, 57.7; $\beta = 107.85°$	131.5, 131.5, 188	207.4, 207.4, 117.9	87.1, 87.1, 201.6
No. of reflections				
Total	355,272 (26,772)	1,068,014 (102,923)	324,544 (32,355)	1,366,812 (129,645)
Unique	99,136 (9,762)	36,775 (3,625)	39,362 (3,898)	62,808 (6,068)
Multiplicity	3.6 (2.7)	29.0 (28.4)	8.2 (8.3)	21.8 (21.4)
Completeness (%)	99.71 (98.82)	99.63 (99.42)	99.96 (100.00)	98.4 (96.98)
Mean $I/\sigma\langle I\rangle$	15.57 (2.29)	24.93 (6.71)	20.98 (2.36)	38.52 (5.03)
Wilson B-factor	11.91	31.14	43.91	17.86
R_{merge}	0.04759 (0.4513)	0.1428 (0.7178)	0.09159 (1.197)	0.05559 (0.7748)
$CC_{1/2}$[b]	0.999 (0.759)	0.999 (0.982)	0.999 (0.794)	1.000 (0.933)
CC*[c]	1.000 (0.929)	1.000 (0.995)	1.000 (0.941)	1.000 (0.982)
R_{work}	0.1468 (0.2597)	0.2178 (0.2788)	0.1975 (0.3018)	0.1560 (0.2008)
R_{free}	0.1738 (0.2632)	0.2482 (0.2978)	0.2260 (0.3219)	0.1712 (0.2019)
No. of non-hydrogen atoms				
All	4,562	4,319	3,678	2,328
Macromolecules	4,079	3,974	3,425	1,985
Ligands	39	116	127	25
Water	444	229	126	318
No. of protein residues	506	492	446	260
RMSD				
Bond length (Å)	0.008	0.007	0.005	0.009
Bond angle (°)	1.15	0.96	0.87	1.18
Ramachandran statistics				
Favored (%)	98	95	97	98
Outliers (%)	0	0.41	0.23	0
Clash score	0.5	2.13	0.86	1.27
Avg B-factors				
All	16.1	53.2	53	25.4
Macromolecules	15.2	53.5	52.5	22.9
Ligands	24.7	61	71.1	47
Solvent	24.4	42.9	47.6	39

[a] Numbers in parentheses are for the highest-resolution shell.
[b] $CC_{1/2}$, Pearson correlation coefficient between the average intensities of each subset.
[c] CC*, Pearson correlation coefficient for correlation between the observed data set and true signal.

length SGBP-B in complex with $XyGO_2$ (2.37 Å, R_{work} = 19.9%, R_{free} = 23.9%, residues 34 to 489) (Table 2) revealed an extended structure composed of three tandem immunoglobulin (Ig)-like domains (domains A, B, and C) followed at the C terminus by a novel xyloglucan-binding domain (domain D) (Fig. 5A). Domains A, B, and C display similar β-sandwich folds; domains B (residues 134 to 230) and C (residues 231 to 313) can be superimposed onto domain A (residues 34 to 133) with RMSDs of 1.1 and 1.2 Å, respectively, for 47 atom pairs (23% and 16% sequence identity, respectively). These domains also display similarity to the C-terminal β-sandwich domains of many GH13 enzymes, including the cyclodextrin glucanotransferase of *Geobacillus stearothermophilus* (Fig. 5B). Such domains are not typically involved in carbohydrate binding. Indeed, visual inspection of the SGBP-B structure, as well as individual production of the A and B domains and affinity PAGE analysis (see Fig. S5 in the supplemental material), indicates that these domains do not contribute to XyG capture. On the other hand, production of the fused domains C and D in tandem (SGBP-B residues 230 to 489) retains complete binding of xyloglucan *in vitro*, with the observed slight increase in affinity likely arising from a reduced potential for steric hindrance of the smaller protein construct during polysaccharide interactions (Table 3). While neither the full-length protein nor domain D displays structural homology to known XyG-binding proteins, the topology of SGBP-B resembles the xylan-binding protein Bacova_04391 (PDB 3ORJ) encoded within a xylan-targeting PUL of *B. ovatus* (22) (Fig. 5C). The structure-based alignment of these proteins reveals 17% sequence identity, with a core RMSD of 3.6 Å for 253 aligned residues. While there is no substrate-complexed structure of Bacova_04391 available, the binding site is predicted to include W241 and Y404 (22), which are proximal to the XyGO binding site in SGBP-B. However, the opposing, clamp-like ar-

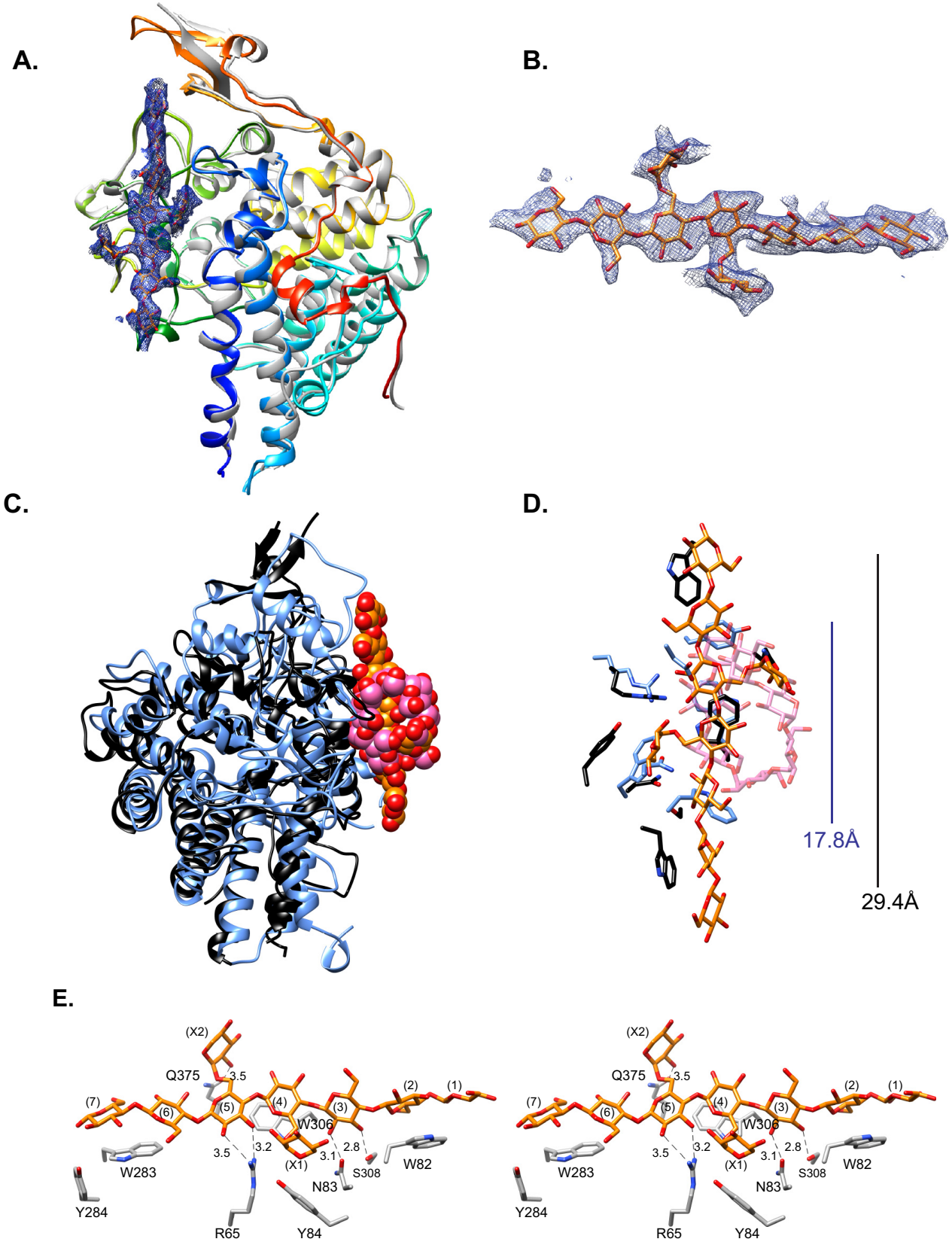


FIG 4 Molecular structure of SGBP-A (Bacova_02651). (A) Overlay of SGBP-A from the *apo* (rainbow) and $XyGO_2$ (gray) structures. The *apo* structure is color ramped from blue to red. An omit map (2σ) for XyGO2 (orange and red sticks) is displayed. (B) Close-up view of the omit map as in panel A, rotated 90° clockwise. (C) Overlay of the Cα backbones of SGBP-A (black) with $XyGO_2$ (orange and red spheres) and *Bt*SusD (blue) with maltoheptaose (pink and red

(Continued)

TABLE 3 Summary of thermodynamic parameters for site-directed mutants of SGBP-A and SGBP-B obtained by ITC with XyG at 25°C[a]

Protein name	K_a		ΔG (kcal · mol^{-1})	ΔH (kcal · mol^{-1})	$T\Delta S$ (kcal · mol^{-1})
	Fold change[b]	M^{-1}			
SGBP-A(W82A W283A W306A)	ND	NB	NB	NB	NB
SGBP-A(W82A)[c]	4.9	9.1×10^4	−6.8	−6.3	0.5
SGBP-A(W306)	ND	NB	NB	NB	NB
SGBP-B(230–489)	0.7	$(8.6 \pm 0.20) \times 10^4$	−6.7	−14.9 ± 0.1	−8.2
SGBP-B(Y363A)	19.7	$(2.9 \pm 0.10) \times 10^3$	−4.7	−18.1 ± 0.1	−13.3
SGBP-B(W364A)	ND	Weak	Weak	Weak	Weak
SGBP-B(F414A)	3.2	$(1.80 \pm 0.03) \times 10^4$	−5.8	−11.4 ± 0.1	−5.6

[a] Shown are average values ± standard deviations from two independent titrations, unless otherwise indicated. Binding thermodynamics are based on the concentration of the binding unit, $XyGO_2$. Weak binding represents a K_a of <500 M^{-1}. ND, not determined; NB, no binding.
[b] K_a fold change = K_a of wild-type protein/K_a of mutant protein for xyloglucan binding.
[c] Values from a single titration.

rangement of these residues in Bacova_04391 is clearly distinct from the planar surface arrangement of the residues that interact with XyG in SGBP-B (described below).

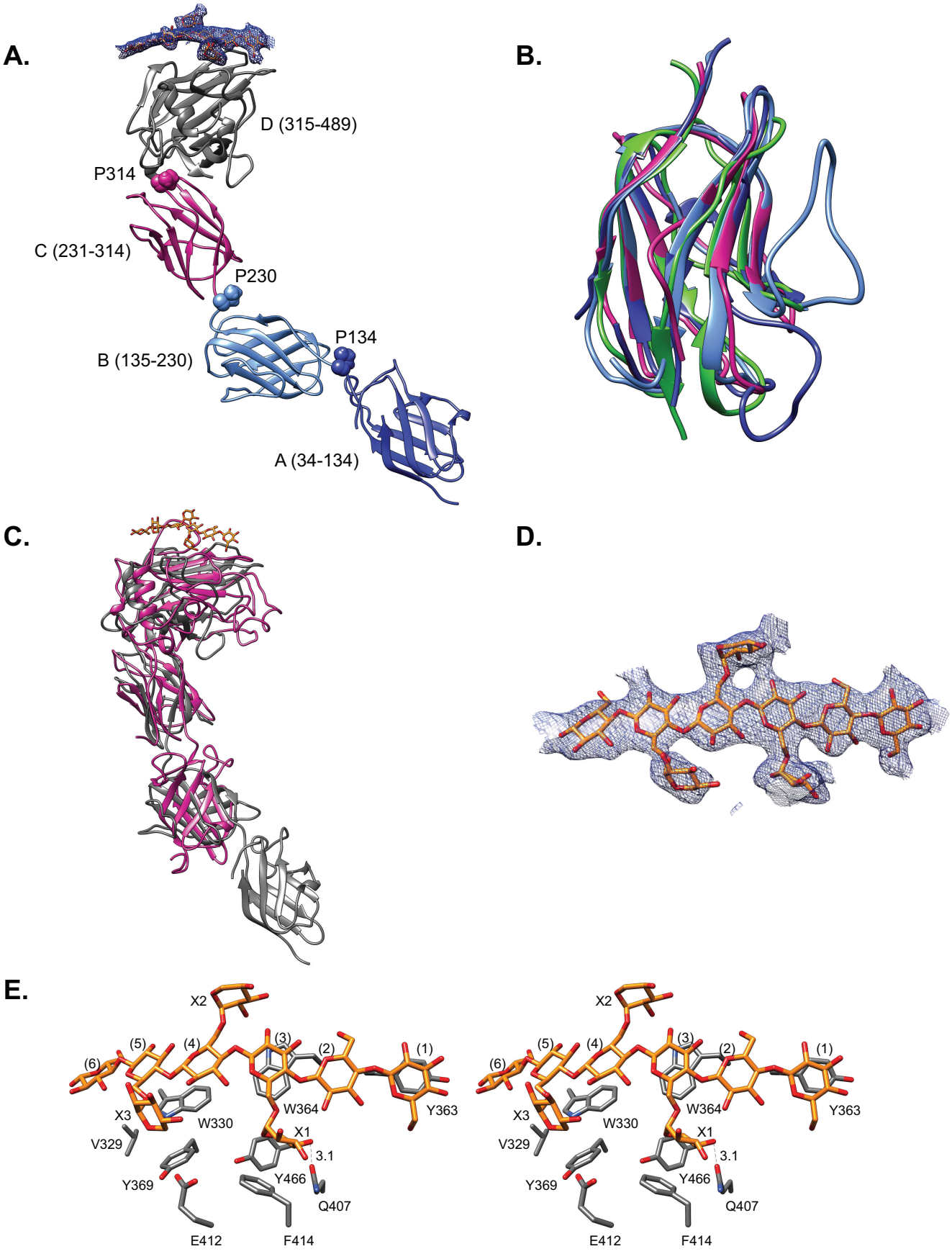
Inspection of the tertiary structure indicates that domains C and D are effectively inseparable, with a contact interface of 396 Å^2. Domains A, B, and C do not pack against each other. Moreover, the five-residue linkers between these first three domains all feature a proline as the middle residue, suggesting significant conformational rigidity (Fig. 5A). Despite the lack of sequence and structural conservation, a similarly positioned proline joins the Ig-like domains of the xylan-binding Bacova_04391 and the starch-binding proteins SusE and SusF. We speculate that this is a biologically important adaptation that serves to project the glycan binding site of these proteins far from the membrane surface. Any mobility of SGBP-B on the surface of the cell (beyond lateral diffusion within the membrane) is likely imparted by the eight-residue linker that spans the predicted lipidated Cys (C28) and the first β-strand of domain A. Other outer membrane proteins from various Sus-like systems possess a similar 10- to 20-amino-acid flexible linker between the lipidated Cys that tethers the protein to the outside the cell and the first secondary structure element (15, 16, 33). Analogously, the outer membrane-anchored endo-xyloglucanase *Bo*GH5 of the XyGUL contains a 100-amino-acid, all-β-strand, N-terminal module and flexible linker that imparts conformational flexibility and distances the catalytic module from the cell surface (23).

XyG binds to domain D of SGBP-B at the concave interface of the top β-sheet, with binding mediated by loops connecting the β-strands. Six glucosyl residues, comprising the main chain, and three branching xylosyl residues of $XyGO_2$ can be modeled in the density (Fig. 5D; cf. Fig. 1A). The backbone is flat, with less of the "twisted-ribbon" geometry observed in some cello- and xyloglucooligosaccharides (34–36). The aromatic platform created by W330, W364, and Y363 spans four glucosyl residues, compared to the longer platform of SGBP-A, which supports six glucosyl residues (Fig. 5E). The Y363A site-directed mutant of SGBP-B displays a 20-fold decrease in the K_a for XyG, while the W364A mutant lacks XyG binding (Table 3; see Fig. S6 in the supplemental material). There are no additional contacts between the protein and the β-glucan backbone and surprisingly few interactions with the side-chain xylosyl residues, despite that fact that ITC data demonstrate that SGBP-B does not measurably bind the cellohexaose (Table 1). F414 stacks with the xylosyl residue of Glc3, while Q407 is positioned for hydrogen bonding with the O4 of xylosyl residue Xyl1. Surprisingly, an F414A mutant of SGBP-B displays only a mild 3-fold decrease in the K_a value for XyG, again suggesting that glycan recognition is primarily mediated via contact with the β-glucan backbone (Table 3; see Fig. S6). Additional residues surrounding the binding site, including Y369 and E412, may contribute to the recognition of more highly decorated XyG, but precisely how this is mediated is presently unclear. Hoping to achieve a higher-resolution view of the SGBP-B–xyloglucan interaction, we solved the crystal structure of the fused CD domains in complex with $XyGO_2$ (1.57 Å, R_{work} = 15.6%, R_{free} = 17.1%, residues 230 to 489) (Table 2). The CD domains of the truncated and full-length proteins superimpose with a 0.4-Å RMSD of the Cα backbone, with no differences in the position of any of the glycan-binding residues (see Fig. S7A in the supplemental material). While density is observed for $XyGO_2$, the ligand could not be unambiguously modeled into this density to achieve a reasonable fit between the X-ray data and the known stereochemistry of the sugar (see Fig. S7B and C). While this may occur for a number of reasons in crystal structures, it is likely that the poor ligand density even at higher resolution is due to movement or multiple orientations of the sugar averaged throughout the lattice.

SGBP-A and SGBP-B have distinct, coordinated functions *in vivo*. The similarity of the glycan specificity of SGBP-A and SGBP-B presents an intriguing conundrum regarding their individual roles in XyG utilization by *B. ovatus*. To disentangle the functions of SGBP-A and SGBP-B in XyG recognition and uptake, we created individual in-frame deletion and complementation mutant strains of *B. ovatus*. In these growth experiments, overnight cultures of strains grown on minimal medium plus glucose were back-diluted 1:100-fold into minimal medium containing 5 mg/ml of the reported carbohydrate. Growth on glucose displayed the shortest lag time for each strain, and so lag times were

Figure Legend Continued

spheres), highlighting the conservation of the glycan-binding site location. (D) Close-up of the SGBP-A (black and orange) and SusD (blue and pink) glycan-binding sites. The approximate length of each glycan-binding site is displayed, colored to match the protein structures. (E) Stereo view of the xyloglucan-binding site of SGBP-A, displaying all residues within 4 Å of the ligand. The backbone glucose residues are numbered from the nonreducing end; xylose residues are labeled X1 and X2. Potential hydrogen-bonding interactions are shown as dashed lines, and the distance is shown in angstroms.


A.
D (315-489)
P314
C (231-314)
P230
P134
B (135-230)
A (34-134)
B.
C.
D.
E.
X2
(5)
(4)
(3)
(2)
(1)
(6)
X3
W330
W364
Y363
V329
X1
3.1
Y369
Y466
Q407
E412
F414

normalized for each carbohydrate by subtracting the lag time of that strain in glucose (Fig. 6; see Fig. S8 in the supplemental material). A strain in which the entire XyGUL is deleted displays a lag of 24.5 h during growth on glucose compared to the isogenic parental wild-type (WT) Δ*tdk* strain, for which exponential growth lags for 19.8 h (see Fig. S8D). It is unknown whether this is because cultures were not normalized by the starting optical density (OD) or viable cells or reflects a minor defect for glucose utilization. The former seems more likely as the growth rates are nearly identical for these strains on glucose and xylose. The ΔXyGUL and WT Δ*tdk* strains display normalized lag times on xylose within experimental error, and curiously some of the mutant and complemented strains display a nominally shorter lag time on xylose than the WT Δ*tdk* strain. Complementation of the ΔSGBP-A strain (ΔSGBP-A::SGBP-A) restores growth to wild-type rates on xyloglucan and $XyGO_1$, yet the calculated rate of the complemented strain is ~72% that of the WT Δ*tdk* strain on $XyGO_2$; similar results were obtained for the SGBP-B complemented strain despite the fact that the growth curves do not appear much different (see Fig. S8C and F). The reason for this observation on $XyGO_2$ is unclear, as the ΔSGBP-B mutant does not have a significantly different growth rate from the WT on $XyGO_2$. Therefore, we limit our discussion to those mutants that displayed the most obvious defects in growth on particular substrates.

The ΔSGBP-A (ΔBacova_02651) strain (cf. Fig. 1B) was completely incapable of growth on XyG, $XyGO_1$, and $XyGO_2$, indicating that SGBP-A is essential for XyG utilization (Fig. 6). This result mirrors our previous data for the canonical Sus of *B. thetaiotaomicron*, which revealed that a homologous Δ*susD* mutant is unable to grow on starch or malto-oligosaccharides, despite normal cell surface expression of all other PUL-encoded proteins (16, 26). More recently, we demonstrated that this phenotype is due to the loss of the physical presence of SusD; complementation of Δ*susD* with SusD*, a triple site-directed mutant (W96A W320A Y296A) that ablates glycan binding, restores *B. thetaiotaomicron* growth on malto-oligosaccharides and starch when *sus* transcription is induced by maltose addition (26). Similarly, the function of SGBP-A extends beyond glycan binding. Complementation of ΔSGBP-A with the SGBP-A* (W82A W283A W306A) variant, which does not bind XyG, supports growth on XyG and XyGOs (Fig. 6; ΔSGBP-A::SGBP-A*), with growth rates that are ~70% that of the WT. In previous studies, we observed that carbohydrate binding by SusD enhanced the sensitivity of the cells to limiting concentrations of malto-oligosaccharides by several orders of magnitude, such that the addition of 0.5 g/liter maltose was required to restore growth of the Δ*susD*::SusD* strain on starch, which nonetheless occurred following an extended lag phase (26). In contrast, the ΔSGBP-A::SGBP-A* strain does not display an extended lag time on any of the xyloglucan substrates compared to the WT (Fig. 6). The specific glycan signal that upregulates *Bo*XyGUL is currently unknown. From our present data, we cannot eliminate the possibility that the glycan binding by SGBP-A enhances transcriptional activation of the XyGUL. However, the modest rate defect displayed by the SGBP-A::SGBP-A* strain suggests that recognition of XyG and product import is somewhat less efficient in these cells.

Intriguingly, the ΔSGBP-B strain (ΔBacova_02650) (cf. Fig. 1B) exhibited a minor growth defect on both XyG and $XyGO_2$, with rates 84.6% and 93.9% that of the WT Δ*tdk* strain. However, growth of the ΔSGBP-B strain on $XyGO_1$ was 54.2% the rate of the parental strain, despite the fact that SGBP-B binds this substrate ca. 10-fold more weakly than $XyGO_2$ and XyG (Fig. 6; Table 1). As such, the data suggest that SGBP-A can compensate for the loss of function of SGBP-B on longer oligo- and polysaccharides, while SGBP-B may adapt the cell to recognize smaller oligosaccharides efficiently. Indeed, a double mutant, consisting of a crippled SGBP-A and a deletion of SGBP-B (ΔSGBP-A::SGBP-A*/ΔSGBP-B), exhibits an extended lag time on both XyG and $XyGO_2$, as well as $XyGO_1$. Taken together, the data indicate that SGBP-A and SGBP-B functionally complement each other in the capture of XyG polysaccharide, while SGBP-B may allow *B. ovatus* to scavenge smaller XyGOs liberated by other gut commensals. This additional role of SGBP-B is especially notable in the context of studies on *Bt*SusE and *Bt*SusF (positioned similarly in the archetypal Sus locus) (Fig. 1B), for which growth defects on starch or malto-oligosaccharides have never been observed. Beyond SGBP-A and SGBP-B, we speculated that the catalytically feeble endo-xyloglucanase GH9, which is expendable for growth in the presence of GH5, might also play a role in glycan binding to the cell surface (23). However, combined deletion of the genes encoding GH9 (encoded by Bacova_02649) and SGBP-B does not exacerbate the growth defect on $XyGO_1$ (Fig. 6; ΔSGBP-B/ΔGH9).

The necessity of SGBP-B is elevated in the SGBP-A* strain, as the ΔSGBP-A::SGBP-A*/ ΔSGBP-B mutant displays an extended lag during growth on XyG and xylogluco-oligosaccharides, while growth rate differences are more subtle. The precise reason for this lag is unclear, but recapitulating our findings on the role of SusD in malto-oligosaccharide sensing in *B. thetaiotaomicron*, this extended lag may be due to inefficient import and thus sensing of xyloglucan in the environment in the absence of glycan binding by essential SGBPs. Our previous work demonstrates that *B. ovatus* cells grown in minimal medium plus glucose express low levels of the XyGUL transcript (23). Thus, in our experiments, we presume that each strain, initially grown in glucose, expresses low levels of the XyGUL transcript and thus low levels of the XyGUL-encoded surface proteins, including the vanguard GH5. Presumably without glycan binding by the SGBPs, the GH5 protein cannot efficiently process xyloglucan, and/or the lack of SGBP function prevents efficient capture and import of the processed oligosaccharides. It may then be that only after a sufficient amount of glycan is processed and imported by the cell is XyGUL upregulated and exponential growth on the glycan can begin. We hypothesize that during exponential growth the essential role

FIG 5 Multimodular structure of SGBP-B (Bacova_02650). (A) Full-length structure of SGBP-B, color coded by domain as indicated. Prolines between domains are indicated as spheres. An omit map (2σ) for $XyGO_2$ is displayed to highlight the location of the glycan-binding site. (B) Overlay of SGBP-B domains A, B, and C (colored as in panel A), with a C-terminal Ig-like domain of the *G. stearothermophilus* cyclodextrin glucanotransferase (PDB 1CYG [residues 375 to 493]) in green. (C) Cα overlay of SGBP-B (gray) and Bacova_04391 (PDB 3ORJ) (pink). (D) Close-up omit map for the $XyGO_2$ ligand, contoured at 2σ. (E) Stereo view of the xyloglucan-binding site of SGBP-B, displaying all residues within 4 Å of the ligand. The backbone glucose residues are numbered from the nonreducing end, xylose residues are shown as X1, X2, and X3, potential hydrogen-bonding interactions are shown as dashed lines, and the distance is shown in angstroms.

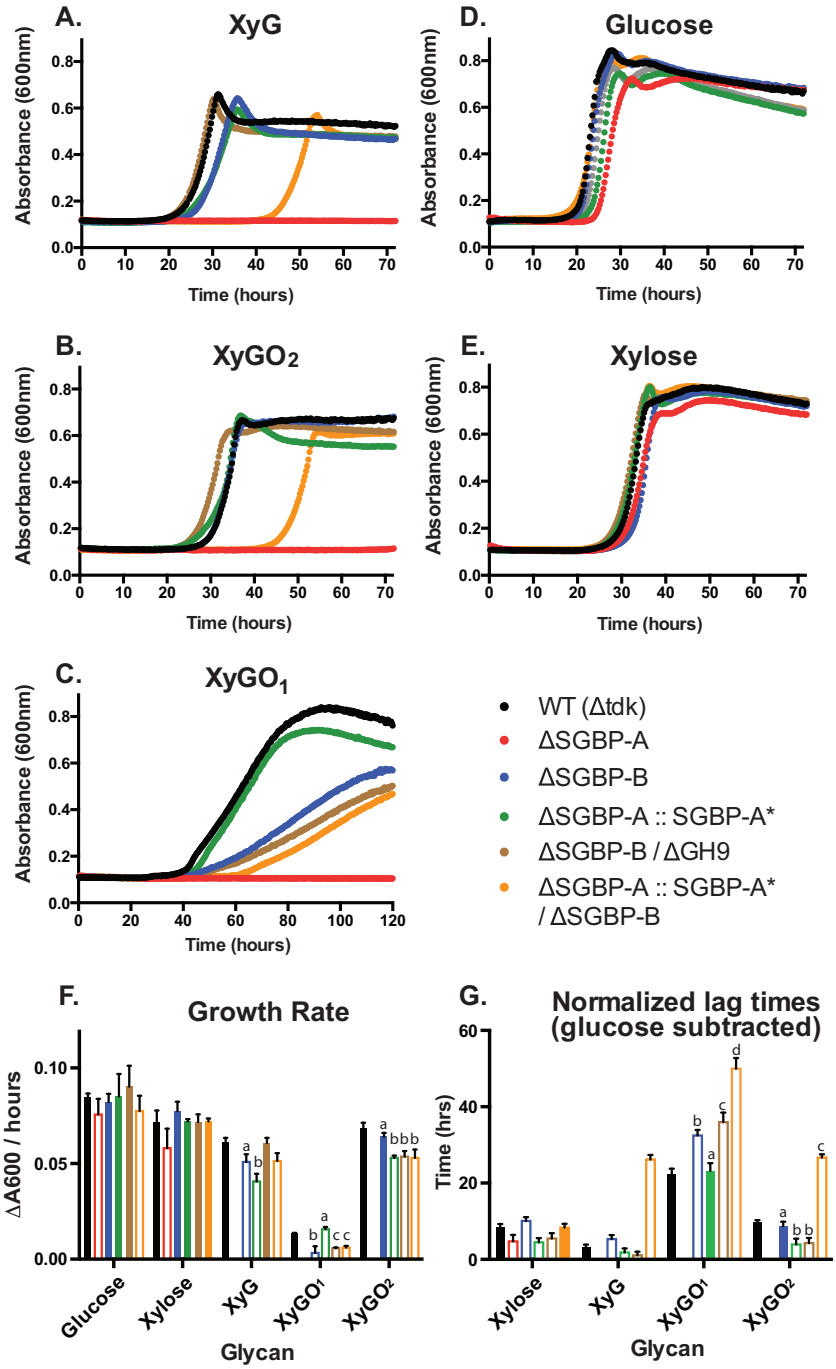


FIG 6 Growth of select XyGUL mutants on xyloglucan and oligosaccharides. *B. ovatus* mutants were created in a thymidine kinase deletion (Δ*tdk*) mutant as described previously (23). SGBP-A* denotes the Bacova_02651 (W82A W283A W306A) allele, and the GH9 gene is Bacova_02649. Growth was measured over time in minimal medium containing (A) XyG, (B) $XyGO_2$, (C) $XyGO_1$, (D) glucose, and (E) xylose. In panel F, the growth rate of each strain on the five carbon sources is displayed, and in panel G, the normalized lag time of each culture, relative to its growth on glucose, is displayed. Solid bars indicate conditions that are not statistically significant from the WT Δ*tdk* cultures grown on the indicated carbohydrate, while open bars indicate a P value of <0.005 compared to the WT Δ*tdk* strain. Conditions denoted by the same letter (b, c, or d) are not statistically significant from each other but are significantly different from the condition labeled "a." Complementation of ΔSGBP-A and ΔSBGP-B was performed by allelic exchange of the wild-type genes back into the genome for expression via the native promoter: these growth curves, quantified rates and lag times are displayed in Fig. S8 in the supplemental material.

of SGBP-A extends beyond glycan recognition, perhaps due to a critical interaction with the TBDT. In the *Bt*Sus, SusD and the TBDT SusC interact (37), and we speculate that this interaction is necessary for glycan uptake, as suggested by the fact that a Δ*susD* mutant cannot grow on starch (16, 26), but a Δ*susD*::SusD* strain regains this ability if a transcriptional activator of the *sus* operon is supplied (26). Likewise, such cognate interactions between homologous protein pairs such as SGBP-A and its TBDT may un-

derlie our observation that a ΔSGBP-A mutant cannot grow on xyloglucan. However, unlike the Sus, in which elimination of SusE and SusF does not affect growth on starch, SGBP-B appears to have a dedicated role in growth on small xyloglucooligosaccharides.

Conclusions. The ability of gut-adapted microorganisms to thrive in the gastrointestinal tract is critically dependent upon their ability to efficiently recognize, cleave, and import glycans. The human gut, in particular, is a densely packed ecosystem with hundreds of species, in which there is potential for both competition and synergy in the utilization of different substrates. Recent work has elucidated that *Bacteroidetes* cross-feed during growth on many glycans; the glycoside hydrolases expressed by one species liberate oligosaccharides for consumption by other members of the community (38, 39). Thus, understanding glycan capture at the cell surface is fundamental to explaining, and eventually predicting, how the carbohydrate content of the diet shapes the gut community structure as well as its causative health effects. Here, we demonstrate that the surface glycan binding proteins encoded within the *Bo*XyGUL play unique and essential roles in the acquisition of the ubiquitous and abundant vegetable polysaccharide xyloglucan. Yet, a number of questions remain regarding the molecular interplay of SGBPs with their cotranscribed cohort of glycoside hydrolases and TonB-dependent transporters.

A particularly understudied aspect of glycan utilization is the mechanism of import via TBDTs (SusC homologs) (Fig. 1), which are ubiquitous and defining components of all PUL (28, 40). PUL-encoded TBDTs in *Bacteroidetes* are larger than the well-characterized iron-targeting TBDTs from many *Proteobacteria* and are further distinguished as the only known glycan-importing TBDTs coexpressed with an SGBP (41). A direct interaction between the *Bt*SusC TBDT and the SusD SGBP has been previously demonstrated (13, 37), as has an interaction between the homologous components encoded by an N-glycan-scavenging PUL of *Capnocytophaga canimorsus* (42). Our observation here that the physical presence of the SusD homolog SGBP-A, independent of XyG-binding ability, is both necessary and sufficient for XyG utilization further supports a model of glycan import whereby the SusC-like TBDTs and the SusD-like SGBPs must be intimately associated to support glycan uptake (Fig. 1C). It is yet presently unclear whether this interaction is static or dynamic and to what extent the association of cognate TBDT/SGBPs is dependent upon the structure of the carbohydrate to be imported. On the other hand, there is clear evidence for independent TBDTs in *Bacteroidetes* that do not require SGBP association for activity. For example, it was recently demonstrated that expression of *nanO*, which encodes a SusC-like TBDT as part of a sialic-acid-targeting PUL from *B. fragilis*, restored growth on this monosaccharide in a mutant strain of *E. coli* (43). In this instance, coexpression of the *susD*-like gene *nanU* was not required, nor did the expression of the *nanU* gene enhance growth kinetics. Similarly, the deletion of BT1762 encoding a fructan-targeting SusD-like protein in *B. thetaiotaomicron* did not result in a dramatic loss of growth on fructans (44). Thus, the strict dependence on a SusD-like SGBP for glycan uptake in the *Bacteroidetes* may be variable and substrate dependent. Furthermore, considering the broader distribution of TBDTs in PUL lacking SGBPs (sometimes known as carbohydrate utilization containing TBDT [CUT] loci; see reference 45 and reviewed in reference 40) across bacterial phyla, it appears that the intimate biophysical association of these substrate-transport and -binding proteins is the result of specific evolution within the *Bacteroidetes*.

Equally intriguing is the observation that while SusD-like proteins such as SGBP-A share moderate primary and high tertiary structural conservation, the genes for the SGBPs encoded immediately downstream (Fig. 1B [sometimes referred to as "*susE* positioned"]) encode glycan-binding lipoproteins with little or no sequence or structural conservation, even among syntenic PUL that target the same polysaccharide. Such is the case for XyGUL from related *Bacteroides* species, which may encode either one or two of these predicted SGBPs, and these proteins vary considerably in length (23). The extremely low similarity of these SGBPs is striking in light of the moderate sequence conservation observed among homologous GHs in syntenic PUL. This, together with the observation that these SGBPs, as exemplified by *Bt*SusE and *Bt*SusF (26) and the XyGUL SGBP-B of the present study, are expendable for polysaccharide growth, implies a high degree of evolutionary flexibility to enhance glycan capture at the cell surface. Because the intestinal ecosystem is a dense consortium of bacteria that must compete for their nutrients, these multimodular SGBPs may reflect ongoing evolutionary experiments to enhance glycan uptake efficiency. Whether organisms that express longer SGBPs, extending further above the cell surface toward the extracellular environment, are better equipped to compete for available carbohydrates is presently unknown. However, the natural diversity of these proteins represents a rich source for the discovery of unique carbohydrate-binding motifs to both inform gut microbiology and generate new, specific carbohydrate analytical reagents (46).

In conclusion, the present study further illuminates the essential role that surface-glycan binding proteins play in facilitating the catabolism of complex dietary carbohydrates by *Bacteroidetes*. The ability of our resident gut bacteria to recognize polysaccharides is the first committed step of glycan consumption by these organisms, a critical process that influences the community structure and thus the metabolic output (i.e., short-chain fatty acid and metabolite profile) of these organisms. A molecular understanding of glycan uptake by human gut bacteria is therefore central to the development of strategies to improve human health through manipulation of the microbiota (47, 48).

MATERIALS AND METHODS

Protein production and purification. The gene fragments corresponding to Bacova_02650 (encoding SGBP-B residues 34 to 489) and Bacova_02651 (encoding SGBP-A residues 28 to 546) were amplified from *Bacteroidetes ovatus* ATCC 8483 genomic DNA by PCR using forward primers, including NdeI restriction sites, and reverse primers, including XhoI. The gene products were ligated into a modified version of pET-28a (EMD Biosciences) containing a recombinant tobacco etch virus (rTEV) protease recognition site (pET-28aTEV) preceding an N-terminal 6-His tag for affinity purification. The expression vector (pET-28TEV) containing SGBP-B was used for subsequent cloning of the domains A (residues 34 to 133), B (residues 134 to 229), and CD (residues 230 to 489). The pET-28TEV vector expressing residues 28 to 546 of SGBP-A was utilized for carbohydrate-binding experiments and crystallization of the *apo* structure. To obtain crystals of SGBP-A with $XyGO_2$, the DNA sequence coding for residues 36 to 546 was PCR amplified from genomic DNA for ligation-independent cloning into the pETite N-His vector (Lucigen, Madison, WI) according to the manufacturer's instructions. The N-terminal primer for pETite N-His insertion contained a TEV cleavage site immediately downstream of the complementary 18-bp overlap (encoding the His tag) to create a TEV-cleavable His-tagged protein. The

site-directed mutants of SGBP-A and SGBP-B in pET-28TEV were created using the QuikChange II site-directed mutagenesis kit (Stratagene) according to the manufacturer's instructions. The sequences of all primers to generate these constructs are displayed in Table S1 in the supplemental material.

The plasmids containing the SGBP-A and SGBP-B genes were transformed into *Escherichia coli* BL21(DE3) or Rosetta(DE3) cells. For native protein expression, cells were cultured in Terrific Broth containing kanamycin (50 μg/ml) and chloramphenicol (20 μg/ml) at 37°C to the mid-exponential phase (A_{600} of ≈0.6), induced by the addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and then incubated for 2 days at 16°C or 1 day at 20°C. Cells were harvested by centrifugation and frozen at −80°C prior to protein purification. For selenomethionine-substituted SGBP-B, the pET-28TEV-SGBP-B plasmid was transformed into *E. coli* Rosetta(DE3)/pLysS and plated onto LB supplemented with kanamycin (50 μg/ml) and chloramphenicol (20 μg/ml). After 16 h of growth at 37°C, colonies were harvested from the plates, used to inoculate 100 ml of M9 minimal medium supplemented with kanamycin (30 μg/ml) and chloramphenicol (20 μg/ml), and then grown at 37°C for 16 h. This overnight culture was used to inoculate a 2-liter baffled flask containing 1 liter of Molecular Dimensions SelenoMet premade medium supplemented with 50 ml of the recommended sterile nutrient mix, chloramphenicol, and kanamycin. Cultures were grown at 37°C to an A_{600} of ≈0.45 before adjusting the temperature to 20°C and supplementing each flask with 100 mg each of L-lysine, L-threonine, and L-phenylalanine and 50 mg each of L-leucine, L-isoleucine, L-valine, and L-selenomethionine (49). After 20 additional minutes of growth, the cells were induced with 0.5 mM IPTG, and cultures were grown for an additional 48 h.

For the purification of native and selenomethionine-substituted protein, cells were thawed and lysed via sonication in His buffer (25 mM NaH_2PO_4, 500 mM NaCl, 20 mM imidazole, pH 7.5) and purified via immobilized nickel affinity chromatography (His-Trap; GE Healthcare) using a gradient of 20 to 300 mM imidazole, according to the manufacturer's instructions. The His tag was removed by incubation with TEV protease (1:100 molar ratio relative to protein) at room temperature for 2 h and then overnight at 4°C while being dialyzed against His buffer. The cleaved protein was then repurified via nickel affinity chromatography to remove undigested target protein, the cleaved His tag, and His-tagged TEV protease. Purified proteins were dialyzed against 20 mM HEPES–100 mM NaCl (pH 7.0) prior to crystallization and concentrated using Vivaspin 15 (10,000-molecular-weight-cutoff [MWCO]) centrifugal concentrators (Vivaproducts, Inc.).

Glycans. Xyloglucan from tamarind seed, β-glucan from barley, and konjac glucomannan were purchased from Megazyme. Starch, guar, and mucin were purchased from Sigma. Hydroxyethyl cellulose was purchased from AMRESCO. Carboxymethyl cellulose was purchased from Acros Organics. Xylogluco-oligosaccharides $XyGO_1$ and $XyGO_2$ for biophysical analyses were prepared from tamarind seed XyG according to the method of Martinez-Fleites et al. (61) with minor modifications. $XyGO_2$ for cocrystallization was purchased from Megazyme (O-XGHDP).

Affinity gel electrophoresis. Affinity PAGE was performed as described previously (50), with minor modification. Various polysaccharides were used at a concentration of 0.05 to 0.1% (wt/vol), and electrophoresis was carried out for 90 min at room temperature in native 10% (wt/vol) polyacrylamide gels. BSA was used as noninteracting negative-control protein.

ITC. Isothermal titration calorimetry (ITC) of glycan binding by the SGPB-A mutants was performed using the TA Nano isothermal titration calorimeter calibrated to 25°C. Proteins were dialyzed against 20 mM HEPES–100 mM NaCl (pH 7.0), and sugars were prepared using the dialysis buffer. The protein (45 to 50 μM) was placed in the sample cell, and the syringe was loaded with 2.5 to 4 mg/ml XyG polysaccharide. Following an initial injection of 0.5 μl, 26 subsequent injections of 2 μl were performed with stirring at 350 rpm, and the resulting heat of reaction was recorded. Data were analyzed using the Nano Analyze software. All other ITC experiments were performed using a MicroCal VP-ITC titration calorimeter calibrated to 25°C. Proteins were dialyzed into 20 mM HEPES–100 mM NaCl (pH 7.0), and polysaccharides were prepared using the dialysis buffer. Proteins (micromolar concentrations) were placed in the sample cell, and a first injection of 2 μl was performed followed by 24 subsequent injections of 10 μl of 2 to 20 mM oligosaccharide (cellotetraose, cellohexaose, $XyGO_1$, or $XyGO_2$) or 1 to 2.5 mg/ml XyG polysaccharide. The solution was stirred at 280 rpm, and the resulting heat of reaction was recorded. Data were analyzed using the Origin software program.

DSC. Structural integrity of the SGBP-B mutants was verified by differential scanning calorimetry (DSC). DSC studies were performed on a MicroCal VP-DSC (Malvern Instruments). Experiments were carried out in 50 mM HEPES (pH 7.0) at a scan rate of 60°C/h. All samples (40 μM protein) were degassed for 7 min with gentle stirring under vacuum prior to being loaded into the calorimeter. Background excess thermal power scans were obtained with buffer in both the sample and reference cells and subtracted from the scans for each sample solution to generate excess heat capacity versus temperature thermograms.

The melting temperature decreased from 57.8 ± 0.9°C for the wild-type SGBP-B protein to 54.6 ± 0.1°C for the Y363A mutant, 54.2 ± 0.1°C for the W364A mutant, and 52 ± 1°C for the F414A mutant. All proteins were therefore in their stable folded state for the ITC measurements (see Fig. S9 in the supplemental material).

***Bacteroides ovatus* mutagenesis.** Gene deletions and complementations were performed via allelic exchange in a *Bacteroides ovatus* thymidine kinase gene (Bacova_03071) deletion (Δ*tdk*) derivative strain of ATCC 8483 using the vector pExchange-tdk, as previously described (23). Primers for the construction of *B. ovatus* mutants are listed in Table S1. The *B. ovatus* Δtdk strain and the *B. ovatus* ΔXyGUL mutant were a generous gift from Eric Martens, University of Michigan Medical School.

***Bacteroides* growth experiments.** All *Bacteroides ovatus* culturing was performed in a Coy anaerobic chamber (85% N_2, 10% H_2, 5% CO_2) at 37°C. Prior to growth on minimal medium plus the carbohydrates indicated (Fig. 6; see Fig. S8 in the supplemental material), each strain was grown for 16 h from a freezer stock in tryptone-yeast extract-glucose (TYG) medium (51) and then back diluted 1:100 into *Bacteroides* minimal medium supplemented with 5 mg/ml glucose, as previously described (52). After growth for 24 h, cultures were back-diluted 1:100 into *Bacteroides* minimal medium supplemented with 5 mg/ml of glucose, xylose, XyG, $XyGO_1$, or $XyGO_2$. Growth experiments were performed in replicates of 12 (glucose, xylose, and xyloglucan) or 5 ($XyGO_1$ and $XyGO_2$) as 200-μl cultures in 96-well plates. Plates were loaded into a Biostack automated plate handling device coupled to a Powerwave HT absorbance reader (both devices from Biotek Instruments, Winooski, VT). Absorbance at 600 nm (A_{600}; i.e., optical density at 600 nm [OD_{600}]) was measured for each well at 20-min intervals. Data were processed using Gen5 software (BioTek) and Microsoft Excel. Growth was quantified in each assay by first identifying a minimum time point (A_{min}) at which A_{600} had increased by 15% over a baseline reading taken during the first 500 min of incubation. Next, we identified the time point at which A_{600} reached its maximum (A_{max}) immediately after exponential growth. The growth rate for each well was defined by $(A_{max} - A_{min})/(T_{max} - T_{min})$, where T_{max} and T_{min} are the corresponding time values for each absorbance. To account for variations in inoculum density, for each strain, the lag time (T_{min}) on glucose was subtracted from the lag time for the substrate of interest; in all cases, cultures had shorter lag times on glucose than other glycans.

Immunofluorescence. Custom rabbit antibodies to recombinant SGBP-A and SGBP-B were generated by Cocalico Biologicals, Inc. (Reamstown, PA). The *B. ovatus* ATCC 8483 Δ*tdk* and ΔSGBP-B strains were grown in 1 ml minimal *Bacteroides* medium (11) supplemented with 5 mg/ml tamarind xyloglucan to an A_{600} of ≈0.6 and then harvested via centrifugation (7,000 × *g* for 3 min) and washed twice with phosphate-buffered saline (PBS). Cells were resuspended in 0.25 ml PBS, and 0.75 ml

of 6% formalin in PBS was added. Cells were incubated with rocking at 20°C for 1.5 h and then washed twice with PBS. Cells were resuspended in 0.5 to 1 ml blocking solution (2% goat serum, 0.02% NaN_3 in PBS) and incubated for 16 h at 4°C. Cells were centrifuged and resuspended in 0.5 ml of a 1/100 dilution of custom rabbit antibody sera in blocking solution and incubated by rocking for 2 h at 20°C. Cells were washed with PBS and then resuspended in 0.4 ml of a 1/500 dilution of Alexa Fluor 488 goat anti-rabbit IgG (Life Technologies) in blocking solution and incubated with rocking for 1 h at 20°C. Cells were washed three times with an excess of PBS and then resuspended in 20 μl of PBS plus 1 μl of ProLong Gold antifade (Life Technologies). Cells were spotted on 0.8% agarose pads and imaged at the Center for Live Cell Imaging at the University of Michigan Medical School, using an Olympus IX70 inverted confocal microscope. Images were processed with Metamorph Software.

Crystallization and data collection. All X-ray diffraction data for both native and selenomethionine-substituted protein crystals were collected at the Life Science Consortium (LS-CAT) at the Advance Photon Source at Argonne National Laboratory in Argonne, IL. The native protein SGBP-B (residues 34 to 489) was concentrated to an A_{280} of 12.25 prior to crystallization and mixed with 10 mM $XyGO_2$ (Megazyme, O-XGHDP). Hanging drop vapor diffusion was performed against mother liquor consisting of 1.1 to 1.5 M ammonium sulfate and 30 to 70 mM sodium cacodylate (pH 6.5). To decrease crystal nucleation, 0.3 ml of paraffin oil was overlaid on top of 0.5 ml of mother liquor yielding diffraction-quality crystals within 2 weeks. Selenomethionine-substituted crystals of SGBP-B were generated using the same conditions as the native crystals. Crystals of the truncated SGBP-B (domains CD, residues 230 to 489) were obtained by mixing concentrated protein (A_{600} of 20.6) with 10 mM $XyGO_2$ for hanging drop vapor diffusion against a solution containing 2 M sodium formate and 0.1 M sodium acetate (pH 4.6). All SGBP-B crystals were flash-frozen prior to data collection by briefly soaking in a solution of 80% mother liquor–20% glycerol plus 10 mM xylogluco-oligosaccharide. Data were processed and scaled using HKL2000 and Scalepack (53). SAD phasing from a selenomethionine-substituted protein crystals was used to determine the structure of SGBP-B. The AutoSol (54) and Autobuild (55) algorithms within the Phenix (56) suite of programs were used to locate and refine the selenium positions and automatically build an initial model of the protein structure, respectively. Successive rounds of manual model building and refinement in Coot (57) and Phenix, respectively, were utilized to build a 2.7-Å model of the selenomethionine-substituted protein, which then was placed in the unit cell of the native data set. Additional rounds of manual model building and refinement were performed to complete the 2.37-Å structure of SGBP-B with $XyGO_2$. The structure of the truncated protein (CD domains, residues 230 to 489) was solved via molecular replacement with Phaser (58) using the CD domains of the full-length protein as a model.

The native protein SGBP-A (residues 28 to 546) was concentrated to an A_{280} of 28.6 and crystallized via hanging drop vapor diffusion from the Morpheus crystal screen (Molecular Dimensions). Crystals formed in well A1 (30 mM $MgCl_2$, 30 mM $CaCl_2$, 20% polyethylene glycol [PEG 500], 10% PEG 20K, 0.1 M imidazole-MES [morpholinethanesulfonic acid], pH 6.5), and were flash-frozen in liquid nitrogen without additional cryoprotectant. The truncated SGBP-A (residues 36 to 546) concentrated to an A_{280} of 38.2 yielded crystals with 10 mM $XyGO_2$ via hanging drop vapor diffusion against 1.2 to 1.8 M sodium citrate (pH 6.15 to 6.25), and were flash-frozen in a cryoprotectant solution of 80% mother liquor–20% ethylene glycol with the glycan. Data were processed and scaled using HKL2000 and Scalepack (53). The structure of the *apo* protein was solved via molecular replacement with BALBES (59) using the homologous structure PDB 3JYS, followed by successive rounds of automatic and manual model building with Autobuild and Coot. The structure of SGBP-A with $XyGO_2$ was solved via molecular replacement with Phaser (58) and refined with Phenix (56). X-data collection and refinement statistics are presented in Table 2.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.02134-15/-/DCSupplemental.

Figure S1, PDF file, 0.6 MB.
Figure S2, PDF file, 0.1 MB.
Figure S3, PDF file, 0.1 MB.
Figure S4, PDF file, 0.1 MB.
Figure S5, PDF file, 0.1 MB.
Figure S6, PDF file, 0.1 MB.
Figure S7, PDF file, 1.7 MB.
Figure S8, PDF file, 0.2 MB.
Figure S9, PDF file, 0.1 MB.
Table S1, PDF file, 0.1 MB.

ACKNOWLEDGMENTS

We thank Eric C. Martens and his laboratory at the University of Michigan Medical School for the use of their anaerobic chamber and plate reader for kinetic growth assays, as well as for the providing the *B. ovatus* Δtdk strain and *B. ovatus* ΔXyGUL strain for genetic manipulation and growth experiments, respectively. We thank Eric C. Martens and Darrell W. Cockburn at the University of Michigan Medical School for helpful discussions on experimental design and data interpretation.

Work in Vancouver was initially supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada and UBC Faculty funding, followed by Canadian Institutes for Health Research Operating Grants (FRN: MOP-137134 and MOP-142472). Infrastructure support was provided by the Canadian Foundation for Innovation and the British Columbia Knowledge Development Fund. Work in Michigan was supported by the University of Michigan Medical School Host Microbiome Initiative. This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract no. DE-AC02-06CH11357. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (grant 085P1000817). None of the funders had any role in study design, data collection and interpretation, or the decision to submit the work for publication.

FUNDING INFORMATION

This work, including the efforts of Nicole Marie Koropatkin, was funded by University of Michigan Medical School Host Microbiome Initiative. This work, including the efforts of Zdzislaw Wawrzak, was funded by United States Department of Energy. This work, including the efforts of Zdzislaw Wawrzak, was funded by Michigan Technology Tri-Corridor. This work, including the efforts of Zdzislaw Wawrzak, was funded by Michigan Economic Development Corporation (MEDC). This work, including the efforts of Harry Brumer, was funded by Gouvernement du Canada | Natural Sciences and Engineering Research Council of Canada (NSERC). This work, including the efforts of Harry Brumer, was funded by Gouvernement du Canada | Canadian Institutes of Health Research (CIHR).

REFERENCES

1. **Mazmanian SK, Liu CH, Tzianabos AO, Kasper DL.** 2005. An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. Cell **122:**107–118. http://dx.doi.org/10.1016/j.cell.2005.05.007.
2. **Sommer F, Nookaew I, Sommer N, Fogelstrand P, Bäckhed F.** 2015. Site-specific programming of the host epithelial transcriptome by the gut microbiota. Genome Biol **16:**62. http://dx.doi.org/10.1186/s13059-015-0614-4.
3. **Ridaura VK, Faith JJ, Rey FE, Cheng J, Duncan AE, Kau AL, Griffin NW, Lombard V, Henrissat B, Bain JR, Muehlbauer MJ, Ilkayeva O, Semenkovich CF, Funai K, Hayashi DK, Lyle BJ, Martini MC, Ursell LK, Clemente JC, Van Treuren W.** 2013. Gut microbiota from twins

discordant for obesity modulate metabolism in mice. Science **341:** 1241214. http://dx.doi.org/10.1126/science.1241214
4. **Bäckhed F, Manchester JK, Semenkovich CF, Gordon JI.** 2007. Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. Proc Natl Acad Sci U S A **104:**979–984. http://dx.doi.org/10.1073/pnas.0605374104.
5. **Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, Relman DA.** 2005. Diversity of the human intestinal microbial flora. Science **308:**1635–1638. http://dx.doi.org/10.1126/science.1110591.
6. **Ding T, Schloss PD.** 2014. Dynamics and associations of microbial community types across the human body. Nature **509:**357–360. http://dx.doi.org/10.1038/nature13178.
7. **El Kaoutari A, Armougom F, Gordon JI, Raoult D, Henrissat B.** 2013. The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. Nat Rev Microbiol **11:**497–504. http://dx.doi.org/10.1038/nrmicro3050.
8. **Xu J, Mahowald MA, Ley RE, Lozupone CA, Hamady M, Martens EC, Henrissat B, Coutinho PM, Minx P, Latreille P, Cordum H, Van Brunt A, Kim K, Fulton RS, Fulton LA, Clifton SW, Wilson RK, Knight RD, Gordon JI.** 2007. Evolution of symbiotic bacteria in the distal human intestine. PLoS Biol **5:**e156. http://dx.doi.org/10.1371/journal.pbio.0050156.
9. **Xu J, Bjursell MK, Himrod J, Deng S, Carmichael LK, Chiang HC, Hooper LV, Gordon JI.** 2003. A genomic view of the human-Bacteroides thetaiotaomicron symbiosis. Science **299:**2074–2076. http://dx.doi.org/10.1126/science.1080029.
10. **Martens EC, Kelly AG, Tauzin AS, Brumer H.** 2014. The devil lies in the details: how variations in polysaccharide fine-structure impact the physiology and evolution of gut microbes. J Mol Biol **426:**3851–3865. http://dx.doi.org/10.1016/j.jmb.2014.06.022.
11. **Martens EC, Lowe EC, Chiang H, Pudlo NA, Wu M, McNulty NP, Abbott DW, Henrissat B, Gilbert HJ, Bolam DN, Gordon JI.** 2011. Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. PLoS Biol **9:**e1001221. http://dx.doi.org/10.1371/journal.pbio.1001221.
12. **Tancula E, Feldhaus MJ, Bedzyk LA, Salyers AA.** 1992. Location and characterization of genes involved in binding of starch to the surface of Bacteroides thetaiotaomicron. J Bacteriol **174:**5609–5616.
13. **Shipman JA, Berleman JE, Salyers AA.** 2000. Characterization of four outer membrane proteins involved in binding starch to the cell surface of Bacteroides thetaiotaomicron. J Bacteriol **182:**5365–5372. http://dx.doi.org/10.1128/JB.182.19.5365-5372.2000.
14. **Shipman JA, Cho KH, Siegel HA, Salyers AA.** 1999. Physiological characterization of SusG, an outer membrane protein essential for starch utilization by Bacteroides thetaiotaomicron. J Bacteriol **181:**7206–7211.
15. **Koropatkin NM, Smith TJ.** 2010. SusG: A unique cell-membrane-associated alpha-amylase from a prominent human gut symbiont targets complex starch molecules. Structure **18:**200–215. http://dx.doi.org/10.1016/j.str.2009.12.010.
16. **Koropatkin NM, Martens EC, Gordon JI, Smith TJ.** 2008. Starch catabolism by a prominent human gut symbiont is directed by the recognition of amylose helices. Structure **16:**1105–1115. http://dx.doi.org/10.1016/j.str.2008.03.017.
17. **Reeves AR, D'Elia JN, Frias J, Salyers AA.** 1996. A Bacteroides thetaiotaomicron outer membrane protein that is essential for utilization of maltooligosaccharides and starch. J Bacteriol **178:**823–830.
18. **Cameron EA, Maynard MA, Smith CJ, Smith TJ, Koropatkin NM, Martens EC.** 2012. Multidomain carbohydrate-binding proteins involved in Bacteroides thetaiotaomicron starch metabolism. J Biol Chem **287:** 34614–34625. http://dx.doi.org/10.1074/jbc.M112.397380.
19. **D'Elia JN, Salyers AA.** 1996. Contribution of a neopullulanase, a pullulanase, and an alpha-glucosidase to growth of Bacteroides thetaiotaomicron on starch. J Bacteriol **178:**7173–7179.
20. **Hehemann JH, Kelly AG, Pudlo NA, Martens EC, Boraston AB.** 2012. Bacteria of the human gut microbiome catabolize red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes. Proc Natl Acad Sci U S A **109:**19786–19791. http://dx.doi.org/10.1073/pnas.1211002109.
21. **Cuskin F, Lowe EC, Temple MJ, Zhu Y, Cameron EA, Pudlo NA, Porter NT, Urs K, Thompson AJ, Cartmell A, Rogowski A, Hamilton BS, Chen R, Tolbert TJ, Piens K, Bracke D, Vervecken W, Hakki Z, Speciale G, Munōz-Munōz JL, Day A, Pena MJ, McLean R, Suits MD, Boraston AB, Atherly T, Ziemer CJ, Williams SJ, Davies GJ, Abbott DW, Martens EC, Gilbert HJ.** 2015. Human gut Bacteroidetes can utilize yeast mannan through a selfish mechanism. Nature **517:**165–169. http://dx.doi.org/10.1038/nature13995.
22. **Rogowski A, Briggs JA, Mortimer JC, Tryfona T, Terrapon N, Lowe EC, Basle A, Morland C, Day AM, Zheng H, Rogers TE, Thompson P, Hawkins AR, Yadav MP, Henrissat B, Martens EC, Dupree P, Gilbert HJ, Bolam DN.** 2015. Glycan complexity dictates microbial resource allocation in the large intestine. Nat Commun **6:**7481. http://dx.doi.org/10.1038/ncomms8481.
23. **Larsbrink J, Rogers TE, Hemsworth GR, McKee LS, Tauzin AS, Spadiut O, Klinter S, Pudlo NA, Urs K, Koropatkin NM, Creagh AL, Haynes CA, Kelly AG, Cederholm SN, Davies GJ, Martens EC, Brumer H.** 2014. A discrete genetic locus confers xyloglucan metabolism in select human gut Bacteroidetes. Nature **506:**498–502. http://dx.doi.org/10.1038/nature12907.
24. **McDougall GJ, Morrison IM, Stewart D, Hillman JR.** 1996. Plant cell walls as dietary fibre: range, structure, processing and function. J Sci Food Agric **70:**133–150.
25. **Schultink A, Liu L, Zhu L, Pauly M.** 2014. Structural diversity and function of xyloglucan sidechain substituents. Plants **3:**526–542. http://dx.doi.org/10.3390/plants3040526.
26. **Cameron EA, Kwiatkowski KJ, Lee BH, Hamaker BR, Koropatkin NM, Martens EC.** 2014. Multifunctional nutrient-binding proteins adapt human symbiotic bacteria for glycan competition in the gut by separately promoting enhanced sensing and catalysis. mBio **5:**e01441-01414. http://dx.doi.org/10.1128/mBio.01441-14.
27. **Terrapon N, Lombard V, Gilbert HJ, Henrissat B.** 2015. Automatic prediction of polysaccharide utilization loci in Bacteroidetes species. Bioinformatics **31:**647–655. http://dx.doi.org/10.1093/bioinformatics/btu716.
28. **Martens EC, Koropatkin NM, Smith TJ, Gordon JI.** 2009. Complex glycan catabolism by the human gut microbiota: the Bacteroidetes Sus-like paradigm. J Biol Chem **284:**24673–24677. http://dx.doi.org/10.1074/jbc.R109.022848.
29. **Bolam DN, Sonnenburg JL.** 2011. Mechanistic insight into polysaccharide use within the intestinal microbiota. Gut Microbes **2:**86–90. http://dx.doi.org/10.4161/gmic.2.2.15232.
30. **Ellrott K, Jaroszewski L, Li W, Wooley JC, Godzik A.** 2010. Expansion of the protein repertoire in newly explored environments: human gut microbiome specific protein families. PLoS Comput Biol **6:**e1000798. http://dx.doi.org/10.1371/journal.pcbi.1000798.
31. **Bolam DN, Koropatkin NM.** 2012. Glycan recognition by the Bacteroidetes Sus-like systems. Curr Opin Struct Biol **22:**563–569. http://dx.doi.org/10.1016/j.sbi.2012.06.006.
32. **Zhou Q, Rutland MW, Teeri TT, Brumer H.** 2007. Xyloglucan in cellulose modification. Cellulose **14:**625–641. http://dx.doi.org/10.1007/s10570-007-9109-0.
33. **Koropatkin N, Martens EC, Gordon JI, Smith TJ.** 2009. Structure of a SusD homologue, BT1043, involved in mucin O-glycan utilization in a prominent human gut symbiont. Biochemistry **48:**1532–1542. http://dx.doi.org/10.1021/bi801942a.
34. **von Schantz L, Håkansson M, Logan DT, Nordberg-Karlsson E, Ohlin M.** 2014. Carbohydrate binding module recognition of xyloglucan defined by polar contacts with branching xyloses and CH-Pi interactions. Proteins **82:**3466–3475. http://dx.doi.org/10.1002/prot.24700.
35. **Luís AS, Venditto I, Temple MJ, Rogowski A, Baslé A, Xue J, Knox JP, Prates JA, Ferreira LM, Fontes CM, Najmudin S, Gilbert HJ.** 2013. Understanding how noncatalytic carbohydrate binding modules can display specificity for xyloglucan. J Biol Chem **288:**4799–4809. http://dx.doi.org/10.1074/jbc.M112.432781.
36. **Tsukimoto K, Takada R, Araki Y, Suzuki K, Karita S, Wakagi T, Shoun H, Watanabe T, Fushinobu S.** 2010. Recognition of cellooligosaccharides by a family 28 carbohydrate-binding module. FEBS Lett **584:**1205–1211. http://dx.doi.org/10.1016/j.febslet.2010.02.027.
37. **Cho KH, Salyers AA.** 2001. Biochemical analysis of interactions between outer membrane proteins that contribute to starch utilization by Bacteroides thetaiotaomicron. J Bacteriol **183:**7224–7230. http://dx.doi.org/10.1128/JB.183.24.7224-7230.2001.
38. **Rakoff-Nahoum S, Coyne MJ, Comstock LE.** 2014. An ecological network of polysaccharide utilization among human intestinal symbionts. Curr Biol **24:**40–49. http://dx.doi.org/10.1016/j.cub.2013.10.077.
39. **Elhenawy W, Debelyy MO, Feldman MF.** 2014. Preferential packing of

acidic glycosidases and proteases into Bacteroides outer membrane vesicles. mBio **5:**e00909-14. http://dx.doi.org/10.1128/mBio.00909-14.
40. **Hemsworth GR, Déjean G, Davies GJ, Brumer H**. 2016. Learning from microbial strategies for polysaccharide degradation. Biochem Soc Trans **44:**94–108. http://dx.doi.org/10.1042/BST20150180.
41. **Noinaj N, Guillier M, Barnard TJ, Buchanan SK**. 2010. TonB-dependent transporters: regulation, structure, and function. Annu Rev Microbiol **64:** 43–60. http://dx.doi.org/10.1146/annurev.micro.112408.134247.
42. **Renzi F, Manfredi P, Mally M, Moes S, enö P, Cornelis GR**. 2011. The N-glycan glycoprotein deglycosylation complex (Gpd) from Capnocytophaga canimorsus deglycosylates human IgG. PLoS Pathog **7:**e1002118. http://dx.doi.org/10.1371/journal.ppat.1002118.
43. **Phansopa C, Roy S, Rafferty JB, Douglas CW, Pandhal J, Wright PC, Kelly DJ, Stafford GP**. 2014. Structural and functional characterization of NanU, a novel high-affinity sialic acid-inducible binding protein of oral and gut-dwelling Bacteroidetes species. Biochem J **458:**499–511. http://dx.doi.org/10.1042/BJ20131415.
44. **Sonnenburg ED, Zheng H, Joglekar P, Higginbottom SK, Firbank SJ, Bolam DN, Sonnenburg JL**. 2010. Specificity of polysaccharide use in intestinal Bacteroides species determines diet-induced microbiota alterations. Cell **141:**1241–1252. http://dx.doi.org/10.1016/j.cell.2010.05.005.
45. **Blanvillain S, Meyer D, Boulanger A, Lautier M, Guynet C, Denancé N, Vasse J, Lauber E, Arlat M**. 2007. Plant carbohydrate scavenging through tonB-dependent receptors: a feature shared by phytopathogenic and aquatic bacteria. PLoS One **2:**e224. http://dx.doi.org/10.1371/journal.pone.0000224.
46. **Gilbert HJ, Knox JP, Boraston AB**. 2013. Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. Curr Opin Struct Biol **23:**669–677. http://dx.doi.org/10.1016/j.sbi.2013.05.005.
47. **Gordon JI**. 2012. Honor thy gut symbionts redux. Science **336:**1251–1253. http://dx.doi.org/10.1126/science.1224686.
48. **Hutkins RW, Krumbeck JA, Bindels LB, Cani PD, Fahey G, Jr, Goh YJ, Hamaker B, Martens EC, Mills DA, Rastal RA, Vaughan E, Sanders ME**. 2016. Prebiotics: why definitions matter. Curr Opin Biotechnol **37:**1–7. http://dx.doi.org/10.1016/j.copbio.2015.09.001.
49. **Van Duyne GD, Standaert RF, Karplus PA, Schreiber SL, Clardy J**. 1993. Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. J Mol Biol **229:**105–124. http://dx.doi.org/10.1006/jmbi.1993.1012.
50. **Freelove AC, Bolam DN, White P, Hazlewood GP, Gilbert HJ**. 2001. A novel carbohydrate-binding protein is a component of the plant cell wall-degrading complex of Piromyces equi. J Biol Chem **276:**43010–43017. http://dx.doi.org/10.1074/jbc.M107143200.
51. **Holdeman LV, Cato ED, Moore WEC**. 1977. Anaerobe laboratory manual. Virginia Polytechnic Institute and State University Anaerobe Laboratory, Blacksburg, VA.
52. **Martens EC, Chiang HC, Gordon JI**. 2008. Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. Cell Host Microbe **4:**447–457. http://dx.doi.org/10.1016/j.chom.2008.09.007.
53. **Otwinowski Z, Minor W**. 1997. Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol **276:**307–326.
54. **Terwilliger TC, Adams PD, Read RJ, McCoy AJ, Moriarty NW, Grosse-Kunstleve RW, Afonine PV, Zwart PH, Hung LW**. 2009. Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. Acta Crystallogr D Biol Crystallogr **65:**582–601. http://dx.doi.org/10.1107/S0907444909012098.
55. **Terwilliger TC, Grosse-Kunstleve RW, Afonine PV, Moriarty NW, Zwart PH, Hung LW, Read RJ, Adams PD**. 2008. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. Acta Crystallogr D Biol Crystallogr **64:**61–69. http://dx.doi.org/10.1107/S090744490705024X.
56. **Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC**. 2002. PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr D Biol Crystallogr **58:** 1948–1954. http://dx.doi.org/10.1107/S0907444902016657.
57. **Emsley P, Cowtan K**. 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr **60:**2126–2132. http://dx.doi.org/10.1107/S0907444904019158.
58. **McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ**. 2007. Phaser crystallographic software. J Appl Crystallogr **40:** 658–674. http://dx.doi.org/10.1107/S0021889807021206.
59. **Long F, Vagin AA, Young P, Murshudov GN**. 2008. BALBES: a molecular-replacement pipeline. Acta Crystallogr D Biol Crystallogr **64:** 125–132. http://dx.doi.org/10.1107/S0907444907050172.
60. **Tuomivaara ST, Yaoi K, O'Neill MA, York WS**. 2015. Generation and structural validation of a library of diverse xyloglucan-derived oligosaccharides, including an update on xyloglucan nomenclature. Carbohydr Res **402:**56–66. http://dx.doi.org/10.1016/j.carres.2014.06.031.
61. **Martinez-Fleites C, Guerreiro CI, Baumann MJ, Taylor EJ, Prates JA, Ferreira LM, Fontes CM, Brumer H, Davies GJ**. 2006. Crystal structures of Clostridium thermocellum xyloglucanase, XGH74A, reveal the structural basis for xyloglucan recognition and degradation. J Biol Chem **281:** 24922–24933. http://dx.doi.org/10.1074/jbc.M603583200.