Supporting Appendix:

Bacteria of the human gut microbiome catabolise red seaweed glycans with carbohydrate-active enzyme updates from extrinsic microbes

Jan-Hendrik Hehemann[1], Amelia G. Kelly[2], Nicholas A. Pudlo[2], Eric C. Martens[2] & Alisdair B. Boraston[1]

[1]Department of Biochemistry & Microbiology, University of Victoria, PO Box 3055 STN CSC, Victoria, British Columbia, V8W 3P6, Canada,

[2]Department of Microbiology & Immunology, University of Michigan Medical School, 1150 West Medical Center Drive , 5641 Medical Science II, Ann Arbor, 48109-5620, USA

**Correspondence should be addressed to: Alisdair B. Boraston, Biochemistry & Microbiology, University of Victoria, PO Box 3055 STN CSC, Victoria, BC, V8W 3P6, Canada. Tel: 250.472.4168. Fax: 250.721.8855. Email: boraston@uvic.ca or to Eric C. Martens, Department of Microbiology & Immunology, University of Michigan Medical School, 1150 West Medical Center Drive , 5641 Medical Science II, Ann Arbor, 48109-5620, USA, Tel: 734-647-5800. Email: emartens@umich.edu

Supplementary Figures S1-S5

Supplementary Tables S1-S8

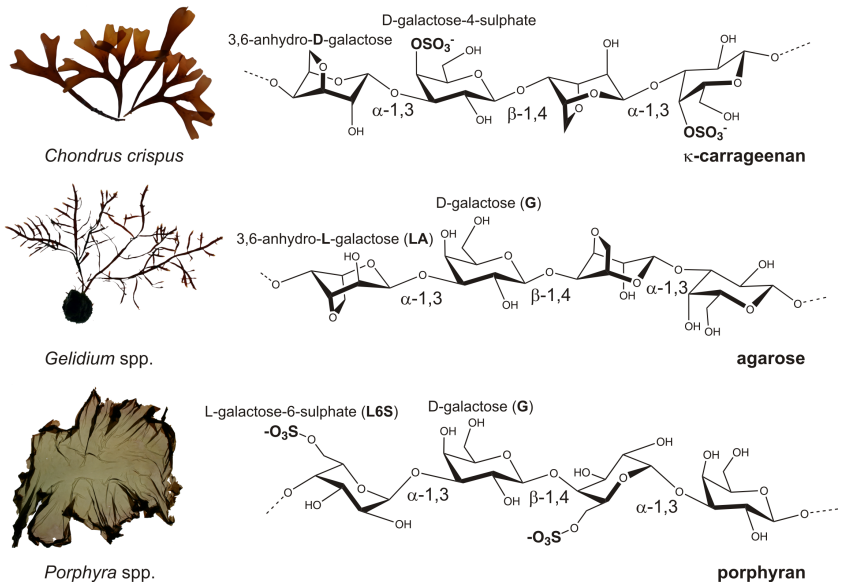

Figure S1: Commercial seaweeds and their major cell wall matrix polysaccharides. Shown are galactans from red seaweeds that are used as food additives (agar, carrageenan) or which are naturally consumed in form of dried seaweed (porphyran). These polysaccharides represent the major cell wall matrix polysaccharides in the shown species. Notably the depicted structures are idealized and these molecules are more heterogeneous in the wild.

Results:

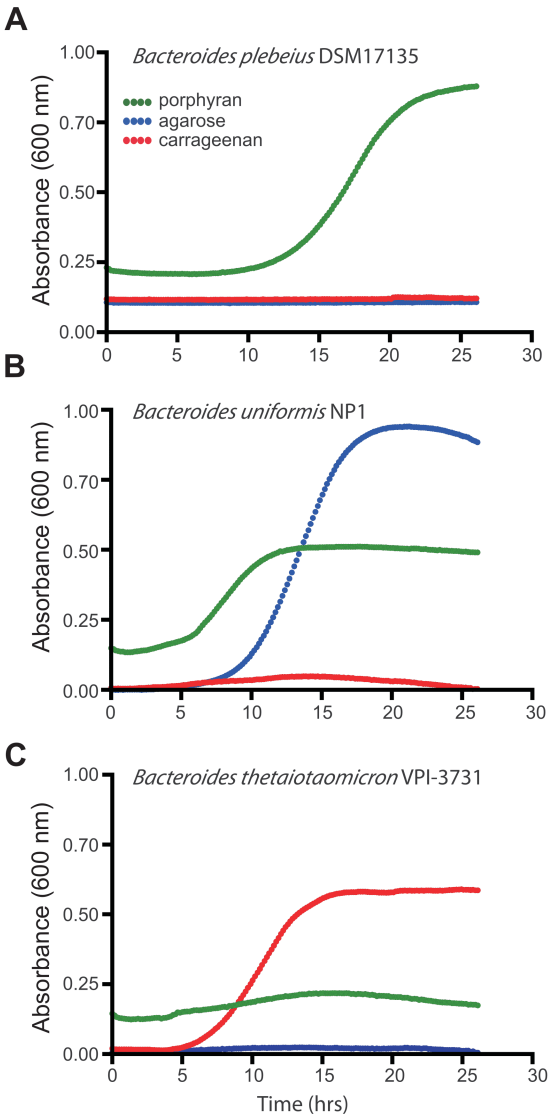

Figure S2: Gut Bacteroides grow on defined galactans from red seaweeds. Intestinal *Bacteroides* species were grown with carrageenan, porphyran or agarose. **(A)** *Bacteroides plebeius* grew selectively on porphyran. (**B)** *Bacteroides uniformis* NP1 grew on agarose and showed minor growth on porphyran. **(C)** *Bacteroides thetaiotaomicron* VP1-3731 did only grow on carrageenan. Note that growth on carrageenan is a trait that is only exhibited by this and not other strains of *B. thetaiotaomicron*.

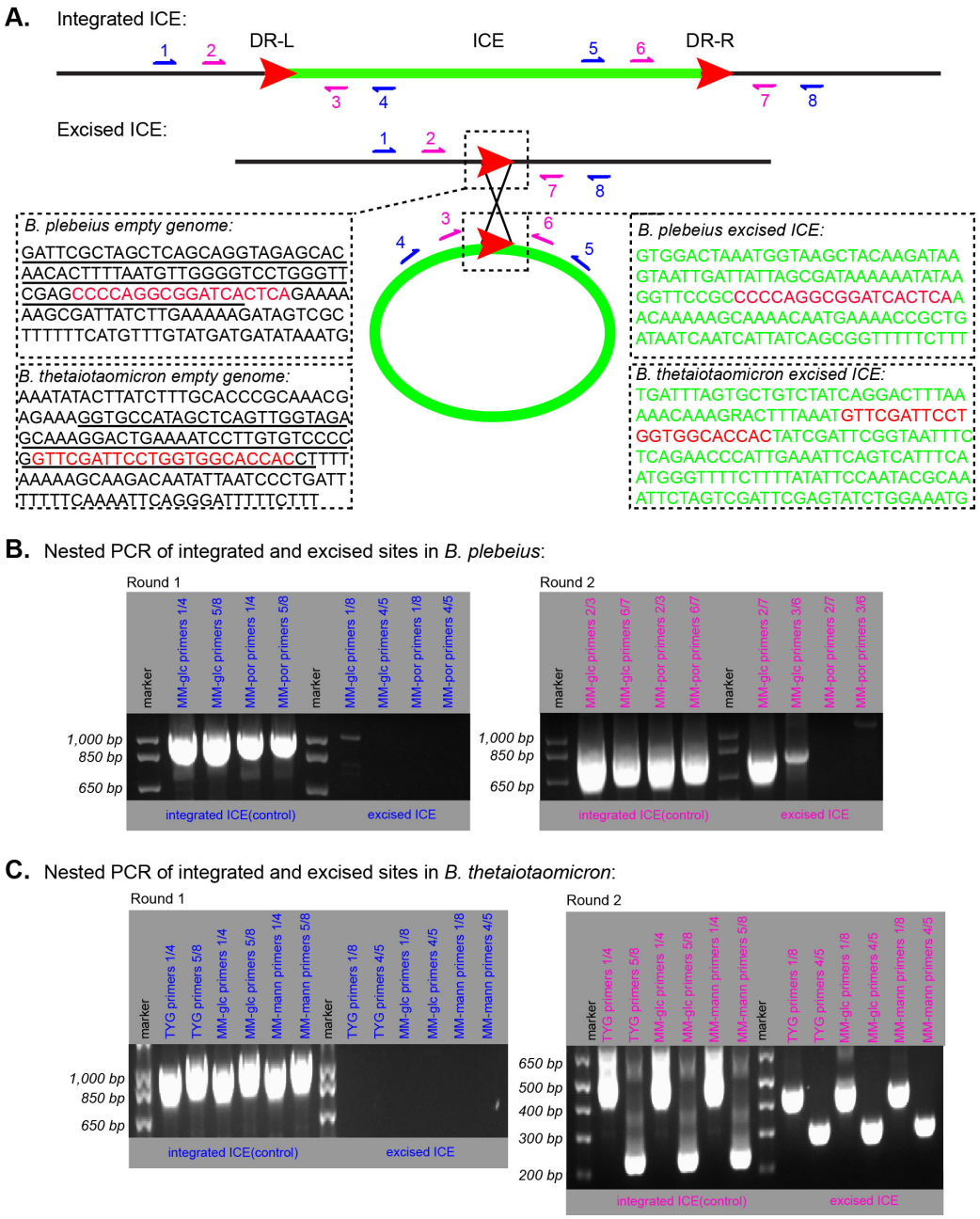

Figure S3. A. Schematic showing the locations of primers used to detect integrated and excised ICE forms: flanking genome region (black line), ICE (heavy green line), direct repeat ends of ICE (red arrow heads). Primers used for the first round of amplification were spaced up- and downstream from the direct repeats (numbered blue lines with barbs); nested primers used in the second round of amplification were spaced similarly but internal to the region amplified in the first round of PCR (numbered red lines with barbs). The DNA sequences in each box are derived from actual sequencing reactions performed on the recombined products in part B. below. Underlined

nucleotide sequences represent the tRNA sequences at which integration occurs and the remaining sequences are color coded according to the schematic.

B-C. Reaction products after round 1 (left gel images) and round 2 (right gel images). Note that only a weak excision product is formed with *B. plebeius* in the MM-glc condition with primers 1/8 after one round; no products were formed for *B. thetaiotaomicron*. After two rounds of amplification, both excised ICE bands are detected in the MM-glc condition only. The recombination event was confirmed by sequencing the 2/7 and 3/6 bands for each species after round 2. In round 1 of amplification for each species ~250ng of genomic DNA was used from cells grown to stationary phase in each condition [minimal medium (MM), plus either glucose (glc), porphyran (por), α-mannan (mann) or tryptone-yeast extract medium (TYG). In round 2, the DNA products derived from he first PCR round were purified using a PCR cleanup kit (Qiagen), eluted in the same volume of water as the original amplification reaction and 5μl used for the subsequent reaction. Each PCR reaction contained: 4mm of each primer, plus Platinum Taq polymerase (Invitrogen) according to the manufacturer's suggested reaction conditions. Reactions were cycled 35X in each round using an annealing temperature of 59°C with a 1.5 min extension time at 72°C.

Attempt to mobilize the *B. plebeius* porphyran ICE into *B. thetaiotaomicron* VPI-5482. The PCR data shown in **Fig. S3** support site-specific excision of both Bacteroides ICEs. The requirement for multiple rounds of PCR to detect these events suggest that they occur at a low frequency under the conditions tested. In order to determine if the *B. plebeius* ICE could be moved into another strain and thereby confer porphyran utilization, we performed an experiment in which a derivative of *B. thetaiotaomicron* VPI-5482 that harbors genomic markers for tetracycline (Tet) and erythromycin (Erm) resistance (both inserted at $tRNA^{ser}$ sites using NBU2-based vectors described in ref. (1) was grown in a mixed lawn with *B. plebeius* (anaerobic atmosphere containing 90% N_2, 5% H_2, 5% CO_2; 3 days at 37°C on BHI-blood agar). The presence of both strains was verified in the mixed culture by extracting genomic DNA from a sample and performing PCR with species-specific primers (qPCR primers used to measure expression of *B. plebeius* PUL genes, or *B. thetaiotaomicron* specific primers that target genes in the α-mannan PUL (1); each primer set was tested on the reciprocal strain as a control). After co-incubation the mixed lawn was resuspended in TYG medium and plated on 20 individual 10 cm petri plates containing BHI-blood agar with Tet and Erm to select for *B. thetaiotaomicron*. After three days of growth, resistant bacteria were again resuspended and the selection process repeated on 20 new BHI-blood plus Tet/Erm plates to insure that all *B. plebeius* was eliminated. The recovered *B. thetaiotaomicron* population was

homogenously suspended from all of the plates and used to inoculate 5ml of MM-porphyran (~10^7 cfu total); after 7 days of growth, no porphyran-degrading *B. thetaiotaomicron* could be enriched based on turbidity measurements. Additional PCR based tests of genomic DNA derived from the recovered *B. thetaiotaomicron* population using porphyran-specific qPCR primers and 35 cycles od amplification failed to show evidence that these genes were present in the recovered *B. thetaiotaomicron* population (pure *B. plebeius* and *B. thetaiotaomicron* DNA served as positive and negative controls, respectively).

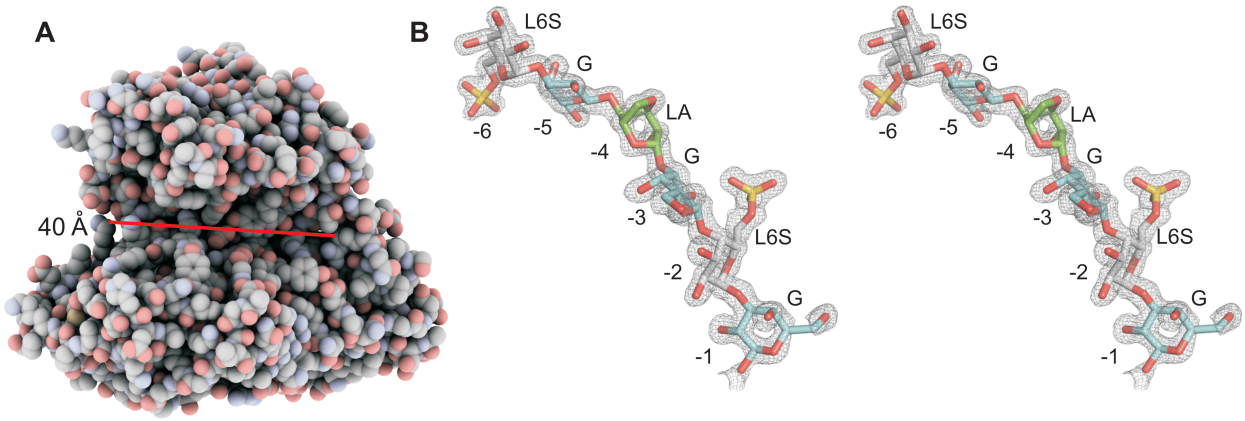

Figure S4: The extended active site of BpGH86A and its interaction with porphyran. (A) Space fill model with a view centered on the substrate binding cleft illustrated with the red bar of 40 Å length. **(B)** Stereo image of the hybrid porphyran/agarose hexa-oligosaccharide trapped in the active site of BpGH86A. The electron density of the hexasaccharide is shown in grey mesh (prior to modelling the sugar) as an unbiased maximum likelihood/σ_a-weighted $2F_o$-F_c map contoured at 1σ (0.19e/Å^3) produced by refinement of the coordinates prior to modeling the sugar. The L-galactose-6-sulphate (L6S) residues are depicted in grey, the D-galactose (G) residues in cyan and the 3,6-anhydro-L-galactose (LA) residue is shown in green. The figures were created with Pymol (2) and QuteMol (3).

The crystal structure of the porphyranase BpGH16B. The structure of BpGH16B was solved by molecular replacement using ZgPorB as search model (data collection and refinement statistics are summarized in **Table S1**). The two molecules in the asymmetric unit were modeled revealing the β-jelly

roll fold and the open substrate binding cleft common to the family 16 glycoside hydrolases (**Fig. S5B**). In addition to the high structural identity with the porphyranases from *Z. galactanivorans* (**Fig. S5C**), BpGH16B possesses the catalytic glutamate residues, Glu173 and Glu178 that are generally invariant in GH16 enzymes. The modeling of a product complex of BpGH16B based of the ZgPorA complex structure also reveals conservation of the active site residues that have defined the porphyranase subfamily of GH16 (**Fig. S5E**). A key feature is the pronounced pocket in the -2 subsite that accommodates the sulphate group of L-galactose-6-sulphate. At its base, His280 is appropriately positioned to form an ion-bond to the sulphate group. Notably, β-porphyranases possess plasticity in the use of cationic side chains at the base of the pocket, which can be arginines and histidines located in different places in the primary sequence. Thus, this histidine is not conserved in primary sequence yet the position of its side chain is structurally conserved with histidines and arginines from other β-porphyranases. A notable difference between BpGH16B and ZgPorA is the presence of a serine residue, Ser151, in BpGH16B which is an arginine in ZgPorA. This arginine approaches the α-face of the L-galactose-6-sulphate in subsite -1 rendering ZgPorA highly specific for this sugar and, through steric restriction, impeding a 3,6-anhydro-L-galactose from occupying this subsite. The active site of BpGH16B is more open at subsite -2 which may increase the tolerance of agarose motifs at this position and possibly expands the potential substrate range of the enzyme (**Fig. S5 E,F**), as it was observed for ZgPorB.

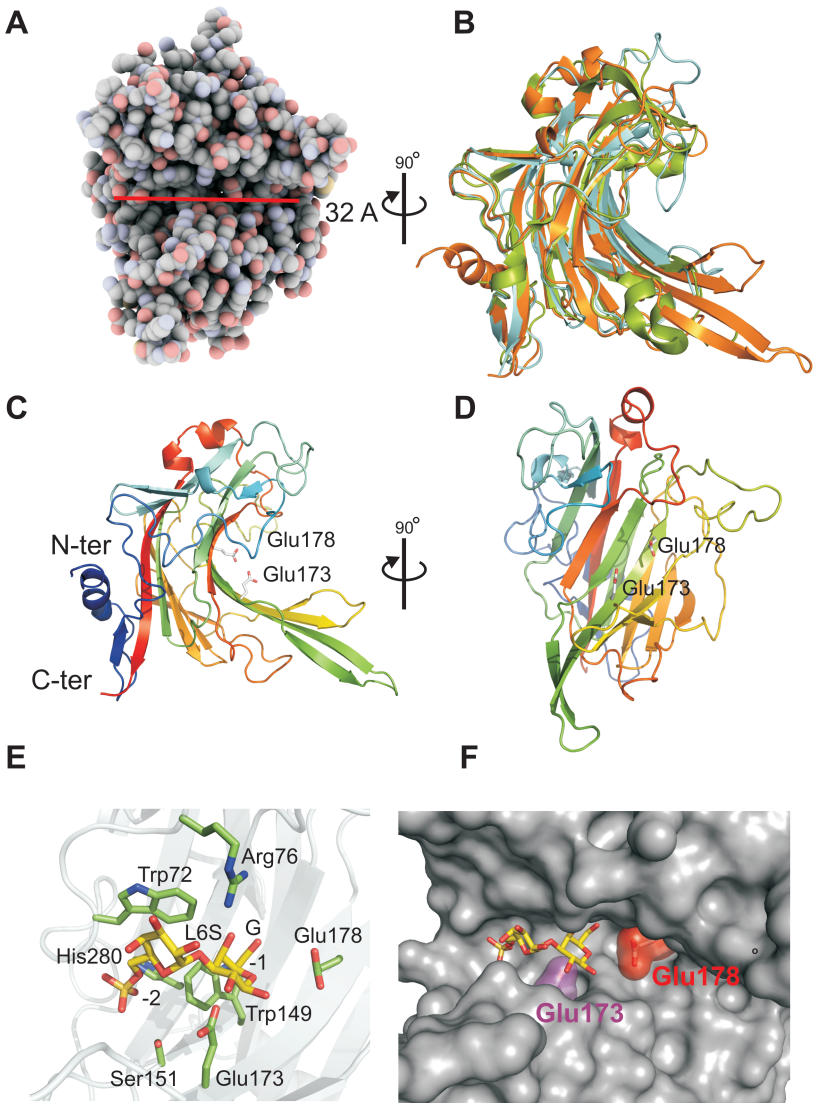

Figure S5: Structural comparison of the porphyranase BpGH16B from *B. plebeius* with homolog enzymes from the marine bacterium *Zobellia galactanivorans*. (**A)** Space fill plot of the overall structure with central view on the substrate binding cleft highlighted by the red line of 32 Å length. (**B**) Cartoon plot comparison of BpGH16B (orange) with ZgPorA (blue) and ZgPorB (green) from *Z. galactanivorans.* (**C)** Cartoon plot of the enzyme showing a view along the substrate binding cleft. (**D)** Frontal view focusing on the catalytic residues in the center of the substrate binding cleft. (**E)** Modelling of L6S-G disaccharide (yellow) into the 2 - binding subsites with the key residues for porphyran recognition and catalysis depicted as green sticks. (**F**). Surface representation of the active site of BpGH16B with the modelled L6S-G disaccharide in yellow bound to the subsites -1, -2 with the catalytic nucleophile in magenta and the acid/base coloured in red.

Materials and Methods:

Unless otherwise stated the chemical reagents were purchased from Sigma-Aldrich

Gene cloning, protein production and purification. The gene encoding residues 30-599 of the BpGH86A (BACPLE_01693) catalytic domain was amplified by polymerase chain reaction from *B. plebeius* DSM 17135 genomic DNA, which was obtained from the DMSZ (www.dmsz.de). The primers used for cloning of the recombinant proteins are listed in **Table S3**. BpGH16B (BACPLE_01689) was cloned encompassing residues 21-321 without the signal peptide (4). BpGH16A (BACPLE_01670) included residues 1-316. All of these PCR products were cloned into pET28a using standard directional cloning procedures with the restriction enzymes NheI and XhoI. The encoded proteins contained a thrombin cleavable N-terminal 6-histidine tag. The vector was transformed into chemically competent *E. coli* BL21(DE3) Star cells (Invitrogen) after verifying the sequences by bi-directional sequencing.

Transformed *E. coli* containing the PET28_BpGH86A, PET28_BpGH16A or PET28_BpGH16B constructs were grown in one litre cultures of autoinduction medium (5) containing 50 μg ml^{-1} kanamycin. Expression and purification by immobilized metal affinity chromatography (IMAC) was performed using previously described methods (6). After IMAC the protein BpGH16B was concentrated and further purified by size exclusion chromatography (Sepharose, S-200) using 20 mM Tris with pH 8. After IMAC the protein BpGH86A was dialyzed overnight against 20 mM Tris pH 8 and further purified by anionic exchange chromatography on a UNOQ column (Biorad), after which the fractions with pure protein were dialyzed against 20 mM Tris pH 8. BpGH16A was purified by IMAC only and the fractions with activity, assayed by agar plate lysis and lugol staining, were pooled, dialysed against 20 mM Tris pH 7.5, 200 mM NaCl and concentrated. All proteins were concentrated by stirred ultra-filtration using 10 kDa molecular weight cutoff membranes.

Protein activity and specificity assays. The activity assays were carried out in PBS buffer with 0.5 % pure porphyran (w/v) or with 0.5 % native porphyran (w/v). Both preparations were purified as previously described (7). For the activity curves 1 ml of porphyran solution (0.5 %) was incubated with 2 μg of the enzymes at 30 °C and samples were taken at time points, boiled for five minutes and analysed by reducing sugar assay (8) (calibrated with glucose as standard) and by fluorophore assisted carbohydrate gel electrophoresis (PACE) with 2-aminoacridone as fluorescent label (9). For the qualitative activity

tests on solid agar, 1 µl of the enzymes (~33 µM) were spotted onto a 1 % (w/v) agar plate prepared with PBS buffer. The agar plate was incubated at 20°C for three hours and activity was revealed by overlaying the plate with Lugol solution for 20 seconds (9). Heat inactivated enzymes served as controls. For end point measurements the porphyran solutions (0.1 % w/v) were incubated, in triplicate, with an excess of enzyme until completion of the reaction and the amount of released sugars was quantified by reducing sugar assay as described above. For the TLC analysis 0.5 % native porphyran solution in 1XPBS buffer was digested with an excess of enzyme, until completion and 3 µl of each reaction were spotted on the plate. The product separation was carried out with n-butanol/formic-acid/water (4:8:1, (v/v)) as solvent system on a POLYGRAM SIL G/UV254 (Macherey-Nagel) plate. The products were visualized with the naphtoresorcinol staining procedure (10) by dipping the dried plate into 0.1 % 1,3-dihydroxy-naphthalene (Sigma) in 5% ethanolic sulfuric acid and heating for 20 min at 110°C. Agarases (AgaA) and porphyranases (ZgPorA) from *Z. galactanivorans,* the reaction products of which have been previously described (7),(11),(12), served as comparative controls in some of the assays.

Protein crystallization. Crystallization conditions were screened at 18^{o}C by the vapour diffusion method in a 96 well sitting drop format with proteins at 10 mg ml^{-1} in 20 mM Tris pH 8. Optimization of initial hits was carried out in hanging drops at 18°C. BpGH16B crystallized by mixing equal volumes of protein (10 mg ml^{-1}) with 23-25 % PEG 3350, 100 mM Tris pH 8.5 and 5 mM calcium chloride. A suitable crystal for data collection was cryo protected by soaking the crystal in mother liquor with 20 % glycerol following by flash freezing in liquid nitrogen.

Initial attempts to obtain a high-resolution dataset of BpGH86A failed due to the low diffraction quality of crystals grown in 10 % PEG 10000, 0.1 M ammonium acetate, 0.1M BisTris pH 5.5. Pure porphyran (lyophilized) was added to the protein (10 mg ml^{-1}) to a final concentration of 0.4% (w/v), which crystallized in 30% PEG MME 2000 and 0.1 M potassium thiocyanate. Crystals for data collection were grown in 27% PEG MME 2000, 0.1 M potassium thiocyanate and 150 mM sodium iodide. Prior to data collection a crystal of BpGH86A was soaked for 10 minutes in mother liquor to which 1 M sodium iodide was added. For cryo protection the same solution was supplemented with 20% ethylene glycol. The crystals were flash frozen in a nitrogen stream and a dataset was collected in-house at 1.5419 Å. All data were processed with iMosflm (13) and Scala (14) as part of the CCP4 suite (15). Data collections at higher resolution were carried out either at BL9-2 at the Stanford Synchrotron radiation Laboratories

(Stanford, USA) or the 08ID-1 beam line at the Canadian Light Source (Saskatoon, Canada). Data collection and Refinement statistics are listed in **Table S1**.

Structure solution and Refinement. The structure of BpGH86A was solved by single wavelength anomalous dispersion method (SAD) using the iodide derivative. ShelXC/D (16) was used to identify 20 sites, which gave sufficient anomalous signal for phasing with ShelxE. The resulting electron density was modified with DM (15). Automatic model building was performed with ARP/wARP (17). The model was manually corrected in Coot (18) and completed by refinement with Refmac5 (19) against a high resolution dataset collected at the CLS 08ID-1 at a wavelength of 0.95369 Å. The dataset for BpGH16B was collected at SSRL BL9-2 at a wavelength of 0.97911 Å. The crystal structure of BpGH16B was solved by molecular replacement using PorB (PDB id: 3JUU) as search model and the program Molrep (20); the initial phases were sufficient for automated model building with ARP/wARP (17), which was manually completed in Coot and refined with Refmac5. In all cases water molecules were added using Coot Find Waters and manually inspected prior to deposition. Five percent of the observations were flagged as "free" and used to monitor refinement procedures. For the BpGH86A structure the Ramachandran statistics were 97.58%, 1.86% and 0.56% for preferred, allowed regions and outliers respectively. For the BpGH16B structure the Ramachandran statistics were 95.88%, 2.43% and 0.69% for preferred, allowed regions and outliers respectively. Data and refinement statistics are given in Table 1. Model validation was performed with MolProbity (21). The graphical representations of protein structures were created with Pymol (2) and QuteMol (3).

Screening for new red algae degrading gut isolates. The *B. thetaiotaomicron* VPI-3731 isolate which uses carrageenan for growth was identified during a screen of 292 different human gut-derived Bacteroidetes. Briefly, each isolate was screened for growth in minimal media supplemented with the algal polysaccharide kappa-carrageenan at 5mg/ml as sole carbon source along within a panel of other common plant and animal polysaccharides as previously described (22). *B. thetaiotaomicron* VPI-3731 was one of only three strains capable of degrading carrageenan and exhibited the largest growth increase under the conditions tested. *B. uniformis* NP1 was isolated from the same collection, but by enrichment culture where the entire community was simultaneously inoculated into a bottle of 5 mg/ml molten low-melt agarose (37°C) in minimal medium (this condition was not screened in assay described above with individual polysaccharides). The isolate was purified by serial sub-culture, followed by

streaking single colonies on BHI-10% horse blood plates to identify the agarose degraders, which were apparent by their agarolytic phenotype indicated by depletion of the agar surrounding the colonies.

Quantitative PCR (qPCR) of *B. plebeius* PUL genes. *B. plebeius* DSM 17135 was grown in triplicate on either galactose or porphyran as a sole carbon source and cells were harvested during exponential growth (A_{600} values of 0.41, 0.42, 0.43 for galactose and 0.57, 0.56, 0.57 for porphyran). Bacterial samples were preserved using RNAProtect (Qiagen), total cellular RNA extracted using the RNeasy kit (Qiagen) and cDNA synthesized using SuperScript III Reverse Transcriptase kit (Invitrogen). Transcript measurement of each PUL gene was performed on a Mastercycler realplex^2 (Eppendorf) using a SYBR green-based master mix (Kapa Biosystems, Woburn, MA) and oligonucleotides listed in **Table S4**. Transcript abundance in each sample was normalized using primers specific for the 16S rRNA transcript.

Molecular analysis of a putative excision mechanism of the PUL. To test for excision of both ICEs, we used a series of 8 primers for each element that were designed to align around the direct repeats at each ICE end. This strategy is similar to several other studies of ICE mobilization (23). Two of these primers align in the genome flanking the ICE, one each on the left and right side, and point toward the direct repeat. The remaining two primers align within the ICE facing outward towards the genomic flanks. A matrix of four PCR reactions was performed using each possible primer combination. If the ICE is integrated, then two products are generated that span the left and right junctions with the direct repeat in the middle of the product. If the ICE excises, then additional products would be detected that correspond to the circularized right and left ICE ends joining together and the corresponding excision site that remains in the genome. In this scheme the primers themselves serve as positive controls as they must be validated to at least give products of the anticipated size in the integrated ICE conformation. The PCR products derived from the hypothesized recombination events (primer combinations 1/8 and 4/5 in **Fig. S3**) were directly sequenced using both end primers to validate the proposed recombination.

Bacterial growth. *B. plebeius* DSM 17135, *B. thetaiotaomicron* VPI-3731 and *B. uniformis* NP1 cultures were routinely grown in a custom chopped meat broth (CMB) medium (**Table S5**). Minimal medium (MM), for testing growth on algal polysaccharides, was prepared as previously described (1), but with the addition of 0.5% chopped meat extract (**Table S5**), Balch's vitamin mixture (24) and three additional supplements that were necessary for growth of *B. plebeius* (**Table S6,7,8**). MM was filter sterilized at 2X

concentration and combined with an equivalent amount of 2X autoclaved polysaccharide (10 mg ml^{-1} in pure water). Growth profiles for each species were measured in 96-well plates (Costar) in an anaerobic chamber (Coy Lab Products, Grass Lake, MI) and using a robotic plate handling device and absorbance reader (Biotek, Winooski, VT) as previously described (22). For each growth profile, a total of 12 replicate 200 µl cultures were averaged. Carbohydrate-free MM assays were included for each strain to control for background growth on the supplements required to cultivate *B. plebeius*. A low amount of growth in the absence of added carbohydrate was only observed for *B. thetaiotaomicron* and this was subtracted from each growth profile for this species.

Table S1: Data collection and refinement statistics.

	BpGH86A-iodide	BpGH86A	BpGH16B
Data collection			
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2\ 2_12_1$
Cell dimensions (Å)			
a, *b*, *c*, (Å)	71.3, 87.7, 114.5	71.3, 87.7, 114.5	51.1, 107.2, 137.3
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	20.0-2.2 (2.32-2.2)	31.7-1.33 (1.41-1.33)	47.9-2.4 (2.5-2.4)
R_{merge}	0.148 (0.42)	0.062 (0.44)	0.13 (0.44)
$I/\sigma I$	15.0 (6.9)	15.8 (3.9)	10.5 (4.2)
Completeness (%)	99.8 (99.7)	99.7 (97.8)	99.9 (99.7)
Redundancy	13.0 (12.4)	7.3 (6.8)	6.3 (6.3)
Refinement			
Resolution (Å)		1.33	2.4
No. reflections		154892	28740
R_{work} / R_{free}		0.14/0.15	0.20/0.25
No. atoms			
Protein		4563	2403 (A); 2385 (B)
Ligand[b]		75	N/A
Water		594	374
B-factors			
Protein		14.078	17.5 (A); 21.3 (B)
Ligand[b]		19.291	N/A
Water		27.562	33.2
R.m.s. deviations			
Bond lengths (Å)		0.015	0.01
Bond angles (°)		1.6	1.3

Values in parentheses are for the highest-resolution shell.

Table S2. Table of the 35 genes analysed by RT-PCR located within the PUL with annotations

Locus Tag	Annotation (functionally verified)	Fold-induction (porphyran/galactose)	Std. dev. (n=3)
BACPLE_01667	RteB response regulator	1.5	0.8
BACPLE_01668	RteB response regulator	11.5	2.0
BACPLE_01669	RteA two component sensor histidine kinase	13.1	4.4
BACPLE_01670	β-agarase GH16A	21.0	4.0
BACPLE_01671	α-agarase GH117 (6)	41.2	5.9
BACPLE_01672	β-galactosidase-like GH2	61.4	3.2
BACPLE_01673	β-galactosidase-like GH2	33.4	6.3
BACPLE_01674	D-altronate dehydratase	181.8	31.9
BACPLE_01675	Mannitol-1-phosphate/altronate dehydrogenases	187.5	31.2
BACPLE_01676	Threonine dehydrogenase	321.3	71.6
BACPLE_01677	Fucose permease	77.8	17.1
BACPLE_01678	Predicted metal dependent hydrolase of TIM-barrel fold	366.4	124.5
BACPLE_01679	Predicted oxidoreductase (aryl-alcohol dehydrogenase related)	942.0	246.3
BACPLE_01680	Uncharacterized conserved protein	1056.0	214.1
BACPLE_01682	Hypothetical protein	144.7	62.7
BACPLE_01683	Distantly related to GH50 β-agarases (TIM-barrel fold)	115.1	22.5
BACPLE_01684	Predicted unsaturated glucuronyl hydrolase (GH88-related)	150.8	49.1
BACPLE_01685	Hypothetical protein	353.8	162.6
BACPLE_01686	Hypothetical protein	481.7	66.4
BACPLE_01689	β-porphyranase GH16 (BpGH16A)	180.6	36.8
BACPLE_01692	Hypothetical protein	3.3	1.0
BACPLE_01693	β-porphyranase GH86 (BpGH86A)	5.8	2.2
BACPLE_01694	β-porphyranase-like GH86	7.2	2.8
BACPLE_01695	Hypothetical protein	6.1	1.8
BACPLE_01696	Contains domain very distantly related to CBM22	8.4	2.2
BACPLE_01697	SusD-like	10.7	1.6
BACPLE_01698	SusC-like	10.0	2.0
BACPLE_01699	HTCS sensor/regulator (TM)	2.7	0.7
BACPLE_01700	Alcohol dehydrogenase (Class IV)	117.2	27.6
BACPLE_01701	Arylsulfatase	163.0	48.3
BACPLE_01702	GH29 (TIM-barrel)	223.6	31.0
BACPLE_01703	β-glucanase (GH16)	250.5	75.0
BACPLE_01704	SusD-like	127.0	39.0
BACPLE_01705	SusC-like	304.6	38.1
BACPLE_01706	β-galactosidase (GH2)	357.6	48.0

Table S3. **Oligonucleotide primers used for recombinant cloning of porphyranases and agarases**

Name	Species	Sequence (5'-3')
BpGH86A_FW	*Bp*	CATATGGCTAGCGTCGACTATAATACCCGTCG
BpGH86A_RV	*Bp*	GTGTGTCTCGAGTTAAGAGACAGTATATTTCTCAACC
BpGH16B_FW	*Bp*	CATATGGCTAGCAAAAATGATAAGGAATATTCTTTGG
BpGH16B_RV	*Bp*	GTGTGTCTCGAGTCACTCTTCTATTGGAACCAACT
BpGH16A_FW	*Bp*	CATATGGCTAGCATGAAGAGAAAACTGTTTACTATT
BpGH16A_RV	*Bp*	GTGTGTCTCGAGCTATTCTTCTGGGACCAGTG

Table S4. **Oligonucleotide primers used for RT-PCR and ICE detection/sequencing.**

Name	Species	Sequence (5'-3')
BP1667F	*Bp*	AGGTCGCTTTCGTGAGGATTTG
BP1667R	*Bp*	CCGCACATTACCAGGCCAGTCATA
BP1668F	*Bp*	AAGACAAATCGGACCATAATACAATAACTA
BP1668R	*Bp*	AGAGCGGCCACTACTTCACG
BP1669F	*Bp*	TACTGGCCAATCTACTTACAAACGCTATC
BP1669R	*Bp*	AGTCCCAAACCAAAACCTTCAGA
BP1670F	*Bp*	ATTGTGGCGCGATGATTTTC
BP1670R	*Bp*	CCCGTGCCACCAGTGTAGTTGT
BP1671F	*Bp*	GAGCGTAAGGCCGATTCATTGG
BP1671R	*Bp*	AGCATTGCACTCGGGTCTCTTCTTACTA
BP1672F	*Bp*	ATGGAAGGCGGTGATTGAAAGAGA
BP1672R	*Bp*	GCCGCACATATCCAGGAGACCACAT
BP1673F	*Bp*	CTCCGGATATACAGACGTGCTTGAT
BP1673R	*Bp*	GATGCCCGCTATGTACTTCTTGTC
BP1674F	*Bp*	AACAAGGCGGGAGGTATTTC
BP1674R	*Bp*	GCCCGGTGCGCTCAACAGATTA
BP1675F	*Bp*	GCTTGCCGGCATGAGGTGATAGGT
BP1675R	*Bp*	CGGAAGGTCGCGTGTGGCATACT
BP1676F	*Bp*	TCAATGCCTGCGAACACAACGAGAC
BP1676R	*Bp*	CGCAACCTATCACGGCCACAGTATCA
BP1677F	*Bp*	CTTGGCGGCAGATGGATAGGT
BP1677R	*Bp*	GCCGAGGTCGCGTAGAGCAAGTG
BP1678F	*Bp*	TGCTTGCCCGCCATAAAAAT
BP1678R	*Bp*	GTGCGCGGATAGTCGGAACC
BP1679F	*Bp*	GACCATGCAGCTCCGGACACAGTT
BP1679R	*Bp*	GCGTTTATCACCCCTACACCCTTTTCA
BP1680F	*Bp*	TCCGGAATTGATTGAAGCCTACAG
BP1680R	*Bp*	ATTTTGCCACATATTCCTCCCATTC
BP1682F	*Bp*	GTGGCCGACGGACTGATTCTCT
BP1682R	*Bp*	AGCTCCTTATAACTCCCCCTGGTGT

BP1683F	*Bp*	TCGAGTATGCGGAAAAACCTGACA
BP1683R	*Bp*	GAATTTTTGCCCTTTTGCGTATGAGC
BP1684F	*Bp*	ATATGGCGAAAGGGGAACTACAGA
BP1684R	*Bp*	CGTCAGGCAGGCTCAACCAT
BP1685F	*Bp*	GTGTGGCCGGTTCTTGA
BP1685R	*Bp*	AGTCCATTCTGCTCCGTTGTTGA
BP1686F	*Bp*	TGATGGAACACCGGCTGAGGAG
BP1686R	*Bp*	TTGTAACCGGTATAAGGGAGATT
BP1689F	*Bp*	CGGCCGCCTGCAACTTTCAA
BP1689R	*Bp*	ATATTTATCGCCCGGCTTTCCATCA
BP1692F	*Bp*	ATCCTTTGCTCCGTTCATATCTTTACA
BP1692R	*Bp*	TATCCTCTTCGTGACATCAATCTCCA
BP1693F	*Bp*	GGTGTTCCTCCCGGTATTTATTGTC
BP1693R	*Bp*	TTGCCTACCTCATGCGTCTTGTTAT
BP1694F	*Bp*	CCGCTCTTTCTGGGGACCGTTCTCTTAT
BP1694R	*Bp*	AGCCCATGCACCCGCCTTCACT
BP1695F	*Bp*	GGTTTTGTTCGGAGCATGTGATAAT
BP1695R	*Bp*	AAACTTGGAAACCTCGCCTGTCT
BP1696F	*Bp*	TCCGCGTGTATGATCCTAACCTC
BP1696R	*Bp*	ACGCCAGTATCCGCACATTTG
BP1697F	*Bp*	CTTTGCGTGGCTTGTTTTATTTTTATCT
BP1697R	*Bp*	ATCGGTTTTCTCCCATTTGTCATTC
BP1698F	*Bp*	GTCAATAATAACGCGCCAATACAGAA
BP1698R	*Bp*	ACCCCCAAGCCGGAGATAACT
BP1699F	*Bp*	GCATGCCATCGAGTCAAAACCAC
BP1699R	*Bp*	AGCCCTGAGCCTTCTGTACCTAAATA
BP1700F	*Bp*	GGTGCGGCAATCTGTCGTGAGG
BP1700R	*Bp*	CGGTTCCGGCTGTTGTAGG
BP1701F	*Bp*	AATGGGGCTCCGAAGGCACAA
BP1701R	*Bp*	CTTTACGCAACGCAGATACGATTTCAG
BP1702F	*Bp*	TCTGCGTTTCGGACTTTCTTCTCACA
BP1702R	*Bp*	ACCTTCGGGCTCATGCAGTCTTTC

BP1703F	*Bp*	CACTGCCCGAAGGATATACTCACAC
BP1703R	*Bp*	CGACATCCAGAAACCGAAAACC
BP1704F	*Bp*	TGATGCGGCGATATTTACTTACGA
BP1704R	*Bp*	CAGGCTTGTTTAGATTGGGGTTTC
BP1705F	*Bp*	TTGACAGATGGCGGAAAGAAGGT
BP1705R	*Bp*	CATAAACGGATAGCTGGTGAAAGTGG
BP1706F	*Bp*	TACCGGCGGGCTGAAAGAATACTG
BP1706R	*Bp*	TAATGGAGCGGTCGGCACTGATAAC
B. plebeius 16SF	*Bp*	GGTAGTCCACACGGTAAACGATGGA
B. plebeius 16SR	*Bp*	CCCGTCAATTCCTTTGAGTTTC
ICE PCR (numbered according to schematic in Fig. S3)		
B. plebeius ICE 1	*Bp*	AGGAGAACGTACCCGTCTGG
B. plebeius ICE 2	*Bp*	CATTCTGAAGCAGGCGCTGG
B. plebeius ICE 3	*Bp*	CCAATCGCGCATTCTTTTTGCG
B. plebeius ICE 4	*Bp*	TGCTGTAATAGGTTTACCGTCC
B. plebeius ICE 5	*Bp*	AGAAAAGCATAACAGTCATGAGC
B. plebeius ICE 6	*Bp*	AAGTATTCTGAAAGATAATACCCC
B. plebeius ICE 7	*Bp*	ATGCGAATTGGCTATGCCTGC
B. plebeius ICE 8	*Bp*	TCTGCCCGGAGGATGTAGCC
B. theta ICE 1	*Bt*	CAGTGTATAATAATTGGTATTGTGG
B. theta ICE 1	*Bt*	ATTGTTTGTAGTGCCCTGATCG
B. theta ICE 1	*Bt*	CCATTACATTCAAGCAGGTATGC
B. theta ICE 1	*Bt*	CATTAGAATAGGTGCTTCACC
B. theta ICE 1	*Bt*	CGGACAGATACAGACCTATCC
B. theta ICE 1	*Bt*	GATTTAGTGCTGTCTATCAGGAC
B. theta ICE 1	*Bt*	GAAAAGCTGTTGCTTAACATGAG
B. theta ICE 1	*Bt*	CAGTCTATGCATGTCTTCTG

Table S5. Chopped Meat Broth. All components were dissolved into ddH_20 and brought to 250 ml, pH 7.2. Medium was filter sterilized, wrapped in foil and stored at 37°C in an anaerobic chamber prior to use.

Beef Extract	2.5 g
Pancreatic digest of casein	7.5 g
Yeast Extract	1.25 g
K_2HPO_4	1.25 g
L-Cysteine	250 mg
Fructose	250 mg
Galactose	250 mg
Glucose	250 mg
Mannose	250 mg
N-Acetyl Glucosamine	250 mg
Xylose	250 mg
K_3	250 µg
B_{12}	2.5 µg
Hematin/L-Histidine stock	7.6 µM/800 µM final
Balch's Vitamins	2.5 ml
Trace Mineral Solution	2.5 ml
Purine and Pyrimidine Solution	2.5 ml
Amino Acid Solution	2.5 ml

Table S6. Trace Mineral Solution. All components were dissolved into ddH_20 and brought to 1L, pH 7.0. Solution was filter sterilized and stored at room temperature.

EDTA	0.5
$MgSO_4*7H_2O$	3
$MnSO_4*H_2O$	0.5
NaCl	1
$FeSO_4*7H_2O$	0.1
$CaCl_2$	0.1
$ZnSO_4*7H_2O$	0.1
$CuSO_4*5H_2O$	0.01
H_3BO_3	0.01
$Na_2MoO_4*2H_2O$	0.01
$NiCl_2*6H_2O$	0.02

Table S7. Purine and Pyrimidine Solution. All components were dissolved into ddH_20 and brought to 1L, pH 7.0. Solution was filter sterilized and stored at room temperature.

Adenine	200 mg
Guanine	200 mg
Thymine	200 mg
Cytosine	200 mg
Uracil	200 mg

Table S8. Amino Acid Solution. All amino acids (62.5 mg each) were dissolved into ddH_20 and brought to 1L, pH 7.0. Solution was filter sterilized and stored at room temperature.

Alanine	Glutamic Acid	Leucine	Serine
Arginine	Glutamine	Lysine	Threonine
Asparagine	Glycine	Methionine	Tryptophan
Aspartic Acid	Histidine	Phenylalanine	Tyrosine
Cysteine	Isoleucine	Proline	Valine

REFERENCES

1. 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.

2. Delano W (2002) The PyMOL Molecular Graphics System. Available at: http://www.pymol.org.

3. Tarini M, Cignoni P, Montani C (2006) Ambient Occlusion and Edge Cueing for Enhancing Real Time Molecular Visualization. *IEEE Transactions on Visualization and Computer Graphics* 12:1237 –1244.

4. Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 340:783–795.

5. Studier FW (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* 41:207–234.

6. Hehemann J-H, Smyth L, Yadav A, Vocadlo DJ, Boraston AB (2012) Analysis of Keystone Enzyme in Agar Hydrolysis Provides Insight into the Degradation (of a Polysaccharide from) Red Seaweeds. *J Biol Chem* 287:13985–13995.

7. Correc G, Hehemann J-H, Czjzek M, Helbert W (2011) Structural analysis of the degradation products of porphyran digested by Zobellia galactanivorans [beta]-porphyranase A. *Carbohydrate Polymers* 83:277–283.

8. Lever M (1972) A new reaction for colorimetric determination of carbohydrates. *Analytical Biochemistry* 47:273–279.

9. Hehemann J-H et al. (2010) Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature* 464:908–912.

10. Duckworth M, Yaphe W (1970) Thin-layer chromatographic analysis of enzymic hydrolysates of agar. *Journal of Chromatography A* 49:482–487.

11. Jam M et al. (2005) The endo-beta-agarases AgaA and AgaB from the marine bacterium Zobellia galactanivorans: two paralogue enzymes with different molecular organizations and catalytic behaviours. *Biochem J* 385:703–713.

12. Allouch J et al. (2003) The three-dimensional structures of two beta-agarases. *J Biol Chem* 278:47171–47180.

13. Battye TGG, Kontogiannis L, Johnson O, Powell HR, Leslie AGW (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* 67:271–281.

14. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* 62:72–82.

15. The CCP4 suite: programs for protein crystallography (1994) *Acta Crystallogr D Biol Crystallogr* 50:760–763.

16. Sheldrick GM (2007) A short history of *SHELX*. *Acta Crystallogr A Found Crystallogr* 64:112–122.

17. Langer G, Cohen SX, Lamzin VS, Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* 3:1171–1179.

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

19. Murshudov GN et al. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D Biol Crystallogr* 67:355–367.

20. Vagin A, Teplyakov A (1997) MOLREP : an Automated Program for Molecular Replacement. *Journal of Applied Crystallography* 30:1022–1025.

21. Davis IW et al. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Research* 35:W375–W383.

22. Martens EC et al. (2011) Recognition and Degradation of Plant Cell Wall Polysaccharides by Two Human Gut Symbionts. *PLoS Biol* 9:e1001221.

23. Flannery EL, Antczak SM, Mobley HLT (2011) Self-Transmissibility of the Integrative and Conjugative Element ICEPm1 between Clinical Isolates Requires a Functional Integrase, Relaxase, and Type IV Secretion System. *J Bacteriol* 193:4104–4112.

24. Balch WE, Fox GE, Magrum LJ, Woese CR, Wolfe RS (1979) Methanogens: reevaluation of a unique biological group. *Microbiol Rev* 43:260–296.