How members of the human gut microbiota overcome the sulfation problem posed by glycosaminoglycans


To cite this version:
Alan Cartmell, Elisabeth Lowe, Arnaud Basle, Susan Firbank, Didier Ndeh, et al.. How members of the human gut microbiota overcome the sulfation problem posed by glycosaminoglycans. Proceedings of the National Academy of Sciences of the United States of America , National Academy of Sciences, 2017, 114 (27), pp.7037-7042. <10.1073/pnas.1704367114>. <hal-01802836>

HAL Id: hal-01802836
https://hal.archives-ouvertes.fr/hal-01802836
Submitted on 8 Jun 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
How members of the human gut microbiota overcome the sulfation problem posed by glycosaminoglycans

Alan Cartmell1,a, Elisabeth C. Lowe1,a, Arnaud Baslé4, Susan J. Firbank3, Didier A. Ndeh3, Heath Murray3, Nicolas Terraponb, Vincent Lombardb, Bernard Henriussab,c,d, Jeremy E. Turnbull4, Mirjam Czjzek1,g, Harry J. Gilbert4, and David N. Bolam2,a

1Institute for Cell and Molecular Biosciences, Newcastle University, Newcastle upon Tyne NE2 4HH, United Kingdom; 2Architecture et Fonction des Macromolécules Biologiques, CNRS, Aix-Marseille University, F-13288 Marseille, France; 3Institut National de la Recherche Agronomique, USCA108 Architecture et Fonction des Macromolécules Biologiques, F-13288 Marseille, France; 4Department of Biological Sciences, King Abdulaziz University, Jeddah 21589, Saudi Arabia; 5Centre for Glyobiology, Department of Biochemistry, Institute of Integrative Biology, University of Liverpool, Liverpool L69 72B, United Kingdom; 6Sorbonne Universités, Université Pierre-et-Marie-Curie, Université Paris 06, F-29688 Roscoff cedex, Bretagne, France; and 7CNRS, UMR 8227, Integrative Biology of Marine Models, Station Biologique de Roscoff, F-29688 Roscoff cedex, Bretagne, France

Edited by Carolyn R. Bertozzi, Stanford University, Stanford, CA, and approved May 23, 2017 (received for review March 17, 2017)

The human microbiota, which plays an important role in health and disease, uses complex carbohydrates as a major source of nutrients. Utilization hierarchy indicates that the host glycosaminoglycans heparin (Hep) and heparan sulfate (HS) are high-priority carbohydrates for Bacteroides thetaiotaomicron, a prominent member of the human microbiota. The sulfation patterns of these glycosaminoglycans are highly variable, which presents a significant enzymatic challenge to the polysaccharide lyases and sulfatases that mediate degradation. It is possible that the bacterium recruits lysates with highly plastic specificities and expresses a repertoire of enzymes that target substrates of the glycosaminoglycans with variable sulfation or that the glycans are desulfated before cleavage by the lyases. To distinguish between these mechanisms, the components of the B. thetaiotaomicron Hep/HS degrading apparatus were analyzed. The data showed that the bacterium expressed a single-surface endo-acting lyase that cleaved HS, reflecting its higher molecular weight compared with Hep. Both Hep and HS oligosaccharides imported into the periplasm were degraded by a repertoire of lyases, with each enzyme displaying specificity for substructures within these glycosaminoglycans that display a different degree of sulfation. Furthermore, the crystal structures of a key surface glycan binding protein, which is able to bind both Hep and HS, and periplasmic sulfatases reveal the major specificity determinants for these proteins. The locus described here is highly conserved within the human gut Bacteroides, indicating that the model developed is of generic relevance to this important microbial community.

The human colonic microbiota (CM) is crucial to health (1–3). The composition of the CM depends on its ability to access nutrients, which are primarily dietary and host glycans. Dissecting the mechanisms by which complex carbohydrates are used by the CM is critical to understanding the drivers of the ecology of this microbial community and how this process relates to human health. The major glycan degraders in the CM are the Bacteroidetes (4–6). These organisms access their target polysaccharides through endo-acting enzymes on the bacterial surface followed by import of the oligosaccharides generated, which are depolymerized in the periplasm. The genes encoding these enzyme systems are physically linked into loci termed polysaccharide utilization loci (PULs) (7). A significant proportion of the complex carbohydrates available to the CM is mammalian in origin (4). Despite this knowledge, our understanding of how the CM accesses host/mammalian glycans is fragmentary. Models for the breakdown of high-mannose and complex N-glycans by gut Bacteroides have been proposed (8), and the ecological significance of removing terminal sialic acid from the epithelial mucosa is established (9). There is, however, a paucity of information on the mechanism by which glycosaminoglycans (GAGs), such as heparin (Hep) and heparan sulfate (HS), are used by the CM. HS is a major component of the extracellular matrix of mammalian cells and therefore, likely to be available to the gut microbiota via turnover of epithelial cells, whereas Hep is released from mast cells at sites of injury and therefore, may not be as prevalent as HS in the gut (10, 11).

Microbial utilization of Hep and HS poses significant biological challenges. Both glycans differ significantly in their degree of polymerization (DP), which suggests that degradation may occur in different cellular locations, whereas sulfation patterns and the uronic acid (UA) are also variable (Fig. 1) (10). This substantial heterogeneity indicates that either a complex portfolio of enzymes is required to deconstruct these acidic glycans or the enzymes that mediate this process display significant substrate promiscuity. The depolymerization of UA-containing glycans, such as Hep/HS, is mediated by glycoside hydrolases (GHs) and/or polysaccharide lyases (PLs) that are grouped into sequence-based families in the CAZy database (12). Hep/HS are degraded by bacterial PLs belonging to families PL12, -13, and -21. Based on specificity, these PLs can be further broadly grouped into three functional groups, which is required to deconstruct these acetylated glycans. The data showed that the bacterium expressed a single-surface endo-acting lyase that cleaved HS, reflecting its higher molecular weight compared with Hep. Both Hep and HS oligosaccharides imported into the periplasm were degraded by a repertoire of lyases, with each enzyme displaying specificity for substructures within these glycosaminoglycans that display a different degree of sulfation. Furthermore, the crystal structures of a key surface glycan binding protein, which is able to bind both Hep and HS, and periplasmic sulfatases reveal the major specificity determinants for these proteins. The locus described here is highly conserved within the human gut Bacteroides, indicating that the model developed is of generic relevance to this important microbial community.

human gut microbiota | glycosaminoglycan degradation | heparin | heparan sulfate | Bacteroides thetaiotaomicron

Significance

The major nutrients available to the human microbiota are complex carbohydrates. Host glycans are important to this microbial community, particularly when dietary carbohydrates are scarce. The host glycans heparin and heparan sulfate are high-priority carbohydrates for Bacteroides thetaiotaomicron, a member of the human microbiota. The degradation of these complex carbohydrates is challenging, reflecting their highly variable sulfation patterns. How bacteria have adapted to depolymerize the myriad of structures of this important class of glycosaminoglycan is unknown. Here, we show how enzyme consortia, displaying complementary functions, target the different features of these host glycans. Structural data reveal that the acidic groups of the glycans are key specificity determinants for enzymes and binding proteins that make up the degradative apparatus.


The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Frequently available online through the PNAS open access option.

Data deposition: The crystallography, atomic coordinates, and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 4AK1, 4AK2, 5G2T, and 5G2U). A.C. and E.C.L. contributed equally to this work. 1To whom correspondence should be addressed. Email: david.bolam@ncl.ac.uk.

This article contains supporting information online at www.pnas.org/cgi/doi/10.1073/pnas.1704367114/-/DCSupplemental.
Hep I–III. Hep I lyases require sulfation, Hep II PLs exhibit promiscuity with respect to sulfation patterns, and Hep III enzymes cleave low-sulfation regions of these GAGs (13, 14). The lyases generate products capped by Δ4,5-unsaturated UA, which is released from the GAG by GHS from family GH88 (13). In addition, some GAG-specific sulfatases have been characterized (15).

Although enzymes active against Hep and HS have been described, how these PLs, GHS, and sulfatases are tailored to act in unison to address the sulfation problem posed by these GAGs is unclear. Several scenarios can be proposed, including desulfation before backbone cleavage, the exploitation of broad specificity PLs that can cleave Hep and HS independent of their sulfation, or the recruitment of a consortium of lyases, with each enzyme targeting specific substructures in these GAGs that differ in their sulfation profiles. Similarly, the mechanisms by which the cell surface glycan binding proteins (SGBPs), which contribute to the glycan degradation machinery, target conserved features of these heterogeneous glycans are unknown.

Bacteroides thetaiotaomicron (Bt), a member of the CM, uses Hep and HS as high-priority nutrient sources that activate a single PUL (PUL$_{Hep}$) (16). Here, we have dissected the mechanism by which the enzymes and binding proteins encoded by this locus fully deconstruct these highly variable GAGs and thus, solve the sulfation problem. Crystal structures of key proteins provide mechanistic insight into substrate and ligand recognition. The data revealed how the specificity of the apparatus is optimized to target the repertoire of GAG structures presented to the bacterium at the different stages of the degradative process.

**Results**

**Structure of Hep and HS.** Hep and HS are composed of a disaccharide repeating unit comprising a UA, α-gluconic (GlcA), or 1-iduronic acid (IdoA) alternating with α-gluconic (GlcN). GlcN is linked α1,4 to the UA, whereas IdoA and GlcA are linked α1,4 and β1,4, respectively, to the amino sugar (10) (Fig. 1). GlcN can be sulfated on O6 and sulfated or acetylated at N2 (GlcNAc), whereas the UAs are often sulfated on O2. Hep contains significantly more IdoA than HS and is almost completely sulfated, whereas HS varies in its sulfation pattern, containing highly sulfated regions (Hep-like) and areas that contain little or no sulfation (Fig. 1). The DP of Hep (∼40) is substantially less than that of HS (∼80–200).

**PUL$_{Hep}$ Structure.** Bt grows on Hep, HS, and N-acetylated-β-O-sulfated heparin (ΔSHept) (Fig. S1) (16, 17). Transcriptomics of Bt cultured on Hep indicate that only PUL$_{Hep}$, extending from bt4652 to bt4675, was up-regulated in response to this GAG (16). The locus encodes PLs, sulfatases, a GH88, and a member of the repressor, ORF, kinase (ROK, Pfam PF00480) family (1). In addition, the sulfatase BT1596 is also up-regulated by Hep (15). PUL$_{Hep}$ encodes a single SusC/SusD-like outer membrane glycan transporter system and a potential SGBP (Fig. 1). In silico analysis of the occurrence of PUL$_{Hep}$ suggests that it is widely distributed within the *Bacteroides* and retains high sequence conservation and synteny (Fig. S2).

**SGBPs.** The extracellular location of the putative SGBP, BT4661, was revealed by immunofluorescence and proteinase K treatment using antibodies against the protein (Fig. S3 A and B). Immunoprecipitation of BT4661 from Hep-grown cells followed by Western blotting revealed the presence of BT4659$_{SusD-like}$, indicating that the two proteins physically associate (Fig. S3C). The interaction between the PUL-encoded SGBP and SusD-like is consistent with previous data from the starch utilization system of *Bt* (18). The importance of BT4659$_{SusD-like}$ in Hep/HS utilization was highlighted by the severe lag (∼20 h) displayed by Δbt4659$_{SusD-like}$ on these GAGs (Fig. S3A). By contrast, Δbt4661$_{SusC-like}$ had no growth defect on HS or ΔShp but had a noticeable phenotype when grown on Hep-derived oligosaccharides (Fig. S3A, Inset). These data suggest that the SGBP plays a role in oligosaccharide scavenging rather than polysaccharide acquisition, such as observed in *Bacteroides ovatus*-mediated xyloglucan degradation (19).

**Glycan specificity.** Isothermal titration calorimetry (ITC) (Fig. 2 and Table S1) revealed that BT4661$_{SusC-like}$ bound tightly to both Hep and HS (Kd ~ 3 μM) but did not bind to other GAGs. Notably, coverage on both GAGs was similar, suggesting the protein can tolerate both sulfated and unsulfated regions and could interact or accommodate the different UAs and neutral sugars in Hep and HS (Fig. 1). BT4661$_{SusC-like}$ bound to oligosaccharides with a DP ≥ 4, with affinity increasing up to DP = 6 to a level similar to Hep. These data indicate that the BT4661$_{SusC-like}$ binding site can accommodate a hexasaccharide and supports the ability of the SGBP to tolerate/recognize sulfation and IdoA at all subsites (Table S1). These data show that BT4661$_{SusC-like}$ is an SGBP of the Hep/HS degrading apparatus and that the protein can accommodate a range of Hep-based GAG structures.

BT4659$_{SusD-like}$ bound Hep, HS, and ΔShp with similar affinity, but the Kd was ~20- to 30-fold lower than that for the BT4661$_{SusC-like}$. Differences in affinity between the SusD-like and
cognate SGBP have been reported previously and may reflect the proposed different roles that these proteins play in glycan utilization (5, 19).

**Structure of full-length BT4661**. The crystal structure of BT4661 (Table S2) revealed six discrete domains that adopt an Ig-like fold (Fig. 2B). This multi-Ig-like domain structure is common to the other three SGBPs, which have structures that are known, although little or no sequence identity is evident between these proteins (5, 19, 20). The domains, defined as D1 (N terminus) to D6 (C terminus), are arranged in an extended, curved conformation. Interdomain probe residues may limit the flexibility of the SGBP (19). Small angle X-ray scattering (SAXS) on full-length BT4661<sub>SGBP</sub> in the presence and absence of Hep gave the same R<sub>g</sub> and D<sub>max</sub> values for both conditions (Table S3), indicating that no large conformational changes are imposed by the target GAG (Fig. 2C, *SI Results*, and Fig. S3D).

**Structure of truncated BT4661<sub>SGBP</sub> bound to oligosaccharide**. Attempts to obtain a ligand complex of full-length BT4661 were unsuccessful. A truncated form of the protein comprising only D5 and D6 displayed similar ligand binding to the WT SGBP. The crystal structure of the D5/6 derivative (TrBT4661<sub>SGBP</sub>) was determined in complex with a fully sulfated, Hep-derived hexasaccharide (Δ4,5UA2S-GlcNS6S-IdoA2S-GlcN6S-IdoA2S-GlcNS6S) (Fig. 2D-F and Table S2). The binding site is formed across the D5 and D6 domains, with the reducing end of the hexasaccharide, GlcNS6S, lying in D6 (subsite 6) and the nonreducing end ΔUA2S located in D5 (subsite 6). D6 and D5 house sugars 1–4 and 5 and 6 of the hexasaccharide, respectively. The presence of the glycan binding site on the C-terminal domain of an extended multidomain SGBP has been observed previously and may be an adaptation to enable ligand binding to occur at a distance from the cell surface (19). Similar to other GAG binding proteins, interactions between ligand and protein mainly involved basic amino acids rather than aromatic residues, which is the characteristic signature of protein–carbohydrate recognition of neutral glycans (10, 19–21). Alanine scanning mutagenesis revealed that K505, R581, and R582 are the key residues involved in ligand binding (Fig. S3E). R582 and K505 interact with the carboxylates of the UA at substrates 2 and 6, respectively, whereas R581 interacts with O3 of the reducing end sugar. These interactions will be conserved in all Hep-based GAGs, providing an explanation for the ability of BT4661 to recognize both HS and Hep with similar affinity.

**Enzymes Encoded by PUL<sub>pup</sub>**. The activity of the PLs sulfatases and GH encoded by PUL<sub>pup</sub> was determined against different substrates, and the growth of mutant strains lacking active forms of these enzymes on Hep, HS, and ΔHep was assessed (Fig. S1 F–I).

**Surface lyase BT4662<sub>PL12</sub>**. The lipoprotein BT4662<sub>PL12</sub> was shown to be a surface enzyme (Fig. S3B) that was active on Hep, ΔHep, and HS, with a preference for the latter substrate (Fig. 3A and Table S4). Aerobic whole-cell assays, which report only on the activity of surface enzymes, showed that the pattern of reaction products closely matches that of recombinant BT4662<sub>PL12</sub> (Fig. S4), confirming the cellular location of the PL. The range of oligosaccharides generated during the early stages of GAG degradation points to an endo-activity (Fig. 3A and Fig. S4E). In other characterized PULs, a surface endo-acting glycanase plays a key role in polysaccharide utilization by generating oligosaccharides that are of an appropriate size to be transported by the SusC/D-like apparatus (5). However, the Δbt4662<sub>PL12</sub> Bi mutant displayed only a partial growth defect on HS (Fig. S1B), likely reflecting the heterogeneity of the DP of the GAG; thus, the smaller forms of this glycan were able to be transported into the periplasm of *Bt* without the need for prior depolymerization. Growth on Hep and ΔHep was not impaired in the Δbt4662<sub>PL12</sub> mutant, reflecting their low DP, which likely enables these molecules to be transported into the periplasm without prior enzymatic processing.

End point assays revealed that Hep was almost completely inaccessible to BT4662<sub>PL12</sub> (only ~5% degradation), whereas the depolymerization values of HS and ΔHep were ~40 and ~80%, respectively (Table S5). Against HS, the limit products comprised a wide range of products, whereas complete degradation of ΔHep...
generated oligosaccharides with a DP of 2–10 (Fig. 3A and Fig. S4E). BT4662PL12 was only active against Hep oligosaccharides with a DP \( \geq 10 \) (Table S4). These data suggest BT4662PL12 has a large substrate binding cleft, in which sulfate groups can only be accommodated at specific positions, and the enzyme displays an endo mode of action, with limited processivity. These data are consistent with a role for the enzyme in depolymerizing high molecular weight GAGs, such as HS, at the bacterial surface.

**Periplasmic PLs.** Based on proteinase K treatment and the product profiles of whole-cell assays compared with those of the recombinant enzymes, the majority of the polysaccharide-degrading enzymes encoded by PULHep are located in the periplasm (Figs. S3B and S4A) and display no evidence for metal dependence. A unifying feature of the three periplasmic PLs, BT4652PL15, BT4657PL12, and BT4675PL13, is the appearance of limit disaccharide products during the initial phase of degradation (Fig. 3B–D). This activity is indicative of a degree of processivity and likely results in rapid production of the disaccharides that act as the signaling cue(s) required to up-regulate PULHep (17). The three periplasmic lyases have differing preferences for Hep, HS, and \( \Delta \)Hep (described in detail below), reflecting the ability of Bt to use these three substrates.

BT4657PL12 exhibited the greatest activity against HS and could access \( \sim 30\% \) of the polymer, while against Hep, the lyase cleaved \( \sim 27\% \) of the glycosidic bonds (Tables S4 and S5). The enzyme completely degraded \( \Delta \)Hep to a single unsulfated disaccharide species (Fig. 3B and Tables S4 and S5). During the initial degradation of \( \Delta \)Hep, BT4657PL12 generated a range of oligosaccharides, with the disaccharide being the dominant product (Fig. 3B and Fig. S4E). These data indicate that BT4657PL12 is primarily endo-acting, with a degree of processivity. The products generated from partially sulfated oligosaccharides showed that BT4657PL12 is unable to accommodate 2-O sulfation of UA in its active site (+1 subsite), explaining why the enzyme displayed limited activity against Hep (SI Results).

In contrast, BT4675PL13 displayed highest activity against Hep and was inactive against \( \Delta \)Hep, consistent with a previous study suggesting that sulfation was required for activity (Table S4) (22). The requirement for sulfation was mirrored in the end point assays, which showed almost complete digestion of Hep but only \( \sim 50\% \) of HS (Table S5). The PL13 enzyme generated a range of different size products indicative of an endo mode of action but also showed limited processivity, because UA2S-GlcNS6S was the dominant product early in the reaction (Fig. 3C).

PL15 enzymes have previously only been shown to be alginate lyases (23). BT4652PL15, however, was active on all three GAGs tested. The enzyme was \( \sim 800\text{-fold} \) more active against fully sulfated Hep oligosaccharides compared with \( \Delta \)Hep, indicating that sulfation was key for optimal activity (Table S4), consistent with the production of UA2S-GlcNS6S, UA2S-GlcNS, and UA-GlcNS6S from Hep; however, against HS, an additional disaccharide, UA-GlcNAc6S, was generated, and the unsulfated disaccharide UA-GlcNAc was produced only from \( \Delta \)Hep (Fig. 3D). BT4652PL15 generated disaccharides against all substrates; no intermediate products were observed during initial degradation (Fig. 3D and Fig. S4D). These data indicate the enzyme is exclusively exo-processed and can accommodate sulfates at all positions of the substrate. Although sulfation is not essential, it is a specificity determinant, and as such, BT4652PL15 displays more substrate flexibility than BT4675PL13.

The data described above suggest that BT4652PL15 and BT4662PL13 can target similar heavily O-sulfated regions of Hep and HS, although the PL15 enzyme can, in addition, cleave sparsely sulfated regions of these GAGs. Consistent with this view is the observation that mutant strains lacking BT4652PL15 or BT4675PL13 displayed growth defects on Hep and to a lesser extent, HS but not \( \Delta \)Hep (Fig. S1B–E). The growth properties of \( \Delta \)hbt4652PL12 revealed a significant role for BT4657PL12 in HS degradation. The endo-acting lyase likely cleaves areas low in sulfation, creating nonreducing ends that are targeted by BT4652PL15 and BT4675PL13. Our data thus provide a biological context for the multiple lyases expressed by Bt in response to Hep and HS (Discussion). The primary substrates for BT4657PL13 and BT4657PL12 are regions of the GAGs that were highly and poorly sulfated, respectively. The \( \Delta \)hbt4652PL15 strain had the largest growth defect on Hep, whereas \( \Delta \)hbt4657PL12 displayed little growth on HS and a marked increase in the lag phase when cultured on \( \Delta \)Hep (Fig. S1D). These growth profiles suggest that BT4657PL12 and BT4675PL13 produce oligosaccharides that only BT4657PL13 can degrade, and thus, the additional flexibility in recognition by the PL15 bridges the two complementary activities of the PL12 and PL13 enzymes (discussed in detail below).

**GH88 enzyme.** BT4659GH88 belongs to a family of enzymes that cleave the glycosidic linkage between the A4,5-un sulfated UA and Gln/GlcNAc disaccharides that are produced by GAG lyases. The enzyme was not active when O2 of the UA was sulfated but could accommodate sulfation at N2 or O6 of the neutral sugar, mirroring the binding specificity of the PULHep hybrid two-component system (HTCS) BT4663 (Table S6) (17). The tuning of the specificity of the GH88 enzyme with that of its cognate HTCS is analogous to that observed in the Bt chondroitin sulfate PUL, indicating a conserved role for these enzymes in controlling the rate of signal degradation during growth on GAGs (24). Although BT4658GH88 displayed similar catalytic efficiencies against GlnNS and GlcNAc, the \( k_m \) for the sulfated sugar were lower, suggesting that the N-linked sulfate contributes to substrate binding but restricts product departure, leading to a reduced \( k_{cat} \).

### 0-sulfatases

A previous study showed that BT1596S-sulf and BT4656S-sulf are exo-acting O-sulfatases; BT1596S-sulf targets the O2 sulfation of unsaturated UA at the nonreducing end of di- and oligosaccharides, whereas BT4656S-sulf cleaves the O6-sulfate of the monosaccharide GlnAc6S or GlcNS6S but is inactive against GalNAc6S (Table S6) (15). To explore the mechanism of substrate recognition, the crystal structures of inactive forms of the two Bt O-sulfatases were determined in complex with their substrates. BT1596S-sulf and BT4656S-sulf share a conserved \( \alpha/\beta \) hydrolase fold comprising a single domain (Fig. 4, Upper and Table S2). Both enzymes have a pocket topology with metal binding sites located at the base of the pocket. Calcium was modeled into the BT4656S-sulf ligand-bound structure, whereas no metal was modeled in BT1596S-sulf (Fig. 4, Lower). Alanine substitution of the residues in this pocket inactivated both enzymes, revealing the active site location (SI Results and Table S6).

Inactive forms of BT1596S-sulf and BT4656S-sulf were crystallized with \( \Delta \)UA2S-GlcNS6S and GlcNS6S, respectively (Fig. 4). BT1596S-sulf and BT4656S-sulf were inactive when expressed in *Escherichia coli*, because the enteric bacterium cannot convert S64 and S77, respectively, into formylglycine, which functions as the catalytic nucleophile [to generate the formylglycine in E. coli, the catalytic serine was mutated to cysteine because the bacterium can convert cysteine to formylglycine (25)]. The essential histidines H188 in BT1596S-sulf and H196 in BT4656S-sulf are in close proximity to the scissile bond (Fig. 4, Lower and Table S6). Thus, H188 and H196 are ideal candidates for the catalytic acid that is required to protonate the O2/O6 of the departing sugar after nucleophilic attack by the formylglycine and formation of a sulfatate–enzyme intermediate, which is then cleaved via a \( \beta \)-elimination to complete the catalytic cycle.

In BT1596S-sulf, the carboxylate of \( \Delta \)UA2S makes a bidentate ionic interaction with R237, and loss of this interaction (R237A) results in a 2,000-fold reduction in catalytic efficiency (Fig. 4A and Table S6). This interaction is likely the major specificity determinant used by the enzyme and explains why the sulfate acts on chondroitin disaccharides, which also possess O2-sulfated UA. In BT4656S-sulf, R363 and D361 coordinate O4, and mutation of R363 to Ala caused an \( \sim 500\text{-fold} \) decrease in activity, whereas
D361A was inactive (Table S6). These amino acids are likely specificity determinants for gluco- over galacto-configured substrates. (SI Results) A description of all mutants.

**N-sulfatase.** There are no predicted sulfamidases in PUL\(_{Hep}\) however, GlcNS does not accumulate during growth on Hep, indicating that \(Bt\) uses this sugar (Fig. S4F). We deleted the remaining uncharacterized ORF in PUL\(_{Hep}\) \(bt4655\). This mutant displayed a growth defect on sulfated GAGs, and GlcNS accumulated in the media during growth on Hep (Figs. S1F and S4F). These data indicate that BT4655 cleaves the sulfamate linkage and is a previously uncharacterized class of sulfatase. Unfortunately, the proposed role of BT4655 as a sulfamidase could not be confirmed biochemically, because a soluble recombinant form of the protein could not be generated.

The data show that the three sulfatasases are key for both Hep and HS breakdown. BT1596\(_{2S-sulf}\) is the first to act, removing 2-O sulfation from the limit PL-derived disaccharide and allowing BT4656\(_{6S-sulf}\) to hydrolyze its target linkage, because this enzyme does not use 2-O-sulfated substrates. The sulfated glucosamine monosaccharides are then substrates for BT4656\(_{6S-sulf}\) and the likely sulfamidase BT4655 (Table S6).

**Cytoplasmic Sugar Kinase.** BT4654\(_{ROK}\) is a cytoplasmic member of the ROK protein family that possesses the conserved DxGxT motif and thus, likely an ATP-dependent kinase (26). This activity was confirmed by the capacity of the enzyme to phosphorylate gluco- and manno-configured sugars but not galacto-configured sugars (Table S6). The \(k_{cat}/K_m\) of the enzyme decreased from GlcN > GlcNS > Glc/Man >> GlcNS, which was driven by increases in \(K_m\) (Table S6). These data show that O4 is critical for catalysis, whereas strong selection in affinity is made at C2 for an equatorial amine, indicating that BT4654\(_{ROK}\) uses both an equatorial N2 and an O4 as specificity determinants. Because the \(\Delta bt4654\(_{ROK}\) mutant displayed no growth defect on Hep or HS (Fig. S1F), \(Bt\) seems to display redundancy in GlcN and GlcNAc phosphorylation.

**Discussion**

\(Bt\) PUL\(_{Hep}\) orchestrates the hierarchical degradation of Hep and HS. Both of these GAGs could be derived from dietary meat or host glycans. In vivo expression of PUL\(_{Hep}\) in mice colonized with \(Bt\) has only been observed in bacteria occupying the mucosal layer, suggesting that the source of Hep/HS available to \(Bt\) is from the turnover of epithelial cells rather than the diet (27). HS could also be supplied in the human diet as a component of meat; however, colonic metagenomic data obtained in a study, which swapped subjects from vegetarian diets to meat-rich diets, did not reveal any PL families associated with the degradation of Hep/HS (28). These data support the hypothesis that the source of Hep/HS in the human gut accessible to the microbiota is from the turnover of epithelial cells. Goodman et al. (29) identified through transposon mutagenesis genes in \(Bt\) important for colonization of the mouse gut. Strains with insertions in the \(bt4655\(_{ROK}\) and \(bt4656\(_{SGBP}\) insertion mutants were much lower in abundance than in monocolonized mice, indicating that the ability to degrade HS is under increased selection pressure for \(Bt\) in the presence of other Bacteroidetes (Fig. S2).

Within the Hep-based GAGs, there is substantial structural diversity, especially in terms of the level of sulfation, and it is this heterogeneity that provides an explanation for the distinct but complementary activities of the enzymes encoded by the PUL\(_{Hep}\). A key feature of the PUL\(_{Hep}\) degradative apparatus is the surface lyase, BT4662\(_{PL12}\), which although dispensable for growth on Hep, plays a significant role in optimal HS utilization. This difference likely reflects the high DP of HS relative to Hep, and thus, the GAG needs to be subjected to a degree of extracellular de-polymerization to generate molecules that can be imported into the periplasm (5).

In the periplasm, the substrate specificity of endoprocessive lysates is optimized to target highly sulfated (BT4675\(_{PL13}\) or low/sulfated regions (BT4657\(_{PL11}\)), producing small oligosaccharides that only the exoprocessive lyase (BT4652\(_{PL13}\)) is able to degrade. This role for BT4652\(_{PL13}\) is critical, because loss of the enzyme means that elements of both Hep and HS remain inaccessible to \(Bt\). The presence of a PL15 lyase is a conserved feature of Hep/HS utilization in other Bacteroidetes, supporting its key role in breakdown of these GAGs (Fig. S2). The synergistic specificities displayed by the PUL\(_{Hep}\) lysases enable the bacterium to cleave the backbone of GAGs that display substantial structural variation, particularly in their sulfation patterns, leading to the generation of disaccharides. The disaccharides are broken down by the complementary activities of the sulfatasases and the GH88 enzyme-generating sugars that can then be metabolized by the bacterium.

\(Bt\) places a high priority on Hep and HS utilization shown by the lack of repression of PUL\(_{Hep}\) by other polysaccharides and glucose (30). HS is an abundant source of GlcNAc, which is required for the synthesis of peptidoglycan. Although \(Bt\) seems to contain all of the biosynthetic genes needed to make GlcNAc, it may still prioritize GlcNAc derived from glycans to synthesize its cell wall. As stated above, we propose that \(Bt\) uses HS/Hep from the host mucosa rather than the diet. Because the epithelium is likely to contain much higher amounts of HS compared with Hep, we believe that the biologically relevant glycan targeted by PUL\(_{Hep}\) is HS. Analysis of the genomes of other members of the CM revealed PULs similar to PUL\(_{Hep}\) (Fig. S2). Bacteroides xylanisolvens and Bacteroides ovatus both contain a predicted surface PL12 lyase similar to BT4662\(_{PL12}\), suggesting that these organisms also target HS. In contrast, the GAG-degrading apparatus of Bacteroides intestinalis lacks an obvious surface Hep/HS lyase (Fig. S2). Thus, the bacterium may target oligosaccharides or low molecular weight
GAGs, such as Hep, which can be imported without the need for enzymatic depolymerization.

The structural data of the sulfatases showed that the interactions with the target sulfates are highly conserved in BT1596 [4FDI], BT4656 [4FDI], and the human 6S GalNAc sulfa (GALSase) (Fig. S3F). In BT4656 [4FDI] and GALSase (the only other 6-O PL activity was monitored at VPI-5482 was cultured anaerobically, and genomic null mutations were generated as described previously (16, 32).

**Enzyme Assays.** PL activity was monitored at A232, BT4658[4FDI] glucuronol glycosul hydrolase activity was determined by monitoring loss of signal at A295. Sulfatase and kinase activities were determined through either linked assays or HPLC. Products were analyzed by TLC and/or HPLC.

**Crystruction of BT1596, BT4656, and BT4661.** Crystruction and structure determination were as described in SI Materials and Methods.

**Acknowledgments.** We thank Carl Morland for expert technical assistance. This work was funded by the United Kingdom Biotechnology and Biological Sciences Research Council (Grant BB/F011463/1), the European Research Council (Grant 322820), and the Wellcome Trust (Grant WT097970/M/A).

**Materials and Methods**

**Bacteroides Culture and Genetic Manipulation.** BT VPI-5482 was cultured anaerobically, and genomic null mutations were generated as described previously (16, 32).

**Enzyme Assays.** PL activity was monitored at A232, BT4658[4FDI] glucuronol glycosul hydrolase activity was determined by monitoring loss of signal at A295. Sulfatase and kinase activities were determined through either linked assays or HPLC. Products were analyzed by TLC and/or HPLC.

**Crystallography of BT1596, BT4656, and BT4661.** Crystallography and structure determination were as described in SI Materials and Methods.

**Full details of all experimental procedures used are described in SI Materials and Methods.**