Sulfatases and a Radical S-Adenosyl-L-methionine (AdoMet) Enzyme Are Key for Mucosal Foraging and Fitness of the Prominent Human Gut Symbiont, *Bacteroides thetaiotaomicron*^{*S}

Received for publication, February 10, 2011, and in revised form, April 19, 2011 Published, JBC Papers in Press, April 20, 2011, DOI 10.1074/jbc.M111.228841

Alhosna Benjdia^{‡1}, Eric C. Martens^{§¶}, Jeffrey I. Gordon[§], and Olivier Berteau^{‡2}

From the [‡]INRA, UMR1319 Micalis, F-78352 Jouy-en-Josas, France , the [§]Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, Missouri 63108, and the [¶]Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, Michigan 48109

The large-scale application of genomic and metagenomic sequencing technologies has yielded a number of insights about the metabolic potential of symbiotic human gut microbes. Nevertheless, the molecular basis of the interactions between commensal bacteria and their host remained to be investigated. Bacteria colonizing the mucosal layer that overlies the gut epithelium are exposed to highly sulfated glycans (i.e. mucin and glycosaminoglycans). These polymers can serve as potential nutrient sources, but their high sulfate content usually prevents their degradation. Commensal bacteria such as Bacteroides thetaiotaomicron possess more predicted sulfatase genes than in the human genome, the physiological functions of which are largely unknown. To be active, sulfatases must undergo a critical post-translational modification catalyzed in anaerobic bacteria by the radical AdoMet enzyme anaerobic sulfatase-maturating enzyme (anSME). In the present study, we have tested the role of this pathway in Bacteroides thetaiotaomicron which, in addition to 28 predicted sulfatases, possesses a single predicted anSME. In vitro studies revealed that deletion of the gene encoding its anSME (BT0238) results in loss of sulfatase activity and impaired ability to use sulfated polysaccharides as carbon sources. Co-colonization of formerly germ-free mice with both isogenic strains (*i.e.* wild-type or $\Delta anSME$), or invasion experiments involving introduction of one followed by the other strain established that anSME activity and the sulfatases activated via this pathway, are important fitness factors for B. thetaiotaomicron, especially when mice are fed a simple sugar diet that requires this saccharolytic bacterium to adaptively forage on host glycans as nutrients. Whole genome transcriptional profiling of wild-type and the anSME mutant in vivo revealed that loss of this enzyme alters expression of genes involved in mucin utilization and that this disrupted ability to access mucosal glycans likely underlies the observed pronounced colonization defect. Comparative genomic analysis reveals that 100% of 46 fully



sequenced human gut Bacteroidetes contain homologs of *BT0238* and genes encoding sulfatases, suggesting that this is an important and evolutionarily conserved feature for bacterial adaptation to life in this habitat.

The human gastrointestinal tract is inhabited by a large population of microbes belonging to all three domains of life and their viruses. The microbiota is dominated by members of bacteria, principally, members of the Firmicutes and Bacteroidetes phyla (1–3). Metagenomic studies of fecal microbial communities have indicated that in a healthy adult, this microbiota is composed of several hundred species-level phylogenetic types (phylotypes), although there is considerable interpersonal variation in species composition, even among monozygotic twins (2, 3). Nonetheless, shotgun sequencing of fecal community DNA has shown that these different species assemblages contain shared functional features, including the ability to process otherwise indigestible components of our varied diets such as plant polysaccharides (1-4).

If the dynamic operations and the potential to deliberately manipulate the functional properties of this very complex microbial ecosystem are to be fully deciphered, the mechanisms by which bacteria colonize and persist in our digestive systems need to be defined at a molecular level. One approach is to create model communities, composed of one or more sequenced members of the microbiota, in gnotobiotic mice (5–7). An organism whose properties have been studied in this fashion is Bacteroides thetaiotaomicron, a prominent species in the adult human gut microbiota (1). Its genome is enriched for carbohydrate active enzymes (CAZymes) involved in metabolic processing of complex carbohydrates, including those present in various glycoproteins that comprise the mucus layer overlying the gut epithelium. The ability to forage for both dietary and host-derived glycans appears to be among the mechanisms by which B. thetaiotaomicron is able to survive in this fiercely competitive ecosystem (8, 9). Some mucosal glycans are highly sulfated (e.g. colonic mucins and glycosaminoglycans), an observation that suggests that the capacity to produce sulfatases is instrumental in allowing some bacterial species to process host-derived carbohydrates. Furthermore, comparative genomic analyses have shown that sulfatases are prevalent

^{*} This work was supported in part by grants from Agence Nationale de la Recherche (ANR-08-BLAN-0224-02), ANSSD-2008 from INRA department AlimH, the National Institutes of Health (DK30292 and DK084214), and the Crohn's and Colitis Foundation of America.

The on-line version of this article (available at http://www.jbc.org) contains supplemental Tables S1–S6 and Fig. S1.

¹ Recipient of a PhD grant from the Région Ile-de-France.

² To whom correspondence should be addressed: INRA, UMR1319 Micalis, F-78352 Jouy-en-Josas, France. E-mail: Olivier.Berteau@jouy.inra.fr.

in other commensal Bacteroidetes (10), leading us to hypothesize that sulfatases are critical, evolved fitness factors.

Among hydrolytic enzymes, a unique feature of sulfatases is their requirement for a 3-oxoalanine moiety, typically called C α -formylglycine (11), for catalysis. This C α -formylglycine residue is formed by the post-translational modification of a critical active site cysteinyl or serinyl residue. This oxidation reaction is catalyzed by distinct enzymatic systems in microbes, two of which have been characterized (11). The first, formylglycine-generating enzyme, catalyzes conversion of a cysteine to C α -formylglycine and requires molecular oxygen as a cofactor (12, 13). The second, anaerobic sulfatase-maturating enzyme (anSME),³ is a member of the superfamily of radical S-adenosyl-L-methionine (AdoMet)-dependent enzymes also called radical SAM enzymes (11, 14-17). As we reported recently, this enzyme is likely the only bacterial enzyme able to activate sulfatases under anaerobic conditions (11). We further elucidated the molecular mechanism of anSME (14-17). Interestingly, metagenomic sequencing projects have revealed that anSMEs are enriched in the gut microbiomes of humans compared with non-gut microbial communities (18), further suggesting a key role for sulfatases in colonization of the human gastrointestinal tract. We directly test this notion in the present study using B. thetaiotaomicron. Our approach involved a combination of competitive colonization and invasion assays of wild-type and anSME-deficient strains in gnotobiotic mice fed polysaccharide-rich and -deficient diets, whole genome transcriptional profiling of the strains in their distal gut habitat, and in vitro tests of their sulfatase activities and growth under defined nutrient conditions.

MATERIALS AND METHODS

Bacterial Strains and Culture Conditions-Bacterial strains and plasmids are summarized in supplemental Table S1. The B. thetaiotaomicron type strain VPI-5482 was grown on brainheart infusion (BHI; Beckton Dickinson) agar supplemented with 10% horse blood (Colorado Serum Co.) for genetic experiments. The following antibiotics were added as required: erythromycin (25 μ g ml⁻¹), gentamicin (200 μ g ml⁻¹), and 5-fluoro-2-deoxyuridine (200 μg ml⁻¹). Minimal medium (MM) consisted of 100 mM KH₂PO₄ (pH 7.2), 15 mM NaCl, 8.5 m_{M} (NH₄)₂SO₄, 4 m_M L-cysteine, and 10 mg hemin (prepared as а 1000 \times stock solution in 0.5 м NaOH), 100 μ м MgCl₂, 1.4 μ м ${\rm FeCl}_3,$ and 50 $\mu{\rm M}~{\rm CaCl}_2,$ 1 $\mu{\rm g}~{\rm ml}^{-1}$ vitamin ${\rm K}_3$ and 5 ${\rm ng}~{\rm ml}^{-1}$ vitamin B₁₂. Heparin, chondroitin sulfate from shark cartilage, and mucin from porcine stomach (type III) were purchased from Sigma-Aldrich. All carbon sources were added to the MM at a final concentration of 0.5% (w/v). Media were filter-sterilized using a Millipore Express filter unit (0.22-µm pore diameter). Bacteria were grown at 37 °C in an anaerobic chamber (Bactron IV) under an atmosphere of nitrogen (90%), carbon dioxide (5%), and hydrogen (5%).

A *B. thetaiotaomicron* mutant lacking the single putative anSME gene (*BT0238*) was constructed using a previously

described allelic-exchange method (8). An inactivation cassette was constructed with the primers listed in supplemental Table S2.

Whole Genome Transcriptional Profiling-Cecal contents were collected after being sacrificed and frozen in liquid nitrogen. Transcriptional profiling was performed using custom Affymetrix GeneChip containing probe sets representing >98% of 4,779 predicted protein-coding genes in the B. thetaiotaomicron genome (9). GeneChip targets were prepared from total RNA recovered from cecal contents as described previously (19) and hybridized to the microarrays according to standard Affymetrix protocol. All GeneChip assays were performed in triplicate. Data were normalized using Microarray Suite 5 software (Affymetrix) by adjusting the average B. thetaiotaomicron transcript signal on each GeneChip to an arbitrary value of 500. Subsequent comparisons of GeneChip datasets were performed using Gene-SpringGX software (version 7.3.1, Agilent) and the following workflow: 1) raw intensity values <1.0 were adjusted to 1 prior to calculating fold-differences; 2) genes with average fold differences in expression of ≥ 2 between isogenic wild-type and $\Delta anSME$ strains were identified; and 3) the list of genes exhibiting \geq 2-fold changes was restricted to include only those for which the fold difference had a *p* value < 0.01 (Student's *t* test), a "present" call in all three GeneChips with elevated expression, and an intensity value in the higher expression state of >100. All of the GeneChip data used in this study are available from the Gene Expression Omnibus (GEO) database (www.ncbi. nlm.nih.gov/projects/geo/) under accession no. GSE25415.

Competitive Colonization of Germ-free Mice—All protocols using mice were approved by our Institution's Animal Studies Committee. C57Bl/6J mice were reared in Trexler-type isolators (La Calhène, Vèlizy-Villacoublay, France). Animals were fed ad libitum with a commercial diet sterilized by gamma irradiation (4 kGy) and supplied with sterilized (20 min, 120 °C) tap water. Room temperature was 21 °C, and a photoperiod of 12 h was used. The germ-free status of the animals was verified routinely. Eight-week-old male animals were gavaged once with a suspension of 100 μ l of either the wild-type strain, an isogenic mutant strain, or a mixture containing equal proportions of wild-type and mutant cells (bacteria harvested from mid-log phase cells grown in TYG medium; $\sim^{10} e^8$ cfu of each strain gavaged/animal; n = 3 mice/strain/experiment). One group of mice was maintained for 12 days on the chow diet prior to and after gavage. Another group was switched to a diet that contained simple sugars rather than complex carbohydrates 12 days prior to gavage. This diet, obtained from Bio-Serv and sterilized by irradiation before administration, contained 35% (w/w) sucrose, 35% glucose, and 20% protein (see Ref. 9 and supplemental Table S3 for other components).

Fecal samples were collected from each animal in the days following gavage and immediately frozen after collection at -80 °C until further use. Total DNA was isolated from fecal pellets using bead-beating extraction and was further purified with a GNOME DNA isolation Kit (MP Biomedicals) (20). The representation of each strain in the fecal microbiota was defined by qPCR in 25-µl reactions that contained 200 nM of each primer (see supplemental Table S2), plus 12.5 µl of Master

asemb



³ The abbreviations used are: anSME, anaerobic sulfatase-maturating enzyme; BHI, brain-heart infusion; MM, minimal medium; qPCR, quantitative PCR; PUL, polysaccharide utilization loci.

Mix SYBR Green (Applied Biosystems) and 10 μ l of diluted DNA. PCR was performed with an ABI Prism 7000 instrument and the following cycling conditions: 95 °C for 10 min, 40 cycles of 95 °C for 15 s, and 60 °C for 1 min followed by a dissociation step of 95 °C for 15 s, 60 °C for 20 s, and 95 °C for 15 s. Purified genomic DNA standards (100 – 0.01 ng), prepared from each strain, were included in each qPCR run so that a standard curve could be constructed and used to calculate the relative representation of each strain in each fecal sample. Results were expressed in cfu using a calibration curve between cfu and DNA amounts. All qPCR assays were performed in triplicate for each sample at two DNA dilutions.

Sulfatase Assay—Protein extracts were prepared from *B. thetaiotaomicron* by centrifuging (10,000 × *g* for 10 min) 20-ml aliquots of a bacterial culture grown under anaerobic conditions to mid-log phase in the different media investigated. Bacterial pellets were resuspended in 5 ml of a solution containing 50 mM Tris, 100 mM KCl, 10 mM MgCl₂, 5% glycerol (pH 7.25), and the mixture was centrifuged (10,000 × *g* for 10 min). 500 μ l of a buffered solution consisting of 50 mM Tris, 100 mM KCl, 10 mM MgCl₂ (pH 7.5), and a protease inhibitor mixture (Complete EDTA-free, Roche Applied Science) were added to the bacterial pellets, and the cells were disrupted at 4 °C with an ultrasonic device (VibraCell 72434, Bioblock Scientific). After sonication (30 s), the protein concentration was assayed with the Bradford reagent (Sigma).

Sulfatase activity was measured at 25 °C in a 600- μ l reaction mixture containing 10 mM of the chromogenic substrate *p*-nitrophenyl sulfate in 50 mM Tris, 100 mM KCl, 10 mM MgCl₂ (pH 7.25), and 125 μ g of extracted *B. thetaiotaomicron* proteins. One unit of activity was defined as the release 1 μ mol of product per minute per milligram of protein extract (monitored at 405 nm with a Beckman DU-640 spectrophotometer). All assays were performed in triplicate.

RESULTS

Identification of Putative Sulfatases in Bacteroides Thetaiotaomicron Genome and a Single anSME-Based on the presence of a sulfatase-associated PFAM domain PF00884 and the canonical N-terminal sulfatase signature C/SxP/AxR (21), we identified 28 putative sulfatase genes in the genome of the human gut-derived B. thetaiotaomicron type strain VPI-5481 (BLASTP E-value $<10^{-30}$). This number of genes is almost twice the number present in our human genome. Notably, 20 of the predicted sulfatases (71%) have secretion signals, and 21 (75%) are encoded by genes present in 10 different polysaccharide utilization loci (PULs). B. thetaiotaomicron VPI-5482 has 88 PULs representing 866 genes (18.1% of its genome) that encode proteins critical for the sensing, binding, import, and degradation of various classes of polysaccharides (see Ref. 6 for a review of the features that define PULs). In contrast to the numerous sulfatases found in this bacterium, only one gene (BT0238) encoding an anSME was detected based on a search of all known sulfatase-maturating enzymes (anSME or formylglycine-generating enzyme) present in the B. thetaiotaomicron genome (BLASTP E-value $< 10^{-10}$).

We recently demonstrated that *BT0238* encodes indeed an authentic anSME (14, 16). In other organisms, such as *Esche*-

richia coli K12 and *Vibrio* species, an *anSME* is commonly present in an operon that also contains a co-expressed sulfatase. The finding that *B. thetaiotaomicron* contains only one identifiable anSME gene raised the possibility that its activity is broad and sufficient to catalyze the maturation of all of the sulfatases of the organism. We have collected over 200 GeneChip data sets that profile the transcriptome of *B. thetaiotaomicron* under a variety of *in vitro* and *in vivo* conditions; meta-analysis of these data sets revealed that *BT0238* is expressed at similar levels *in vitro* or *in vivo* independently of the growth conditions surveyed (9, 22). Therefore, we postulated that constitutive expression of this anSME would allow it to activate various PUL-associated sulfatases as they are produced in a substratedependent manner.

B. thetaiotaomicron Requires anSME for Optimal Fitness in Gut—First, we constructed a $\Delta anSME$ B. thetaiotaomicron strain with previously reported procedures (Fig. 1A) (8). As shown, *in vitro*, the growth of the mutated strains was not affected compared with the wild-type strain (Fig. 1B).

To determine whether loss of *BT0238* and its corresponding anSME activity alters the ability of *B. thetaiotaomicron* to compete in the mammalian gut, we introduced the $\Delta anSME$ mutant with or without its isogenic wild-type parent into the intestines of 8-week-old germ-free male C57BL/6J mice.

Each bacterial strain was inoculated alone with a single gavage, or as a 1:1 mixture of both strains ($\sim 2 \times {}^{10}e^8$ cfu total per animal) into mice adapted previously to either a plant gly-can-rich diet or a simple sugar diet (8) where carbohydrates are represented in the form of glucose and sucrose (35% (w/w each)). Our previous work had shown that little dietary carbon reaches the distal gut in mice consuming the latter diet and that *B. thetaiotaomicron* is forced to rely on host mucosal glycans to sustain itself (8, 9).

qPCR assays of fecal DNA, targeting the *BT0238* gene in the wild-type strain or upstream and downstream regions in the mutated strain (Fig. 1 and supplemental Table S2), were used to define the relative fitness of these strains as a function of time after gavage and host diet. When introduced alone into adult germ-free mice, each strain rapidly achieved a stable level of colonization. Moreover, no significant differences in fecal concentrations were noted between these strains regardless of the host diet (Fig. 1, C and D). A different situation arose after mice were gavaged with a 1:1 mixture of both strains. Initially, during the first 5 days after gavage, both strains were present in feces at levels comparable with those documented in mice inoculated with just a single strain ("mono-associated"). This was true for either diet. However, after 5 days, the representation of the $\Delta anSME$ strain decreased progressively, with the difference becoming statistically significant after 8 days; by 11 days, the $\Delta anSME$ strain represented only 6.8 \pm 4% of the total bacterial population in hosts consuming the plant glycan-rich diet and $<1.8 \pm 0.7\%$ on the simple sugar diet (Fig. 2). When competition experiments were performed in vitro in medium that lacks sulfated glycans, a totally different situation arose; when overnight cultures were subcultured daily into fresh complex BHI medium at a 1:100 dilution, after 11 days, both strains remained at very similar levels (no statistically significant differences; see Fig. 2G). Taken together, these results demonstrate that loss of





FIGURE 1. **Comparison of growth properties of wild-type and isogenic** $\Delta anSME$ strains of *B. thetaiotaomicron* VPI-5482 *in vitro* and *in vivo. A*, PCR amplification of the -300 and +300 bp region surrounding the *B. thetaiotaomicron* anSME gene in WT and $\Delta anSME$ strains (Δ). *B*, growth curves of wild-type (\blacksquare) and the $\Delta anSME$ (\diamond) cells in BHI medium. Monocolonization of 8-week-old male C57BL/6J germ-free mice with the wild-type strain (each mouse represented by a different closed symbol; \blacksquare , \blacklozenge , or \blacktriangle) or the isogenic $\Delta anSME$ strain (\square , \diamondsuit , \triangle). The *dashed blue line* represents the average colonization levels achieved with the wild-type strain, whereas the *dashed red line* represents results obtained with the isogenic $\Delta anSME$ strain. Mice were fed a standard plant glycan-rich diet (C) or a diet containing simple sugars (glucose/sucrose) (D). The representation of each strain at the indicated times following gavage was determined by the qPCR assay.

the radical AdoMet enzyme anSME (15, 16) produces an *in vivo* fitness defect even when dietary glycans are present.

 Δ anSME Strain Exhibits Defects in Utilization of Sulfated Glycans and Sulfatase Deficiency in Vitro—To understand the basis for the observed *in vivo* fitness defect of $\Delta anSME$ cells, we measured growth of the isogenic wild-type and mutant strains in MM with various carbohydrates. In MM-glucose, as in complex rich BHI medium, both strains exhibited no significant differences in lag time, exponential growth rate, or growth level (Figs. 1*B* and 3*A*) (23, 24). In contrast, the $\Delta anSME$ strain had a growth defect reaching levels that were 30% of those attained by the wild-type strain in MM porcine mucin and had an average doubling time during log phase growth that was longer than for the wild-type strain (105 versus 84 min; Fig. 3B). The growth defects observed with the glycosaminoglycans, chondroitin and heparin, were even more dramatic. Indeed, whereas both carbon sources sustained efficient growth of the wild-type strain with an average growth rate (doubling time) of 84 and 130 min with chondroitin or heparin respectively, no growth was observed for the $\Delta anSME$ strain (Fig. 3, C and D). Thus, significant growth defects were measured when the $\Delta anSME$ strain was cultured with sulfated macromolecules as obligate carbon sources (Fig. 3E).

Collectively, these results support the notion that the mutant strain is unable to mature the sulfatases required for hydrolysis of the corresponding polysaccharides, thus preventing their use as carbon sources. To confirm this, we assayed sulfatase activity in both strains using the chromogenic substrate, *p*-nitrophenyl sulfate. Although likely not a substrate for all *B. thetaiotaomicron* sulfatases because we have identified no alternative sulfatase maturation enzyme in *B. thetaiotaomicron* genome, *p*-nitrophenyl sulfate hydrolysis reflects both sulfatase induction and maturation under the different conditions assayed. Protein extracts prepared from the wild-type strain grown in BHI medium or MM-glucose exhibited very low levels of sulfatase activity, whereas we were unable to detect any sulfatase activity in $\Delta anSME$ protein extracts (Fig. 4). When wild-type cells were grown on either MM containing mucin or glycosaminoglycans, sulfatase activity was between 16–26 times higher than when cells were grown in BHI or MM-glucose media (Fig. 4), suggesting that growth on these substrates stimulates sulfatase expression. Under these conditions, the mutant did not exhibit sulfatase activity, demonstrating that it cannot catalyze maturation of the expressed enzymes (12). These results support that *B. thetaiotaomicron* has no alternative pathways to activate sulfatases except through anSME.

Transcriptional Analysis of Wild-type and Δ anSME in Mouse Cecum—We used custom *B. thetaiotaomicron* GeneChips containing probe sets to >98% of the known and predicted proteincoding genes of the organism to identify the molecular pathways affected *in vivo*, on either of the two diets, by the absence of anSME and active bacterial sulfatases. RNA was prepared from cecal contents of mice mono-associated with wild-type or mutant strains 11 days after gavage. (These mice represent the last time point shown in Fig. 1, *C* and *D*; *n* = 3 mice/group; total of 12 GeneChip data sets.)

In the ceca of animals fed the plant glycan-rich diet, only three genes, including a putative chitobiase (*BT0865*) and a periplasmic β -glucosidase precursor (*BT1872*), exhibited significantly increased expression in the wild-type compared with mutant strain, whereas 14 genes, including two sulfatases (*BT0756* and *BT4631*), a β -galactosidase (*BT0757*), several reg-







FIGURE 2. **Competition experiments in C57BL/6J germ-free mice co-inoculated with both wild-type and** $\Delta anSME$ strains. Mice were fed a standard plant glycan-rich diet (*A*, time series experiment; *B*, time point analysis after 11 days; *C*, strains ratio in percentages) or a simple sugar diet (*D*, time series experiment; *E*, time point analysis after 11 days; *C*, strains ratio in percentages) or a simple sugar diet (*D*, time series experiment; *E*, time point analysis after 11 days; *C*, strains ratio in percentages) or a simple sugar diet (*D*, time series experiment; *E*, time point analysis after 11 days; *F*, strains ratio in percentages). Each mouse is represented by a different symbol (*filled symbols*, wild-type strain; *open symbols*, $\Delta anSME$ mutant strain). The *dashed blue line* indicates the average colonization levels achieved with the wild-type strain, whereas the *red line* represents the average colonization densities obtained with the $\Delta anSME$ strain. An *asterisk* indicates that the difference in levels between strains at a given time point are statistically significant (p < 0.02 Student's *t* test). *G*, *in vitro* competition experiment in BHI medium. Both wild-type (**D**) and $\Delta anSME$ (\Box) strains were mixed in a 1:1 ratio and grown in BHI medium. After daily serial passages, each strain was quantified using qPCR (see supplemental Table S2). Overnight cultures were subcultured daily into fresh medium at a 1:100 dilution. The experiment was performed in duplicate.

ulatory components (*BT0753*, *BT4643*, *BT4644*) and outer membrane proteins (*BT0754*, *BT4633*, *BT4634*), showed the opposite pattern with increased expression in the $\Delta anSME$ strain (see supplemental Table S4 for a list of genes, the fold difference in their levels of expression, and the statistical significance of the observed differences).

A more dramatic effect was observed in *B. thetaiotaomicron* cells harvested from mice fed the simple sugar diet: 26 genes exhibited increased expression in wild-type strain, whereas 29 genes showed increased expression in the $\Delta anSME$ strain (Fig. 5 and supplemental Table S4). The 26 up-regulated genes encoded notably for three glucosidases (*BT1871*, *BT1872*, and *BT3664*), two transcription factors (*BT4355* and *BT4636*), several putative outer membrane proteins (*BT3519*, *BT0866*, *BT0867*, *BT4081*, and *BT4267*) and hypothetical proteins contained in operons involved in α -mannan/host glycans metabolism (*BT2617–28*, *BT3774–85*, and *BT3786–92*).

In the mutant strain, the 29 genes up-regulated included three sulfatases (*BT0756*, *BT3796*, and *BT3799*), three glycosidases and a galactose oxidase (*BT0683*, *BT0757*, *BT4040*, and *BT4295*), and several hypothetical proteins from operons containing sulfatase genes (*BT0749*–*58*, *BT1615–22*, *BT3796–99*). Several putative outer membrane and hypothetical proteins present in PULs dedicated to host glycans or mucin metabolism

(BT2392-95, BT4038-40, BT4294-300, BT4631-36) were also induced and a new PUL likely involved in sulfated glycans metabolism (BT0752-57). Interestingly, sulfatases or sulfatases containing operons were only induced in the mutant strain underlying their *in vivo* requirement.

Ten of the 17 genes that exhibited differences in their expression between wild-type and mutant strains in mice consuming the plant glycan-rich diet were identical to genes with altered expression on the simple sugar diet. All but four (BT1507, BT4643-46) of the seven genes whose differential expression were unique to the plant glycan diet were components of PULs (Fig. 5 and supplemental Tables S4 and S5).

Thirty-seven of the 55 genes (67%) with altered expression between strains in mice consuming the simple sugar diet were contained in predicted PULs (supplemental Table S5). These PULs encode 51 glycosyl hydrolases, which represent almost one-third of the glycosylhydrolases encoded by the *B. thetaiotaomicron* genome (25).

Altogether, 19 PULs exhibited differential expression between wild-type strain and mutants (21% of all the PULs present in *B. thetaiotaomicron* genome). Among *B. thetaiotaomicron* PULs, 35 have been reported to be involved in mucin type or hosts glycans degradation either *in vivo* or *in vitro* (8) (supplemental Tables S4 and S5). The single disruption of the







FIGURE 3. Growth curves of wild-type (\blacksquare) and $\Delta anSME$ (\diamond) *B. thetaiotaomicron* strains in MM containing glucose (*A*), mucins (*B*), chondroitin (*C*), or heparin (*D*) as the sole carbon source are shown. Each experiment was performed in triplicate. Mean values \pm S.D. are shown for each time point. *E*, ratio of growth (A_{600} ratio) of wild-type (\blacksquare) *versus* $\Delta anSME$ (\Box) mutant strains in the indicated media, as assayed at early stationary phase. Mean values \pm S.D. are plotted. An *asterisk* indicates that the difference between strains is statistically significant at p < 0.01 (Student's *t* test).



FIGURE 4. Sulfatase activity in *Bacteroides thetaiotaomicron* WT and *AanSME* strains assayed on the synthetic substrate *p*-nitrophenyl sulfate. Activity is expressed in mol·min⁻¹·mg⁻¹ of protein extract. Mean values \pm S.D. are plotted. Differences between the wild-type strain grown on BHI medium or MM-glucose compared with MM-chondroitin, MM-heparin, and MM-porcine gastric mucin were statistically significant (p < 0.01, Student's *t* test). The *AanSME* strain was not assayed in MM-heparin or MM-chondroitin, which did not support growth.

anSME gene has thus a major impact on the PULs involved in host glycan metabolism with 48% of them specifically impacted. We also examined a broader set of 33 individual *B. thetaiotaomicron* loci previously implicated in mucosal glycan metabolism (8). This analysis revealed that, beyond the 19 PULs that exhibited altered *in vivo* expression between mutant and wildtype cells, only subtle expression differences in host glycanresponsive genes exist between these two strains (Fig. 5), suggesting that a significant portion of the capacity of the organism to sense host glycans remains intact in the $\Delta anSME$ mutant (26). Still, this apparently subtle disruption in host glycan metabolism due to *anSME* loss is associated with a substantial loss of competitive fitness *in vivo*.

This latter notion was confirmed by a follow-up experiment designed to further assess the extent of the fitness defect exhibited by the anSME-deficient strain. First, mice fed the standard plant glycan-rich diet were gavaged with one or the other of two strains (10^8 cfu). Eleven days later, the second strain was introduced with a single gavage (10^8 cfu). As shown in Fig. 6*A*, when the wild-type strain was already established, the mutated strain could not invade the distal gut (*i.e.* it was not detectable in fecal samples obtained at any of the time points surveyed). In con-





FIGURE 5. Differential expression of host glycan utilization genes in wildtype and $\Delta anSME$ strains in mono-associated gnotobiotic mice consuming a simple sugar diet. The heat map shows normalized in vivo expression intensity for 33 B. thetaiotaomicron VPI-5482 gene clusters implicated in utilization of host-derived glycans. Genes contained in each operon are listed vertically in the column (operon). All operons shown, except BT3796-99 (marked with an asterisk), are components of susC/D-containing PULs. Notably, the BT3796-99 locus is associated with a hybrid two-component system regulator and encodes enzymatic functions, one glycoside hydrolase and two sulfatases, suggesting that it is functionally similar to B. thetaiotaomicron PULs but lacks the defining susC/susD homologs. Three loci with higher expression in the wild-type strain and five loci with higher expression in isogenic $\Delta anSME$ cells are separated at the top of the figure. Average fold change values for each operon are given to the right of the heat map. Negative numbers indicate lower expression in the $\Delta anSME$ mutant, and positive numbers indicate increased expression in the mutant. Fold change values for the four loci with greatest expression changes are highlighted in yellow. Intensity values are calibrated according to the color bar at the right and range from 50–5000; note that the scale is not linear.

trast, the wild-type strain was able to readily establish itself in mice that had been previously colonized with the mutant strain; beginning the first day after gavage, it increased its abundance



FIGURE 6. Invasion experiment designed to assess the relative fitness of the wild-type and mutant strains in the distal gut of gnotobiotic C57BL/6J mice fed the plant-glycan rich diet. Mice were colonized first with the wild-type strain (A) or the $\Delta anSME$ strain (B), or both strains (C) (n = 3 mice/group). After 11 days, mono-associated mice were inoculated with a single gavage of either the $\Delta anSME$ strain (*arrow* in A) or the wild-type strain (*arrow* in B). Each mouse is represented by a symbol (\blacksquare , \blacklozenge , or \blacktriangle). The *blue dashed line* represents average colonization levels attained by the wild-type strain, whereas the *dashed red line* represents average colonization levels observed with the $\Delta anSME$ strain. Data obtained by qPCR assay from the wild-type strain is shown as *open symbols*. Asterisk, difference between strains is statistically significant at a p < 0.02 (Student's *t* test).

reaching a "steady state" level 9 days later, albeit one that was still below that of the preexisting $\Delta anSME$ mutant after 33 days (Fig. 6*B*). If both strains were inoculated at the same time (Fig. 6*C*), after the first exclusion period (*i.e.* the first 11 days), the mutant population remained stable at 3–6% of the level attained by the wild-type strain. Thus, despite the fact that the $\Delta anSME$ mutant had a fitness defect in either the direct competition experiment (Fig. 6*C*) or in the invasion experiment (Fig. 6*A*), once established (Fig. 6*B*), it was never fully lost from the population, suggesting that different functional niches exist in the gut (*e.g.* luminal *versus* mucosal adherent; presence of



sulfated *versus* non-sulfated mucins) and that only some of these niches are selective for bacteria with sulfatase activity.

DISCUSSION

Based on the presence of a characteristic PFAM domain (PF00884) and the canonical N-terminal sulfatase signature C/SxP/AxR (21), we identified 28 potential sulfatase genes in the genome of a prominent saccharolytic bacterium present in the human gut microbiota. Many of these *B. thetaiotaomicron* sulfatase genes are contained in gene clusters termed PULs. Interestingly, we identified only one *B. thetaiotaomicron* gene encoding an anSME, suggesting that organismal sulfatase activity could be eliminated through deletion of this single gene, which is required for sulfatase maturation (11, 14–17, 21).

To inactivate all sulfatase activity in *B. thetaiotaomicron*, we deleted the gene encoding anSME (16). Previously called *chuR*, anSMEbt was originally identified as a regulator of *B. thetaiotaomicron* chondroitin and heparin utilization pathways (27). Our recently reported *in vitro* characterization of anSMEbt (16) and the *in vivo* results presented here establish that anSMEbt is a maturation enzyme catalyzing the essential post-translational modification required for *B. thetaiotaomicron* sulfatase activity (27).

Colonization of germ-free mice consuming a plant polysaccharide-rich or a simple sugar diet with wild-type and anSMEdeficient strains revealed that active sulfatase production by *B. thetaiotaomicron* is essential for competitive colonization of the gut, especially when the organism is forced to adaptively forage on host mucosal glycans because complex dietary polysaccharides are not available. *In vitro* biochemical assays and characterization of the organism's expressed transcriptome in the cecum confirmed that this anSME mutant is defective for sulfatase activity, cannot effectively utilize highly sulfated polysaccharides as carbon sources *in vitro*, and exhibits improper regulation of mucin *O*-glycan catabolism *in vivo*. Thus, anSME activity and subsequent activation of sulfatases represent an important pathway that allows this model Bacteroidetes species to adapt to life in the gut.

Many host-derived glycans (mucins and glycosaminoglycans) are sulfated; fewer dietary glycans from marine origin, such as carrageenan, and porphyran, which are staples in the diets of some human societies, are also heavily sulfated (28). Our previous studies of host glycan utilization by B. thetaiotaomicron in vitro and in vivo found that just two of the organisms 88 PULs are required for utilization of glycosaminoglycans (chondroitin sulfate, dermatan sulfate, hyaluronan, and heparin); these two PULs only contain three sulfatases. Neither of the glycosaminoglycan-specific PULs exhibited increased expression during growth of B. thetaiotaomicron in the distal mouse gut (cecum) compared with growth in vitro in minimal medium, suggesting that glycosaminoglycans are not heavily foraged in vivo, at least in the cecum. In contrast, B. thetaiotaomicron has at least 35 PULs for degrading host and mucin glycans, 17 of which are specifically expressed in vivo (8). Our present study, interestingly shows that the expression of 17 of these PULs (48%) is modulated when the anSME gene is inactivated.

The connection between loss of sulfatase activity in our $\Delta anSME$ mutant and the abnormal expression of some PULs in *vivo* is intriguing. One explanation for the observed alterations in PUL gene expression is that the inability to desulfate host glycans results in changes in how mucins are broken down and which oligosaccharide products are sensed. For example, if B. thetaiotaomicron is unable to remove sulfate groups, which typically occur as 6-O- and 3-O-sulfate moieties on N-acetylglucosamine and N-acetylgalactosamine (29, 30), the underlying linkages may not be sensed by the various environmental sensor/transcriptional regulators associated with PULs and thus will not induce PUL gene expression. We have shown previously that a number of the mucin O-glycan-responsive PULs in B. thetaiotaomicron are controlled by a mechanism termed trans-envelope signaling, in which dedicated outer membrane oligosaccharide transporters make protein-protein interactions with inner membrane-spanning anti- σ factors, which in turn regulate activity of extracytoplasmic function- σ transcription factors (31).

In this environmental sensing paradigm, extracellular oligosaccharides are sensed at the cell surface before they are subjected to final degradation. Thus, specific up-regulation of some PULs in the $\Delta anSME$ mutant could indicate that sulfated oligosaccharides are accumulating in the extracellular space because they cannot be fully degraded, and increased concentrations of these saccharides results in stimulation of PUL regulators that respond to them. Consistent with this idea, three of the 19 PULs that are up-regulated in the $\Delta anSME$ mutant contain sulfatases (see supplemental Table S5), suggesting that they may sense sulfated glycans and, upon induction, serve to degrade these metabolites. An alternative hypothesis to explain why some PULs are up-regulated in the $\Delta anSME$ mutant is that in the absence of the ability to desulfate normally targeted oligosaccharides, B. thetaiotaomicron turns to other, non-sulfated glycans as energy sources. Consistent with this hypothesis is the observation that some PULs up-regulated in the $\Delta anSME$ mutant do not contain sulfatases (supplemental Tables S4 and S5) and therefore may target glycan regions that lack this modification.

Several PULs operons (BT0752-57 and BT4631-36) and operons (BT3796-99 and BT4642-50) were identified here for the first time as being induced in vivo. The functions of these operons are unknown but they encode glycosidases and several sulfatases genes indicating a potential function for sulfated host glycans metabolism. BT0752-57 shows increased expression in the absence of anSME. Increased expression of this locus is consistent with the fact that this system is associated with an extracytoplasmic function- σ /anti- σ regulator that likely activates expression via a trans-envelope signaling mechanism triggered by substrate transport through the specialized SusC homolog BT0754. This receptor might recognize sulfated oligosaccharides (or sulfated glycosylations) which are then desulfated by the single sulfatase encoded by BT0756. This sulfatase contains an N-terminal secretion signal but no predicted lipidation site, suggesting that it functions in the periplasm. Thus, in the $\Delta anSME$ strain, the inability to desulfate and subsequently to hydrolyze the glycan signal that triggers this system likely explains its increased expression (supplemental Fig. S1).



FIGURE 7. Phylogenetic tree and potential functions of sulfatases found in *B. thetaiotaomicron* VPI 5482. *Purple*, sulfatases induced *in vitro* in presence of glycosaminoglycans; *orange*, sulfatase induced *in vivo* and *in vitro* in presence of mucin; *red* and *blue*, sulfatases induced only *in vivo* based on a previous (8) or the current study in the *AanSME* strain, respectively. Sequence alignment was performed with ClustalW. The phylogenetic tree was generated using the program Mega and the neighbor-joining method with the Kimura two-parameter calculation model. An *open star* indicates that the protein contains a predicted signal peptidase I cleavage site, and a *filled star* indicates that the protein contains a predicted signal peptidase II cleavage site, whereas a *gray star* indicates ambiguity in prediction based on SignalP (version 3.0) or LipoP (version 1.0).

Indeed, glycosidases, especially exoglycosidases, are usually unable to modify sulfated oligosaccharides.

In contrast, the PUL BT0865-67 exhibits dramatically diminished expression in the $\Delta anSME$ strain, suggesting that the sensor regulating this locus responds to a glycan signal that is blocked by sulfation. Thus, we speculate that in the absence of sulfatase activity potentially provided by an enzyme unlinked to the BT0865-67 locus, the signal that triggers this PUL is not perceived and the BT0865-67 operon, which is highly expressed *in vivo* under conditions where *B. thetaiotaomicron* is forced to forage host glycans (9, 19), is not induced. Among the PULs repressed in the mutant, we identified new PULs of unknown functions (BT3517-32, BT4145-4183, and BT4266-4272) but encoding numerous glycosidases further expanding the number of genes involved in *B. thetaiotaomicron* adaptation to its host.

Analysis of the genomes of 46 sequenced human gut Bacteroidetes (supplemental Table S6) revealed that all possess an anSME gene and that several genes encoding sulfatases are present within many species. For example, *Bacteroides fragilis, Bacteroides dorei*, or *Parabacteroides distasonis* each encodes \geq 19 sulfatases in their genomes. In contrast, genes encoding predicted sulfatases are notably absent from most of the Firmicutes (21), the other dominant bacterial phylum represented in the adult human gut microbiota. This observation further demonstrates that sulfatases are an important and evolutionary conserved feature among Bacteroidetes inhabiting the human digestive tract.

Sulfatases as Fitness Factors for a Human Gut Symbiont

This current and our previous report on *B. thetaiotaomicron* in vivo colonization (8) allowed us to show that almost half of the 28 encoded sulfatases (Fig. 7) are expressed only in the presence of the two major classes of sulfated macromolecules found in the human body: mucins or glycosaminoglycans. This demonstrates a complex inter-relationship between commensal bacteria and their human host mediated by sulfatases. Expression of active sulfatases is critical for the fitness of B. thetaiotaomicron. This organism, by virtue of its ability to forage sulfated glycans is, in turn, in a position to shape features of its gut habitat, notably modification of host mucins that it likely contacts directly during growth in the mucus layer. Alterations in mucus composition, notably extensive desulfation has been reported in the intestines of individuals with inflammatory bowel diseases (24). As metagenomic studies of the gut microbiomes of individuals with various forms of inflammatory bowel diseases, or other disorders where intestinal mucosal barrier function is disrupted, it will be interesting to ascertain whether the representation and expression of genes encoding sulfatases and the enzymes responsible for their activation correlates with disease type, disease activity, and nutritional status. Such an analysis may yield new potential biomarkers, mediators, or even therapeutic targets.

Acknowledgments—We thank J. J. Gratadoux and S. Blugeon for technical assistance and Laura Kyro for support with graphics.

REFERENCES

- Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., Mende, D. R., Li, J., Xu, J., Li, S., Li, D., Cao, J., Wang, B., Liang, H., Zheng, H., Xie, Y., Tap, J., Lepage, P., Bertalan, M., Batto, J. M., Hansen, T., Le Paslier, D., Linneberg, A., Nielsen, H. B., Pelletier, E., Renault, P., Sicheritz-Ponten, T., Turner, K., Zhu, H., Yu, C., Li, S., Jian, M., Zhou, Y., Li, Y., Zhang, X., Li, S., Qin, N., Yang, H., Wang, J., Brunak, S., Doré, J., Guarner, F., Kristiansen, K., Pedersen, O., Parkhill, J., Weissenbach, J., Bork, P., Ehrlich, S. D., and Wang, J. (2010) *Nature* 464, 59–65
- Turnbaugh, P. J., Hamady, M., Yatsunenko, T., Cantarel, B. L., Duncan, A., Ley, R. E., Sogin, M. L., Jones, W. J., Roe, B. A., Affourtit, J. P., Egholm, M., Henrissat, B., Heath, A. C., Knight, R., and Gordon, J. I. (2009) *Nature* 457, 480–484
- Turnbaugh, P. J., Quince, C., Faith, J. J., McHardy, A. C., Yatsunenko, T., Niazi, F., Affourtit, J., Egholm, M., Henrissat, B., Knight, R., and Gordon, J. I. (2010) *Proc. Natl. Acad. Sci. U.S.A.* 107, 7503–7508
- Gloux, K., Berteau, O., El Oumami, H., Béguet, F., Leclerc, M., and Doré, J. (2011) Proc. Natl. Acad. Sci. U.S.A. 108, 4539 – 4546
- Goodman, A. L., McNulty, N. P., Zhao, Y., Leip, D., Mitra, R. D., Lozupone, C. A., Knight, R., and Gordon, J. I. (2009) *Cell Host. Microbe.* 6, 279–289
- Martens, E. C., Koropatkin, N. M., Smith, T. J., and Gordon, J. I. (2009) J. Biol. Chem. 284, 24673–24677
- Rey, F. E., Faith, J. J., Bain, J., Muehlbauer, M. J., Stevens, R. D., Newgard, C. B., and Gordon, J. I. (2010) *J. Biol. Chem.* 285, 22082–22090
- Martens, E. C., Chiang, H. C., and Gordon, J. I. (2008) *Cell Host. Microbe.* 4, 447–457
- Sonnenburg, J. L., Xu, J., Leip, D. D., Chen, C. H., Westover, B. P., Weatherford, J., Buhler, J. D., and Gordon, J. I. (2005) *Science* **307**, 1955–1959
- Xu, J., Mahowald, M. A., Ley, R. E., Lozupone, C. A., Hamady, M., Martens, E. C., Henrissat, B., Coutinho, P. M., Minx, P., Latreille, P., Cordum, H., Van Brunt, A., Kim, K., Fulton, R. S., Fulton, L. A., Clifton, S. W., Wilson, R. K., Knight, R. D., and Gordon, J. I. (2007) *PLoS Biol.* 5, e156
- 11. Benjdia, A., Dehò, G., Rabot, S., and Berteau, O. (2007) *FEBS. Lett.* **581**, 1009–1014





- Carlson, B. L., Ballister, E. R., Skordalakes, E., King, D. S., Breidenbach, M. A., Gilmore, S. A., Berger, J. M., and Bertozzi, C. R. (2008) *J. Biol. Chem.* 283, 20117–20125
- Dierks, T., Schmidt, B., Borissenko, L. V., Peng, J., Preusser, A., Mariappan, M., and von Figura, K. (2003) *Cell* 113, 435–444
- 14. Benjdia, A., Leprince, J., Guillot, A., Vaudry, H., Rabot, S., and Berteau, O. (2007) J. Am. Chem. Soc. **129**, 3462–3463
- Benjdia, A., Leprince, J., Sandström, C., Vaudry, H., and Berteau, O. (2009) J. Am. Chem. Soc. 131, 8348–8349
- Benjdia, A., Subramanian, S., Leprince, J., Vaudry, H., Johnson, M. K., and Berteau, O. (2008) *J. Biol. Chem.* 283, 17815–17826
- 17. Benjdia, A., Subramanian, S., Leprince, J., Vaudry, H., Johnson, M. K., and Berteau, O. (2010) *FEBS. J.* **277,** 1906–1920
- Kurokawa, K., Itoh, T., Kuwahara, T., Oshima, K., Toh, H., Toyoda, A., Takami, H., Morita, H., Sharma, V. K., Srivastava, T. P., Taylor, T. D., Noguchi, H., Mori, H., Ogura, Y., Ehrlich, D. S., Itoh, K., Takagi, T., Sakaki, Y., Hayashi, T., and Hattori, M. (2007) *DNA. Res.* 14, 169–181
- 19. Bjursell, M. K., Martens, E. C., and Gordon, J. I. (2006) J. Biol. Chem. 281, 36269–36279
- Sokol, H., Pigneur, B., Watterlot, L., Lakhdari, O., Bermúdez-Humarán, L. G., Gratadoux, J. J., Blugeon, S., Bridonneau, C., Furet, J. P., Corthier, G., Grangette, C., Vasquez, N., Pochart, P., Trugnan, G., Thomas, G., Blottière, H. M., Doré, J., Marteau, P., Seksik, P., and Langella, P. (2008) *Proc. Natl. Acad. Sci. U.S.A.* 105, 16731–16736

- 21. Berteau, O., Guillot, A., Benjdia, A., and Rabot, S. (2006) *J. Biol. Chem.* **281**, 22464–22470
- Sonnenburg, E. D., Sonnenburg, J. L., Manchester, J. K., Hansen, E. E., Chiang, H. C., and Gordon, J. I. (2006) *Proc. Natl. Acad. Sci. U.S.A.* 103, 8834–8839
- Hanson, S. R., Best, M. D., and Wong, C. H. (2004) Angew. Chem. Int. Ed. Engl. 43, 5736–5763
- 24. Tsai, H. H., Dwarakanath, A. D., Hart, C. A., Milton, J. D., and Rhodes, J. M. (1995) *Gut.* **36**, 570–576
- 25. Xu, J., Bjursell, M. K., Himrod, J., Deng, S., Carmichael, L. K., Chiang, H. C., Hooper, L. V., and Gordon, J. I. (2003) *Science* **299**, 2074–2076
- Osterås, M., Boncompagni, E., Vincent, N., Poggi, M. C., and Le Rudulier, D. (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95, 11394–11399
- Salyers, A. A., Pajeau, M., and McCarthy, R. E. (1988) *Appl. Environ. Microbiol.* 54, 1970–1976
- Hehemann, J. H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M., and Michel, G. (2010) Nature 464, 908–912
- Lee, J. K., Bistrup, A., van Zante, A., and Rosen, S. D. (2003) *Glycobiology*. 13, 245–254
- Tobisawa, Y., Imai, Y., Fukuda, M., and Kawashima, H. (2010) J. Biol. Chem. 285, 6750-6760
- Martens, E. C., Roth, R., Heuser, J. E., and Gordon, J. I. (2009) J. Biol. Chem. 284, 18445–18457



Sulfatases and a Radical S-Adenosyl-I-methionine (AdoMet) Enzyme Are Key for Mucosal Foraging and Fitness of the Prominent Human Gut Symbiont, Bacteroides thetaiotaomicron

Alhosna Benjdia, Eric C. Martens, Jeffrey I. Gordon and Olivier Berteau

J. Biol. Chem. 2011, 286:25973-25982. doi: 10.1074/jbc.M111.228841 originally published online April 20, 2011

Access the most updated version of this article at doi: 10.1074/jbc.M111.228841

Alerts:

- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material: http://www.jbc.org/content/suppl/2011/04/20/M111.228841.DC1

This article cites 31 references, 17 of which can be accessed free at http://www.jbc.org/content/286/29/25973.full.html#ref-list-1