

Efficient utilization of complex N-linked glycans is a selective advantage for *Bacteroides fragilis* in extraintestinal infections

Yanlu Cao, Edson R. Rocha, and C. Jeffrey Smith¹

Department of Microbiology and Immunology, Brody School of Medicine, East Carolina University, Greenville, NC 27834

Edited by Thomas J. Silhavy, Princeton University, Princeton, NJ, and approved July 30, 2014 (received for review April 22, 2014)

***Bacteroides fragilis* is the most common anaerobe isolated from clinical infections, and in this report we demonstrate a characteristic of the species that is critical to their success as an opportunistic pathogen. Among the *Bacteroides* spp. in the gut, *B. fragilis* has the unique ability of efficiently harvesting complex N-linked glycans from the glycoproteins common to serum and serous fluid. This activity is mediated by an outer membrane protein complex designated as Don. Using the abundant serum glycoprotein transferrin as a model, it has been shown that *B. fragilis* alone can rapidly and efficiently deglycosylate this protein in vitro and that transferrin glycans can provide the sole source of carbon and energy for growth in defined media. We then showed that transferrin deglycosylation occurs in vivo when *B. fragilis* is propagated in the rat tissue cage model of extraintestinal growth, and that this ability provides a competitive advantage in vivo over strains lacking the *don* locus.**

The genus *Bacteroides* comprises Gram-negative, obligate anaerobic bacteria constituting one of the largest stable components of the microbiota in the human large intestine (1, 2). This relationship benefits the host by aiding in the development of healthy immune response and in the maintenance of many physiological and nutritional processes. One major factor that contributes to the predominance of *Bacteroides* is their capacity to use a wide spectrum of polysaccharides ranging from dietary compounds that cannot be digested by the host, to glycans derived from the host in the form of glycoconjugates (3, 4). Polysaccharide digestion in the *Bacteroides* is mediated in large part by outer membrane complexes that bind, cleave, and transport these substrates. These systems have a genetic organization similar to the starch utilization system (Sus) of *Bacteroides thetaiotaomicron*, which are characterized by a regulatory region followed by an operon coding for homologs of the TonB-dependent transporter, SusC, and the accessory binding protein SusD. Additional genes in the operon are for specific substrate-binding and glycohydrolase enzymes (5). The genes for these protein complexes are termed polysaccharide utilization loci (PULs).

In contrast to their beneficial role in the colon, *Bacteroides* species also are the most common opportunistic pathogens isolated from clinical specimens of anaerobic infections. These opportunistic infections can occur when the integrity of the intestinal wall becomes compromised. Predisposing conditions include intestinal surgery, perforated appendix, carcinoma, diverticulitis, and trauma. Peritonitis and intraabdominal abscesses are the most common infections associated with *Bacteroides* often leading to bacteremia (6). Notably, *Bacteroides fragilis* is the most frequently isolated anaerobe from human infections, thus it is regarded as more invasive than the other *Bacteroides* species (7, 8). A number of factors have been identified that may contribute to its enhanced extraintestinal survival. For example, there is a variable polysaccharide capsule which interferes with immune surveillance, and there is a robust oxidative stress resistance system, but these attributes are not exclusive to *B. fragilis* (9, 10). Another challenge faced by organisms invading from the

colon is the need to adapt to different nutritional sources. In the gut, *Bacteroides* rely on dietary polysaccharides and host-derived glycans as sources of carbon/energy, but when outside the gut environment their choice is limited to host-derived glycans. The ability to efficiently harvest glycans present on host glycoproteins in extraintestinal sites may enhance survival and colonization.

In this report we describe a Sus-like PUL unique to *B. fragilis* and that promotes efficient utilization of complex N-linked glycans from the most abundant serum/serous fluid glycoproteins including transferrin. Overall the studies suggest that an ability to harvest these glycans is advantageous for growth at extraintestinal sites and can explain, in part, the success of *B. fragilis* as an opportunistic pathogen relative to the other *Bacteroides* species.

Results

In Vivo Extraintestinal PUL Expression. Considerable progress has been made in understanding how the *Bacteroides* compete for and use limited nutrients in the gut ecosystem (3, 4). In contrast, there is a paucity of knowledge concerning *Bacteroides* nutrient foraging in extraintestinal environments. As an initial strategy to identify significant pathways of catabolism, microarrays were used to measure gene expression of *B. fragilis* cells growing in a rat tissue cage abscess model. The results suggested that the organism may rely on a variety of PULs to harvest glycans from host proteins present in the serous fluid of the model abscess. *B. fragilis* has 47 putative PULs and 22 stand-alone SusC/SusD-like protein pairs, and several of these were significantly induced in vivo compared with growth in defined media (DM) with glucose. Five genes, BF638R3439–BF638R3443, associated with one of

Significance

Human microbiota have a huge impact on health from proper development of the immune system to maintenance of normal physiological processes. The largest concentration of microbes is found in the colon, which is home to hundreds of bacterial species, most of which are obligate anaerobes. This population also poses a significant threat of opportunistic infection, and of all the species present, *Bacteroides fragilis* is the anaerobe most frequently isolated from extraintestinal infections of intestinal origin. The findings presented here describe a unique ability of this species to efficiently deglycosylate complex N-linked glycans from the most abundant glycoproteins found in serum and serous fluid, which gives *B. fragilis* a competitive, nutritional advantage for extraintestinal growth.

Author contributions: Y.C., E.R.R., and C.J.S. designed research; Y.C. performed research; Y.C., E.R.R., and C.J.S. analyzed data; and Y.C. and C.J.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The microarray expression data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE53883).

¹To whom correspondence should be addressed. Email: smithcha@ecu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1407344111/-DCSupplemental.

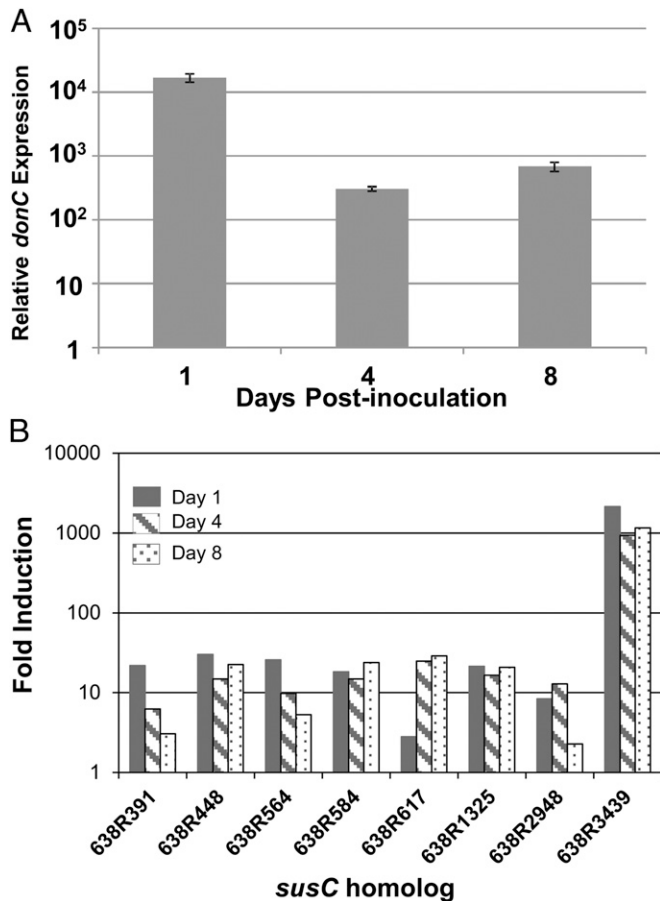


Fig. 1. In vivo PUL gene expression during growth in the rat tissue cage model. Samples for analyses were pooled from five animals 1, 4, and 8 d post-inoculation. (A) Expression of *donC* in samples from the rat tissue cage relative to midlogarithmic cultures grown in DM–glucose. Expression was measured in triplicate samples by qRT-PCR and normalized to the amount of 16S rRNA. Values are shown with SDs. (B) Induction of SusC homologs during growth in vivo relative to DM–glucose. Shown are the highly induced in vivo *susC*-like genes associated with PULs in cluster 1. Clusters were determined from expression microarray data analyzed by *k*-means clustering using the standard Pearson’s correlation coefficient distance metric. All induction values were significant at $P < 0.01$.

these PULs were induced from 2,147- to 765-fold and were the most highly induced genes in vivo. These genes are in an operon that constitutes part of a PUL named *donABCDEFGHIJ* (SI Appendix, Fig. S1).

To more easily compare in vivo expression of all PUL operons, the SusC homologs were used as a proxy for expression of PUL operons (11). This approach was validated by quantitative RT-PCR (qRT-PCR) which confirmed the in vivo induction of *donC* (Fig. 1A). The results from day 1, 4, and 8 samples were compared with in vitro midlogarithmic phase cultures and showed a high level of induction throughout the course of the experiment.

Examination of the SusC homolog microarray expression data by *k*-means clustering revealed four major in vivo PUL gene induction patterns (SI Appendix, Table S1 and Fig. S2). The *donC* gene was associated with the most highly induced genes, cluster 1, which consisted of eight PULs and five stand-alone SusC/SusD homologs. As shown in Fig. 1B, the PULs in this cluster were rapidly induced more than 10-fold and the level of induction remained high throughout the 8 d. A second cluster of 25 SusC homologs was induced more gradually and remained induced 1.5- to 5-fold for the 8 d. The remaining two PUL clusters were repressed in the abscess model and may be more important in the intestinal environment.

In Vitro Induction of the *donCDEFG* Operon. The Don PUL was chosen as a model for extraintestinal PUL function and included several gene products that may be associated with N-linked glycan utilization, such as an endo-S-like endo-*N*-acetyl- β -D-glucosaminidase (DonE) and a lectin-like product with an α -*N*-acetylglucosaminidase domain (DonF). To demonstrate a function of the PUL, it was necessary to find an inducing substrate for in vitro expression studies. This was accomplished by preparing a crude glycan mixture by alkaline β -elimination hydrolysis of porcine stomach mucin. Although the majority of mucin glycans are O-linked, some N-linked glycans also are released by this method (4). A defined medium containing the glycan mixture as the sole carbon/energy source supported rapid growth of the WT strain (Fig. 2). Expression analysis by qRT-PCR showed that *donC* was strongly induced in this medium in a growth-phase-dependent manner with the highest level of induction at mid-logarithmic phase.

Deglycosylation of Transferrin Is Mediated by *donCDEFG*. Don function was explored by incubating serous fluid from uninfected rat tissue cages with WT or $\Delta donCDEFG$ (Δdon) mutant cells grown to midlogarithmic phase in the inducing, mucin glycan medium. Samples were analyzed by SDS/PAGE and the results in Fig. 3A show that one of the major fluid proteins was significantly reduced in size in samples treated with the WT strain but not with the deletion mutant. Analysis by MS revealed that this protein was rat serum transferrin, one of the most abundant glycoproteins in serum. The precise size reduction of the protein suggested the possibility that the transferrin glycans were removed. To test this, human transferrin was incubated with WT or Δdon mutant cells. Duplicate samples were analyzed by SDS/PAGE with either Coomassie blue staining for protein or *Sambucus nigra* agglutinin (SNA) lectin staining for the terminal Neu5Ac- α -2-6Gal(NAc)-R structure common to complex N-linked glycans. As shown in Fig. 3B, transferrin treated with the WT strain was reduced in size after the 3-h incubation, whereas the deletion strain did not have an obvious effect. Also, no degradation of the protein band was observed in any of the samples. Experiments with SNA lectin showed that incubation of transferrin with WT cells resulted in a loss of SNA staining compared with little or no change in assays with the deletion mutant and controls (Fig. 3B). Other serous fluid proteins also were rapidly deglycosylated by *B. fragilis*. In assays using serous fluid, the WT strain had significant deglycosylation activity on many serous fluid proteins but the Δdon mutant activity was considerably reduced

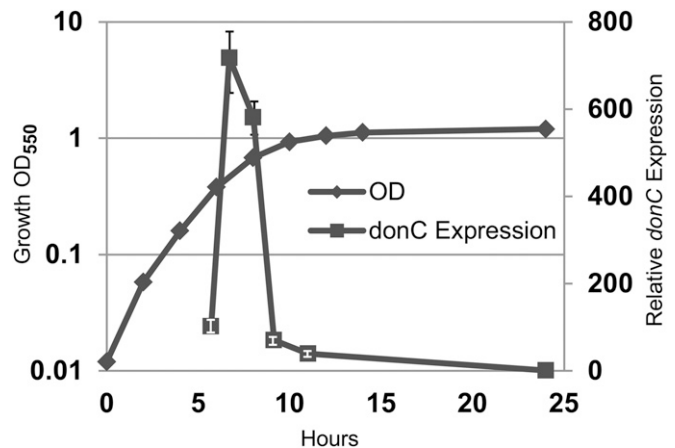


Fig. 2. Induction of *donC* expression during growth on mucin glycans. Strain BF638R was grown in DM containing 2% porcine gastric mucin glycans as the sole carbon/energy source. The induction of *donC* relative to growth on glucose was determined by qRT-PCR. The results are overlaid on a typical growth curve with mucin glycans (OD₅₅₀, gray diamonds). The qRT-PCR results represent two independent experiments (gray squares) with SDs shown.

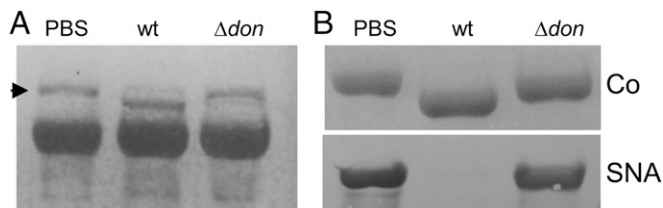


Fig. 3. Deglycosylation of transferrin is mediated by the Don PUL. (A) Serous fluid obtained from uninfected tissues cages was incubated for 3 h with PBS or *B. fragilis* cells induced by growth in DM–mucin glycan medium. Samples were analyzed by 12% SDS/PAGE and stained with Ponceau S. The arrow indicates the location of transferrin above the abundant serum albumin protein. (B) Deglycosylation analysis of human transferrin. Human transferrin was incubated with WT or Δdon cells induced by growth in DM–mucin glycan medium. Samples were analyzed by SDS/PAGE and stained for protein with Coomassie blue (Co) or SNA lectin to detect N-linked glycans as described in *Materials and Methods*.

(*SI Appendix, Fig. S3*). Taken together, the decrease in transferrin mass and its loss of SNA staining in assays with WT cells indicate that the *don* locus mediates deglycosylation of N-glycans from transferrin and other proteins in serous fluid.

The SNA lectin is suitable for the detection of N-linked glycans but its use in the assay does not provide definitive evidence of complete deglycosylation. Therefore, a comprehensive MS glycomics analysis was performed to determine the structure of the glycans remaining on transferrin after incubation with *B. fragilis*. Human transferrin has two glycosylation sites, and there were seven distinct oligosaccharides linked to those sites, five of which made up 98.1% of the total transferrin glycans (Table 1). Incubation of transferrin with WT cells for 3 h resulted in essentially complete deglycosylation. By comparison, nearly 70% of the N-linked glycans remained on transferrin after incubation with the Δdon mutant. There were no new oligosaccharide structures detected in reactions with the WT strain suggesting efficient cleavage of the glycan likely at the β -1,4-di-N-acetylchitobiose core by the endo- β -N-acetylglucosaminidase S (DonE). These results confirm that N-glycans are a substrate for the Don PUL but do not rule out the possibility that this system may also have activity on some O-linked glycans. These genes were named *don* for deglycosylation of N-linked glycans.

Growth on Transferrin as the Sole Carbon and Energy Source. The ability to forage for high-quality carbohydrates in extraintestinal sites would be advantageous to *Bacteroides* spp. translocated from the gut. Transferrin is an abundant glycoprotein present in serum and is estimated to have a carbohydrate content of 5.8%, so we tested if it could support the growth of *B. fragilis*. As shown in Fig. 4, 25 mg/mL of transferrin supported robust growth of the WT strain but not the Δdon mutant. There was no significant difference in the growth rate between WT and mutant strains grown with glucose (Fig. 4). Expression of the *don* operon during growth on transferrin was measured by qRT-PCR with *donC* primers, and the results showed a 10^4 -fold induction over glucose grown cells (*SI Appendix, Fig. S4*). SDS/PAGE analysis of the transferrin medium after 24 or 48 h of growth indicated there was no degradation of the transferrin peptide. The nonglycosylated protein, BSA, did not support growth (Fig. 4). Interestingly, the Δdon mutant was able to grow in the transferrin medium albeit at a much slower rate. This confirms the presence of a second system that enables the slow utilization of transferrin glycans.

In Vivo Role of the Don PUL. To investigate a potential role for the *don* operon in extraintestinal sites, the rat tissue cage model was used to measure growth in vivo. There was no drastic difference in cfu counts between WT and Δdon strains in monoculture experiments over the course of a 15-d period (Fig. 5A). There was however a small, twofold advantage for the WT in the first 2 d following inoculation. To determine if this difference would

translate into a competitive advantage, mixed culture assays were performed using WT and mutant strains coinoculated into rat tissue cages. The results were clear; Δdon was quickly out-competed, and at 1 d postinoculation about 85% of the total population was WT (Fig. 5B). The same trend carried through day 8 when the WT strain reached about 97% of the total population. In control experiments (*SI Appendix, Fig. S5*) strains coinoculated into DM–glucose showed no significant difference through logarithmic phase, and in fact, the Δdon mutant had a slight advantage in stationary phase.

The potential importance of transferrin as an extraintestinal carbohydrate source in vivo was demonstrated by SDS/PAGE analysis of fluids removed from rat tissue cages over the course of the experiment (Fig. 5C). The transferrin from uninfected controls appeared as a single peptide species of about 81 kDa but by 1-d postinoculation there were two transferrin peptides, 78 and 81 kDa, in samples from WT-infected animals. After 2 d nearly all transferrin was converted to the smaller species. In contrast, samples from animals inoculated with the Δdon strain appeared unchanged at day 1 and by day 2 a small portion of the transferrin was deglycosylated. Interestingly all of the transferrin in samples from WT- or mutant-infected animals was deglycosylated by day 8 (Fig. 5C). Overall these results show that the *don* locus plays an important role in the efficient deglycosylation of transferrin in vivo, and that well-developed mechanisms of N-glycan foraging are advantageous in extraintestinal sites.

Efficient Transferrin Deglycosylation Is Unique to *B. fragilis*. *B. fragilis* is the most common *Bacteroides* species isolated from opportunistic infections at extraintestinal sites. The ability of six other intestinal *Bacteroides* species to deglycosylate human transferrin was tested by SDS/PAGE and SNA glycan staining. The results in Fig. 6 show that only the two *B. fragilis* strains were able to efficiently deglycosylate transferrin in the 3-h assays. This was indicated by both the change in mass and the loss of SNA staining of transferrin. No significant loss in SNA staining was seen for the other species except *Bacteroides vulgatus*. In this case the transferrin band appeared as a doublet with a somewhat decreased staining intensity in the SNA blot (Fig. 6, lane 7). These data indicate that *B. vulgatus* may deglycosylate transferrin inefficiently, but of the intestinal *Bacteroides* species tested, only

Table 1. Quantitative glycan analysis of transferrin incubated with *B. fragilis* cells

Glycan ^b	% Total ^c	Glycan Remaining on Transferrin (μ M) ^a		
		PBS	wt	Δdon
	87.5%	15.86 \pm 0.32	ND	10.43 \pm 1.06
	4.0%	0.72 \pm 0.03	ND	0.25 \pm 0.01
	3.3%	0.60 \pm 0.02	ND	0.71 \pm 0.11
	1.9%	0.34 \pm 0.01	0.05	0.27 \pm 0.07
	1.4%	0.25 \pm 0.01	ND	0.21 \pm 0.04
Totals	98.1%	17.77 \pm 0.4	0.05	11.87 \pm 1.11

^aTransferrin was incubated with PBS, strain 638R (wt), or the Δdon mutant for 3 hours as described in the text. The amount of glycan remaining on transferrin is shown normalized to 1 g/L of transferrin and results are the average of 2 independent samples. ND=not detected.

^bThe five major glycans on transferrin are shown and were composed of the following sugars: galactose, ; N-acetylglucosamine, ; mannose, ; fucose, ; and sialic acid .

^cThe % of the total glycans linked to transferrin prior to treatment with *B. fragilis*.

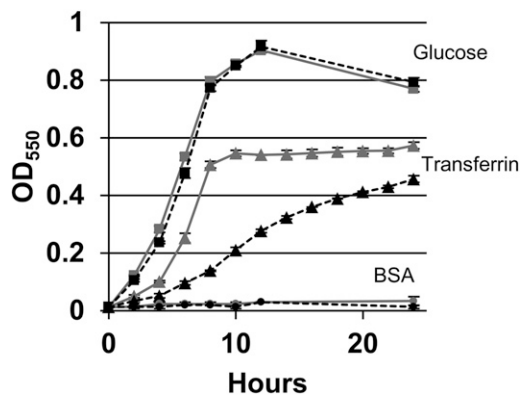


Fig. 4. *B. fragilis* can grow with transferrin as the sole source of carbon/energy. Growth curves are shown for WT or Δdon strains in DM with different carbon/energy sources. An overnight inoculum of 2% was used, and the OD₅₅₀ was measured at specific time intervals. Dashed lines represent Δdon and solid lines represent WT. Squares represent glucose (0.4%), triangles human transferrin (25 mg/mL), and circles BSA (25 mg/mL). Each growth curve represents two biological repeats with the SDs shown.

B. fragilis can rapidly remove the N-linked glycans from this abundant serum protein.

Consistent with these data are results from BLAST searches where the entire Don locus was found to be highly conserved (92–99% identity) for all 16 *B. fragilis* strains in National Center for Biotechnology Information (NCBI) and Human Microbiome Project databases. Further, phylogenetic analyses of DonE, DonF, and DonG indicate that they are significantly diverged from other similar Bacteroidetes proteins, suggesting they have a unique function in *B. fragilis* (SI Appendix, Fig. S6).

Discussion

Intestinal microbiota have a complex, symbiotic relationship with their host that is maintained by a series of physical barriers and immunological processes designed to keep the microbial populations in check. When these barriers are breached, invading organisms must rapidly adapt to an environment with a new set of physical, chemical, and nutritional parameters. In this study, we used a rat tissue cage model to gain a better understanding of the nutritional sources that opportunists such as *B. fragilis* might encounter in extraintestinal habitats. Expression microarrays showed that PUL genes were highly induced in vivo and that *don* genes were induced more than any others. The carbon/energy sources available in the abdominal cavity are quite varied, so we were not surprised to see a robust, complex response in PUL expression. A recent study of *Bacteroides thetaiotaomicron* PUL regulation during growth on mixtures of dietary glycans provides some insight into this response (11). The work showed that there is a rapid response to the influx of new glycan sources, and that multiple PULs were simultaneously expressed to prioritize utilization of the most advantageous substrates. This required coordinated induction and repression of multiple systems working to maximize the utilization of certain glycans, although the rationale for prioritization is not yet entirely clear. Similarly, *B. fragilis* growing in the rat tissue cage elicited a broad PUL response, and in this case it seemed to be directed toward the use of N-linked glycans found on host glycoproteins. This is reasonable considering nearly all proteins in the serous fluid except serum albumin are N-linked glycoproteins, and transferrin, being the most abundant of these, is a significant substrate for the Don PUL (SI Appendix, Table S2). PULs that were induced in the microarray experiments that have protein signatures consistent with the utilization of N-linked glycans were as follows: BF638R0384–BF638R0392 (β -N-1-4 acetylglucosaminidase + GH88 glycosidase), BF638R1323–BF638R1330 (endo- β -N-acetylglucosaminidase + Con

A-like), and BF638R0444–BF638R0448 have member proteins with fibronectin-binding domains.

The hydrolysis of host glycoprotein glycans is an important characteristic of several bacterial pathogens and this has been proposed to be a mechanism for immune evasion as well as for nutrient acquisition. *Streptococcus pyogenes* secretes the archetypical GH18 family glycoprotein hydrolase, EndoS, which can efficiently hydrolyze the IgG glycan and significantly reduce IgG-mediated killing in blood (12, 13). *Streptococcus pneumoniae* possesses several exoglycosidases that in combination deglycosylate IgA1, lactoferrin, human secretory component, and α 1-acid glycoprotein; and it has been shown that deglycosylation of α 1-acid glycoprotein has supported growth in the absence of other carbohydrates (14, 15). This ability may be linked to the persistence of *S. pneumoniae* in the nasopharynx. Likewise transferrin supports the growth of *B. fragilis* (Fig. 4) and is present in serum and serous fluid at 2–4 mg/mL, so this would be an excellent source of fermentable carbohydrate for extraintestinal growth. Glycan harvesting also seems to be important for in vivo growth of *Capnocytophaga canimorsus*, a member of the *Bacteroidetes* phylum and a common inhabitant of the oral cavity of dogs (16). *C. canimorsus* was able to deglycosylate fetuin and IgG in vitro using a Sus-like system, and the deglycosylation activity was necessary to sustain growth with cultured mammalian cells.

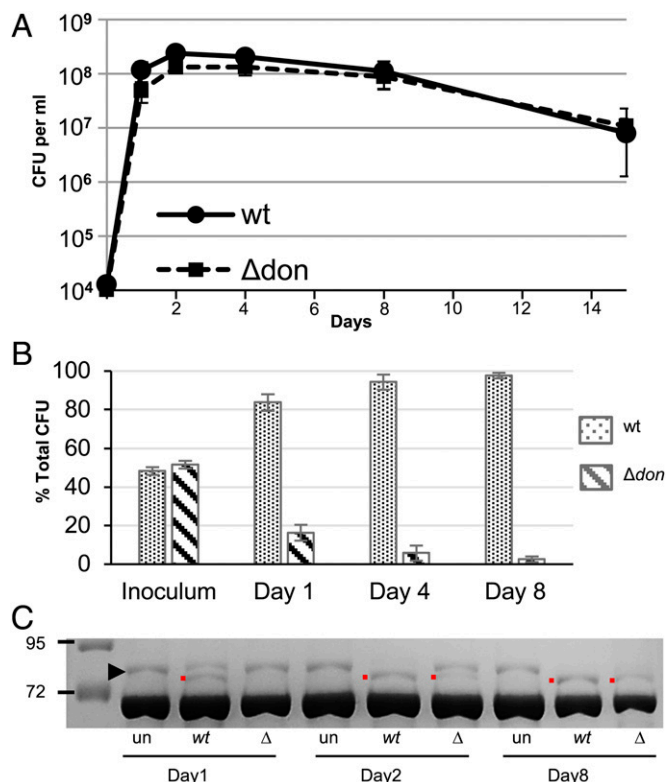


Fig. 5. Growth in vivo is enhanced by deglycosylation of transferrin. (A) Growth curve for WT (circles) and Δdon (squares) strains inoculated separately into rat tissue cages. (B) Mixed culture competition assay. Mixtures of WT (stippled bars) and Δdon (cross-hatched bars) cells were prepared in a 1:1 ratio, inoculated into five rats, and the cfu per milliliter determined. The percentage of WT or Δdon cells was determined by screening the total cell counts for tetracycline-resistant colonies. (C) Coomassie blue-stained SDS/PAGE gel of serous fluid samples obtained following growth of strains in the rat tissue cage. Samples from the WT and Δdon (Δ) strains were compared with uninoculated controls (un). The arrowhead indicates the migration of the glycosylated transferrin, and red squares indicate deglycosylated transferrin. The results in A and B are averaged from two biological repeats with the SDs shown, and C shows a representative result.

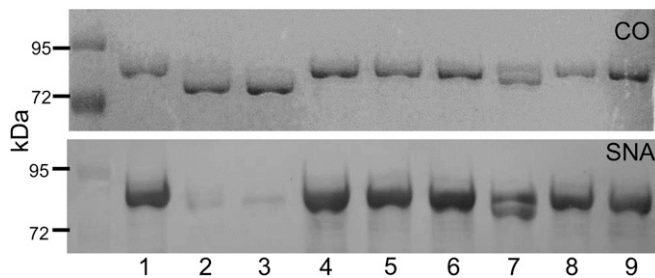


Fig. 6. Deglycosylation of human transferrin by medically important *Bacteroides* species. Pure human transferrin was used in standard 3-h deglycosylation assays with midlogarithmic phase cells of *Bacteroides* species grown in DM–mucin glycan media. Samples were analyzed by SDS/PAGE with Coomassie blue (CO) staining and on duplicate SDS/PAGE gels, followed by SNA glycan staining. Lanes: 1, PBS; 2, *B. fragilis* (638R); 3, *B. fragilis* (American Type Culture Collection 25285); 4, *B. thetaiotaomicron* (IB116); 5, *Bacteroides uniformis*; 6, *Bacteroides ovatus*; 7, *B. vulgatus*; 8, *Parabacteroides distasonis*; 9, *Parabacteroides merdae*.

A notable difference between the Don PUL and glycosidase systems mentioned above is the inability to deglycosylate IgG or IgA. Several attempts with different sources of human IgG and IgA failed to demonstrate activity against these substrates (*SI Appendix, Fig. S7*). Further, it is well known that transferrin and lactoferrin also have antibacterial activity, largely due to their ability to sequester iron from bacteria (17, 18). However, the glycosylation status of transferrin and lactoferrin has little effect on their ability to either bind iron or their receptors (19). Overall it does not seem that Don PUL was designed to circumvent the immune system.

A key finding of the work presented here was that rapid, efficient deglycosylation of transferrin and many other glycoproteins in serous fluid is mediated by the *don* genetic locus, which is unique to *B. fragilis* among the medically important *Bacteroides* species. This was demonstrated specifically for transferrin but may apply to the other serous proteins (Fig. 6 and *SI Appendix, Fig. S3*). Other *Bacteroides* may have N-linked deglycosylation capabilities more similar to the Δdon mutant. For example, *B. vulgatus* showed some deglycosylation of transferrin after 3 h (Fig. 6) and results with a second strain of *B. thetaiotaomicron* showed some activity on transferrin after 3 h but failed to further deglycosylate the substrate even after overnight incubation (*SI Appendix, Fig. S8*).

We propose that the Don PUL is composed of seven genes. DonA and DonB are the regulators of the system encoding an extracytoplasmic function (ECF) sigma factor and antisigma factor, respectively. DonEFG are outer membrane proteins that bind and cleave glycans from the target glycoproteins. DonC and DonD would work in tandem to transport oligosaccharides into the periplasm where the concerted effort of neuraminidases and other glycohydrolases release monosaccharides from the glycan chains (5). A critical requirement of this model is the removal of terminal sialic acid residues from the oligosaccharides transported to the periplasm. This would be accomplished by neuraminidase, NanH, which in a previous study was shown to be necessary for robust growth in a rat pouch model similar to the tissue cage model used here (20). Notably NanH has strong periplasmic localization signals. An extensive region of the *B. fragilis* chromosome (BF638R1715–BF638R1740) is devoted to sialic acid utilization and includes the *nanLET* operon, the *nanH* operon, and five stand-alone pairs of SusC/SusD homologs (21, 22). Our microarray data showed that regulation of these genes was complex but the entire region was induced on day 1 following inoculation of the tissue cage, and for the most part remained elevated throughout (*SI Appendix, Table S1* and *Fig. S2*).

The normal habitat for *B. fragilis* is the colon where it is in a persistent, mutualistic relationship with its host. The Don system must have evolved to assist in the establishment of

a specific niche in the gut or to help the organism maintain its competitive position. Several studies have indicated that *B. fragilis* are closely associated with the mucosa, and that they penetrate the mucus layer to bind receptors deep in the crypts (23, 24). The *don* genes are significantly induced by a crude mixture of gastric mucin glycans and in fact, they were the most highly induced of all genes in mucin glycan medium (*SI Appendix, Table S1*). One potential role for the Don PUL may be to harvest the high quality complex N-linked mucus glycans (which make up about 20% of mucus glycans) (25), as cells transit through the mucus layers to the crypts. Other host-associated N-linked glycan sources may become available to bacteria adherent in the crypts, or the *don* operon may be shut down because *don* transcription is negligible in the absence of an inducer. Koropatkin et al. (3) suggested that in the colon there are many glycan microhabitats and any given *Bacteroides* spp. has evolved to respond to a specific subset of these. If we define the glycan habitat of *B. fragilis*, we may then gain further insight as to why the Don PUL potentiates the organism to so efficiently deglycosylate proteins in extraintestinal sites.

Materials and Methods

Bacterial Strains and Growth. Bacterial strains and plasmids used in this study are listed in *SI Appendix, Table S3*. *B. fragilis* 638R was the WT strain used for genetic analyses. *Bacteroides* strains were grown in an anaerobic chamber in brain heart infusion broth supplemented with hemin and cysteine (BHIS) (10). Rifampicin (20 $\mu\text{g}/\text{mL}$), gentamicin (100 $\mu\text{g}/\text{mL}$), erythromycin (10 $\mu\text{g}/\text{mL}$), and tetracycline (5 $\mu\text{g}/\text{mL}$) were added for the selection of antibiotic resistance markers. DMs were prepared as described previously with the specific carbon/energy sources described (26). Mucin glycans were prepared by proteolysis of porcine gastric mucin (Sigma-Aldrich; catalog no. M2378), followed by alkaline β -elimination to release free glycans as described in detail in *SI Appendix, Table S4* (4). In experiments using DM–transferrin, the transferrin was >98% iron-saturated holotransferrin, and the transferrin-containing media (and controls) contained 100 μM FeSO_4 to ensure a readily available source of iron for growth.

Construction of a *don* Operon-Deletion Mutant. The *donCDEFG* genes were replaced with *tetQ* tetracycline resistance (*SI Appendix, Fig. S1*). Chromosomal fragments of about 2 kb flanking the *donC* and *donE* genes were amplified using PCR primer pairs UF/UR and DF/DR respectively (*SI Appendix, Table S5*). The amplified DNA was cloned into the *Bacteroides* suicide vector pFD516 together with the *tetQ* gene cassette in a three-fragment ligation. The recombinant plasmid was mobilized into *B. fragilis* 638R by conjugation, and exconjugants were selected on BHIS plates containing rifampicin, gentamicin, and tetracycline. Candidate mutants were screened for sensitivity to erythromycin and by PCR to identify the double cross-over allelic exchange.

Animal Model of Infection. The tissue cage infection model has been described previously (27, 28). Briefly, a perforated, sterilized ping pong ball was surgically implanted into the peritoneal cavity of an adult male Sprague–Dawley rat and allowed to encapsulate for 4–5 wk. The ball became encased in connective tissue, developed a blood supply, and filled with sterile serous fluids (~25 mL per ball). Cultures grown overnight in BHIS media were diluted in PBS buffer (50 mM PO_4 , 150 mM NaCl, pH 7.4) to 10^5 cfu/mL, and 4 mL of this suspension was injected into the tissue cage. Samples were aspirated at the indicated time points for viable cell counts and RNA extraction. For viable cell counts, aliquots were serially diluted in PBS buffer, plated on BHIS plates, and incubated in an anaerobic chamber to determine cfu. Samples were also plated on LB agar plates and incubated aerobically to check for contamination.

The in vivo competition assay was performed by mixing overnight cultures of 638R and Δdon at a 1:1 ratio, followed by dilution in PBS buffer to a total viable cell count of 10^5 cfu/mL. Four milliliters of the cell mixture were injected into the tissue cage of each animal, and aliquots were plated to determine cfu per milliliter. Diluted samples were plated on BHIS plates with rifampicin and gentamicin. After 4–5 d incubation, 100–300 colonies from each sample were tested for growth on BHIS plates with and without tetracycline to check the resistance phenotype and determine the ratio of mutant to WT.

All procedures involving animals followed National Institutes of Health guidelines (29) and were approved by the Animal Care and Use Committee

of East Carolina University. For each bacterial strain tested, two trials of five animals each were performed.

Total RNA Extraction. RNA extraction from in vivo samples was done as described previously with a few modifications (27). Fluid was aspirated and immediately mixed with RNAprotect Cell Reagent (Qiagen, Inc.) at a 1:2 ratio. Next, 0.1% sodium deoxycholate was used to lyse host cells, then bacterial cells and debris were collected and washed in PBS containing RNAprotect Cell Reagent and sodium deoxycholate. Total RNA was extracted from the cell pellet using the hot phenol method described previously (30) and stored in 50% formamide at -80°C . RNA extraction from in vitro cultures was performed on washed cell pellets using the same hot phenol method.

cDNA Synthesis, qRT-PCR, and Microarray Analysis. Total RNA was purified using the RNeasy Mini Kit (Qiagen, Inc.) and DNA was removed by treatment with DNase (Ambion/Life Technologies, Inc.). DNA contamination was determined by PCR using 16S rRNA gene primers (SI Appendix, Table S5). First strand cDNA synthesis was carried out using 1 μg total RNA with random hexamer primers and SuperScript III RT (Life Technologies, Inc.). For qRT-PCR, primer pair omp117rL and omp117rR were used to amplify a 140-bp fragment of the *donC* gene in a standard reaction mixture with SYBR[®] Green Supermix (Bio-Rad, Inc.). All sample reactions were run in triplicate, and RNA with no RT was used as a control for DNA contamination. Relative expression values were normalized to 16S rRNA and calculated by the method of Pfaffl (31). Results represent at least two independent experiments performed in triplicate.

Expression microarray analyses were performed essentially as described previously (32). Double stranded cDNA was synthesized with the SuperScript[®] Double-Stranded cDNA Synthesis Kit (Life Technologies, Inc.). One microgram of ds-cDNA was labeled with cy3, hybridized to microarray slides, and processed by the Florida State University Roche/NimbleGen Microarray Facility. For each experimental condition at least two independent trials were performed. Each trial consisted of a high-density oligonucleotide whole-genome expression microarray (Roche/NimbleGen) with eight technical replicates of each probe per slide. RNA for each trial was pooled from five rats before cDNA synthesis. Raw microarray expression data were

normalized by using the robust multiarray average algorithm implemented with Roche DEVA 1.1 software. The microarray expression data have been deposited in the NCBI Gene Expression Omnibus database under accession no. GSE53883.

Whole-Cell Deglycosylation Assay. Cultures were grown in 5 mL DM–mucin glycan, harvested by centrifugation at midlogarithmic phase, and washed; and the cells were then suspended to an OD₅₅₀ of 1.0 in PBS. One hundred microliters of the cell suspension were incubated with 100 μL tissue cage serous fluid (1:10 diluted), purified human transferrin (1 mg/mL in PBS), IgG (1 mg/mL in PBS), or IgA (1 mg/mL in PBS) as described in Results and SI Appendix. All proteins were from Sigma-Aldrich Inc. For controls, 100 μL PBS replaced the bacterial cell suspension. Assays were incubated anaerobically at 37°C for 3 h or overnight, then samples were centrifuged; the supernatants collected, mixed with loading buffer, and electrophoresed on 12% SDS/PAGE gels. The gels were stained for protein, or proteins were transferred to PVDF membranes for glycan determination using SNA according to the manufacturer's instructions (Roche; DIG Glycan Differentiation Kit).

Glycomics Analyses of *B. fragilis*-Treated Human Transferrin. Purified human transferrin was suspended in PBS to 2 mg/mL, and 100 μL used in deglycosylation assays as described above except that following incubation, the assay supernatants were filter sterilized and frozen at -80°C . Samples were analyzed by Ezose Sciences Inc. to quantify N-linked glycans using previously reported methods (33, 34). Briefly, samples were denatured, digested with trypsin and then heat-inactivated. The N-glycans were then enzymatically released from the peptides by treatment with PNGase F, captured on chemoselective beads and then processed for MALDI-TOF MS. Mass spectra were analyzed by Ezose Sciences, Inc. using proprietary bioinformatics programs. Results were normalized to 1 g/L of transferrin and reported as the average of two independent experiments.

ACKNOWLEDGMENTS. We thank D. O'Rourke, A. Coburn, and M. Rosenbaum for help and advice on the animal model. This work was supported by US National Institutes of Health Grant AI40588 (to C.J.S.).

- Eckburg PB, et al. (2005) Diversity of the human intestinal microbial flora. *Science* 308(5728):1635–1638.
- Faith JJ, et al. (2013) The long-term stability of the human gut microbiota. *Science* 341(6141):1237439.
- Koropatkin NM, Cameron EA, Martens EC (2012) How glycan metabolism shapes the human gut microbiota. *Nat Rev Microbiol* 10(5):323–335.
- 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(5):447–457.
- Martens EC, Koropatkin NM, Smith TJ, Gordon JI (2009) Complex glycan catabolism by the human gut microbiota: The Bacteroidetes Sus-like paradigm. *J Biol Chem* 284(37):24673–24677.
- Lassmann B, Gustafson DR, Wood CM, Rosenblatt JE (2007) Reemergence of anaerobic bacteremia. *Clin Infect Dis* 44(7):895–900.
- Brook I (1988) Recovery of anaerobic bacteria from clinical specimens in 12 years at two military hospitals. *J Clin Microbiol* 26(6):1181–1188.
- Polk BF, Kasper DL (1977) *Bacteroides fragilis* subspecies in clinical isolates. *Ann Intern Med* 86(5):569–571.
- Krinos CM, et al. (2001) Extensive surface diversity of a commensal microorganism by multiple DNA inversions. *Nature* 414(6863):555–558.
- Sund CJ, et al. (2008) The *Bacteroides fragilis* transcriptome response to oxygen and H₂O₂: The role of OxyR and its effect on survival and virulence. *Mol Microbiol* 67(1):129–142.
- Rogers TE, et al. (2013) Dynamic responses of *Bacteroides thetaiotaomicron* during growth on glycan mixtures. *Mol Microbiol* 88(5):876–890.
- Collin M, Olseén A (2001) EndoS, a novel secreted protein from *Streptococcus pyogenes* with endoglycosidase activity on human IgG. *EMBO J* 20(12):3046–3055.
- Collin M, et al. (2002) EndoS and SpeB from *Streptococcus pyogenes* inhibit immunoglobulin-mediated opsonophagocytosis. *Infect Immun* 70(12):6646–6651.
- Burnaugh AM, Frantz LJ, King SJ (2008) Growth of *Streptococcus pneumoniae* on human glycoconjugates is dependent upon the sequential activity of bacterial exoglycosidases. *J Bacteriol* 190(1):221–230.
- King SJ, Hippe KR, Weiser JN (2006) Deglycosylation of human glycoconjugates by the sequential activities of exoglycosidases expressed by *Streptococcus pneumoniae*. *Mol Microbiol* 59(3):961–974.
- Renzi F, et al. (2011) The N-glycan glycoprotein deglycosylation complex (Gpd) from *Capnocytophaga canimorsus* deglycosylates human IgG. *PLoS Pathog* 7(6):e1002118.
- Lawrence TH III, Biggers CJ, Simonton PR (1977) Bacteriostatic inhibition of *Klebsiella pneumoniae* by three human transferrins. *Ann Hum Biol* 4(3):281–284.
- Martin CM, Jandl JH, Finland M (1963) Enhancement of acute bacterial infections in rats and mice by iron and their inhibition by human transferrin. *J Infect Dis* 112(2):158–163.
- Zhao N, Enns CA (2013) N-linked glycosylation is required for transferrin-induced stabilization of transferrin receptor 2, but not for transferrin binding or trafficking to the cell surface. *Biochemistry* 52(19):3310–3319.
- Godoy VG, Dallas MM, Russo TA, Malamy MH (1993) A role for *Bacteroides fragilis* neuraminidase in bacterial growth in two model systems. *Infect Immun* 61(10):4415–4426.
- Nakayama-Imaohji H, et al. (2012) Characterization of a gene cluster for sialoglycoconjugate utilization in *Bacteroides fragilis*. *J Med Invest* 59(1-2):79–94.
- Brigham C, et al. (2009) Sialic acid (N-acetyl neuraminic acid) utilization by *Bacteroides fragilis* requires a novel N-acetyl mannosamine epimerase. *J Bacteriol* 191(11):3629–3638.
- Lee SM, et al. (2013) Bacterial colonization factors control specificity and stability of the gut microbiota. *Nature* 501(7467):426–429.
- Namavar F, et al. (1989) Epidemiology of the *Bacteroides fragilis* group in the colonic flora in 10 patients with colonic cancer. *J Med Microbiol* 29(3):171–176.
- Derrien M, et al. (2010) Mucin-bacterial interactions in the human oral cavity and digestive tract. *Gut Microbes* 1(4):254–268.
- Varel VH, Bryant MP (1974) Nutritional features of *Bacteroides fragilis* subsp. *fragilis*. *Appl Microbiol* 28(2):251–257.
- Lobo LA, Jenkins AL, Jeffrey Smith C, Rocha ER (2013) Expression of *Bacteroides fragilis* hemolysins in vivo and role of HlyBA in an intra-abdominal infection model. *Microbiologyopen* 2(2):326–337.
- Bamberger DM, Herndon BL, Fitch J, Florkowski A, Parkhurst V (2002) Effects of neutrophils on cefazolin activity and penicillin-binding proteins in *Staphylococcus aureus* abscesses. *Antimicrob Agents Chemother* 46(9):2878–2884.
- National Research Council (US) Committee for the Update of the Guide for the Care and Use of Laboratory Animals (2011) *Guide for the Care and Use of Laboratory Animals* (National Academies Press, Washington, DC), 8th Ed.
- Rocha ER, Smith CJ (1997) Regulation of *Bacteroides fragilis* katB mRNA by oxidative stress and carbon limitation. *J Bacteriol* 179(22):7033–7039.
- Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29(9):e45.
- Ndamukong IC, Gee J, Smith CJ (2013) The extracytoplasmic function sigma factor EcfO protects *Bacteroides fragilis* against oxidative stress. *J Bacteriol* 195(1):145–155.
- Nishimura S, et al. (2004) High-throughput protein glycomics: Combined use of chemoselective glycoblotting and MALDI-TOF/TOF mass spectrometry. *Angew Chem Int Ed Engl* 44(1):91–96.
- Furukawa J, et al. (2008) Comprehensive approach to structural and functional glycomics based on chemoselective glycoblotting and sequential tag conversion. *Anal Chem* 80(4):1094–1101.