Received: 23 October 2020 | Revised: 10 January 2021 | Accepted: 11 January 2021

DOI: 10.1111/mmi.14683

MICROREVIEW

WILEY

TonB-dependent transporters in the Bacteroidetes: Unique domain structures and potential functions

Rebecca M. Pollet [1] | Lauryn M. Martin [2] | Nicole M. Koropatkin [3]

[1]Department of Biology, Davidson College, Davidson, NC, USA

[2]Department of Biology, Alcorn State University, Alcorn, MS, USA

[3]Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, MI, USA

Correspondence
Nicole M. Koropatkin, Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, MI 48109, USA.
Email: nkoropat@umich.edu

Funding information
National Institutes of Health, Grant/Award Number: R01 GM118475

Abstract

The human gut microbiota endows the host with a wealth of metabolic functions central to health, one of which is the degradation and fermentation of complex carbohydrates. The Bacteroidetes are one of the dominant bacterial phyla of this community and possess an expanded capacity for glycan utilization. This is mediated via the coordinated expression of discrete polysaccharide utilization loci (PUL) that invariantly encode a TonB-dependent transporter (SusC) that works with a glycan-capturing lipoprotein (SusD). More broadly within Gram-negative bacteria, TonB-dependent transporters (TBDTs) are deployed for the uptake of not only sugars, but also more often for essential nutrients such as iron and vitamins. Here, we provide a comprehensive look at the repertoire of TBDTs found in the model gut symbiont *Bacteroides thetaiotaomicron* and the range of predicted functional domains associated with these transporters and SusD proteins for the uptake of both glycans and other nutrients. This atlas of the *B. thetaiotaomicron* TBDTs reveals that there are at least three distinct subtypes of these transporters encoded within its genome that are presumably regulated in different ways to tune nutrient uptake.

KEYWORDS

Bacteroides thetaiotaomicron, Bacteroidetes, microbiota, polysaccharide utilization loci (PUL), TonB-dependent transporters

1 | INTRODUCTION

The human gut microbiota describes a rich community of microorganisms that influences host health. While this community includes fungi, protozoa, viruses, and bacteriophage, the most well-studied members are the Bacteria (Auchtung et al., 2018; Ding & Schloss, 2014; Shkoporov et al., 2019). While there are hundreds of species of bacteria within the mammalian large intestine, the Bacteroidetes is one of the dominant phyla of bacteria and comprise the largest number of Gram-negative organisms in the gut (Ding & Schloss, 2014; MetaHIT Consortium (additional members) et al., 2011; Tap et al., 2009). The Bacteroidetes are endowed with a prolific capacity for complex carbohydrate degradation, including the deconstruction of plant fibers from our diet as well as the host mucin layer and glycosaminoglycans (Lapébie et al., 2019). This capability is encoded within dozens to several hundred discrete operons termed polysaccharide utilization loci (PUL) (Grondin et al., 2017; Terrapon et al., 2015). The Bacteroidetes PUL encode genes for the initial capture, degradation, import, and complete hydrolysis of a target polysaccharide to its component sugars. Some PUL are highly specific for distinct glycan substructures while others can target a range of structures within a broader glycan class (Martens et al., 2014). For example, *Bacteroides ovatus* deploys multiple PUL for the recognition of different types of xylan and fine differences in structure (i.e., corn glucuronoarabinoxylan versus birch glucuronoxylan) drive differential PUL activation (Rogowski et al., 2015). Other PUL are less specific and can recognize multiple related glycan structures as seen in *Bacteroides uniformis* in which a single PUL recognizes multiple discrete β1,3-glucan structures (Déjean et al., 2020). The repertoire of PUL within a species influences its metabolic niche and fitness within the host (Backhed, 2005; Martens et al., 2011; McNulty et al., 2013).

The hallmark feature of the Bacteroidetes PUL that allows their relatively easy identification from genomic data is the presence of a pair of genes encoding a TonB-dependent transporter (TBDT) and a surface lipoprotein (Martens et al., 2009a; Xu et al., 2007). This pairing is usually called out as a *susC/susD* homolog pair after the starch utilization system operon in *Bacteroides thetaiotaomicron* (*B. theta*), the human gut bacterium from which this genetic pairing and later function was first discovered by Abigail Salyers and colleagues (Reeves et al., 1997). Early work on this system suggested that *susC* and *susD* encoded a TBDT and surface lipoprotein, respectively, that play a key role in starch utilization (Reeves et al., 1996; Shipman et al., 2000). Sequencing of the *B. theta* genome revealed at least 101 homologous *susC/D* pairings, many flanked by genes encoding predicted glycoside hydrolases, polysaccharide lyases and other accessory enzymes for carbohydrate degradation (Xu et al., 2007). Over the past two decades and as sequencing of bacterial genomes exponentially increased, the Bacteroidetes PUL for carbohydrate uptake was established, with the SusC/D proteins defining a novel type of TBDT (SusC) that works with a surface lipoprotein (SusD). Since that time many PUL encoded within Bacteroidetes from the human gut, oral cavity, and environment have been biochemically and functionally studied, providing a greater appreciation of both the conserved and novel features found among these systems. The Sus paradigm for carbohydrate uptake in the Bacteroidetes and the mechanistic features of TBDTs have been the subject of several excellent and recent reviews (Bolam & van den Berg, 2018; Brown & Koropatkin, 2020; Grondin et al., 2017). However, the variety of TBDTs encoded within gut Bacteroides genomes—for polysaccharide utilization or other nutrient uptake—have not been fully explored. In this review, we detail the structural features of TBDTs across Gram-negative bacteria, with an emphasis on the novel predicted features of TBDTs found within the model human gut symbiont *Bacteroides thetaiotaomicron*. A reference catalogue of all annotated Bacteroidetes PUL can be found in the PULDB database maintained by the Carbohydrate-Active Enzymes (CAZy) group (Terrapon et al., 2015, 2018) as well as the dbCAN-PUL database (Ausland et al., 2021).

2 | GENERAL FEATURES OF TonB-DEPENDENT TRANSPORTERS

TonB-dependent transporters (TBDTs) were first identified in *Escherichia coli* K12 when mutation of the transporter now named FhuA and the associated inner membrane TonB protein caused bacteria to become resistant to infection by bacteriophage T1 (Luria & Delbruck, 1943). Later the role of FhuA and other homologous proteins as outer membrane transporters was determined (Di Masi et al., 1973; Hantke & Braun, 1975; Luckey et al., 1975; Szmelcman & Hofnung, 1975). Most of these transporters in *E. coli* are involved in the uptake of iron–siderophore complexes and vitamin B12 (Di Masi et al., 1973; Kadner et al., 1980). As homologous transporters have been characterized across Gram-negative bacteria, it has become clear that these transporters play a large role in transporting many different macromolecules across the outer membrane that are too large to diffuse via porins. Because of this key role in nutrient uptake, TBDTs are often essential for sensing and adapting to environmental signals and are associated with pathogenicity in bacteria such as *E. coli*, *Pseudomonas aeruginosa*, and *Serratia marcescens*, among others (Takase et al., 2000; Torres et al., 2001; Weakland et al., 2020).

TBDTs are powered by their interaction with TonB, an inner membrane-associated protein with a significant soluble portion that spans the periplasm (Figure 1a) (Domingo Köhler et al., 2010). Along with the inner membrane proteins ExbB and ExbD, TonB is essential for the function of TBDTs (Higgs et al., 2002; Sverzhinsky

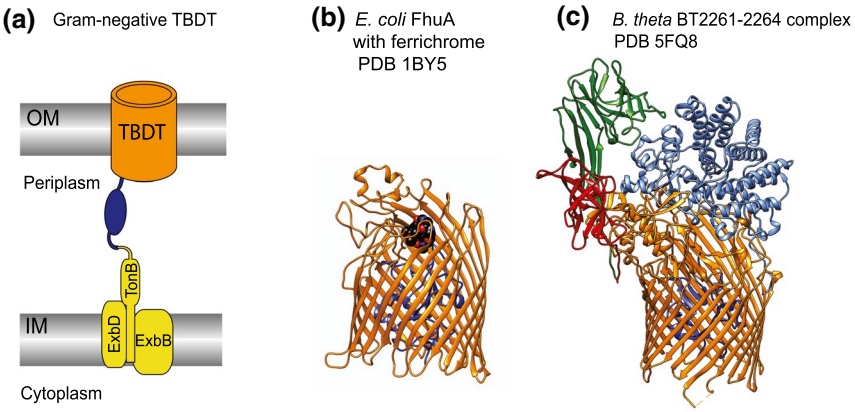

FIGURE 1 TonB-dependent transporter structure. (a) Classic architecture of the TonB-dependent transporters found in Gram-negative bacteria including pairing to the TonB/ExbB/ExbD complex. The barrel domain of the TBDT is displayed in orange and the plug domain in dark blue. (b) Structure of the *E. coli FhuA* TBDT with bound ferrichrome ligand (PDB 1BY5) (Locher et al., 1998). The barrel domain is displayed in orange and the plug domain is colored dark blue. The ferrichrome ligand is displayed in black/red/blue spheres. (c) Structure of the *B. theta* BT2261-2264 SusCD complex. The TBDT (SusC-like) protein BT2264 is displayed as in panel B with an orange barrel and dark blue plug. The SusD protein BT2263 is displayed in blue and associated PUL-encoded lipoproteins BT2261 and BT2262 are displayed in red and green, respectively. Note that only one half of the BT2261-2264 complex is displayed, as a dimeric complex has been observed via crystallography and size exclusion (Glenwright et al., 2017)

et al., 2015). The inner membrane complex formed by a pentamer of ExbB and a dimer of ExbD harnesses proton motive force that energizes the active transport process through the associated TBDT (Celia et al., 2019). At least one copy of TonB spans the periplasm, interacting with the ExbBD complex via an N-terminal membrane spanning α-helix and the TBDT via a well-ordered C-terminal domain (Celia et al., 2016; Domingo Köhler et al., 2010; Higgs et al., 2002; Sverzhinsky et al., 2015). The final β-strand of this TonB C-terminal domain directly contacts a β-strand of the TBDT called the TonB box (Kadner, 1990; Pawelek et al., 2006; Shultis et al., 2006).

Two key domains define the structure of TBDTs, a 22 β-strand barrel that spans the outer membrane (orange) and a plug domain (blue) housed within that barrel (Figure 1b,c) (Ferguson et al., 1998; Locher et al., 1998). The TonB box for pairing to TonB is found immediately N-terminal to the plug domain (Kadner, 1990; Pawelek et al., 2006; Shultis et al., 2006). The general structure of these domains is well defined with more than 50 structures having been determined of at least 15 unique transporters from many different bacteria with and without ligand (Bolam & van den Berg, 2018; Noinaj et al., 2010). For many of the iron and vitamin capturing TBDT systems, the transporter works alone to capture ligand, with the extracellular loops of the barrel and the plug domain providing specific recognition of the substrate (Locher et al., 1998; Noinaj et al., 2010). A notable exception to this is the *Neisseria* TbpAB system for the capture of transferrin iron where both the TBDT TbpA and the surface-exposed lipoprotein TbpB are required for efficient iron transport (Gómez et al., 1998). Despite this structural understanding of classical TBDTs, many details on the function of these transporters are still poorly understood including the rearrangement of the plug domain that is needed to create a channel for the ligand to pass through the transporter. Additionally, as more TBDTs are identified, it has become clear that additional domains are found in these transporters but the impact of these domains on transport is often unclear.

3 | TonB-DEPENDENT TRANSPORTERS IN THE BACTEROIDETES

Abigail Salyers proposed the presence of TBDTs in *Bacteroides* PUL as early as 1995 (Cheng et al., 1995). However, it was not until 10 years later that carbohydrates including sucrose and maltodextrins were confirmed to be transported via TBDTs from *Xanthomonas campestris* and *Caulobacter crescentus* (Blanvillain et al., 2007; Neugebauer et al., 2005). Even then, the number of proteins within this family that would be devoted to carbohydrate transport was under appreciated (Schauer et al., 2008). Additionally, surveys of TBDTs across bacteria have often noted the key differences between characterized TBDT largely from Enterobacteria and predicted TBDTs in the Bacteroidetes (Blanvillain et al., 2007; Koebnik, 2005). Today, biochemical studies of *Bacteroides* PUL and the associated TBDTs have begun to give us a better picture of how the TBDTs encoded by the Bacteroidetes share similar structure and function to the classical TBDT while also allowing for novel structures and functions (Glenwright et al., 2017; Gray et al., 2021; Madej et al., 2020). Here, we use Pfam domain analysis of the 121 predicted TBDTs in *Bacteroides thetaiotaomicron* to review what is currently known and what work remains to be done to better understand sugar transport and the broader role of TBDTs in the Bacteroidetes.

4 | STRUCTURAL DOMAINS WITHIN *B. thetaiotaomicron* TBDTs

The core structure of all TBDTs is the barrel through which the ligand will pass and the plug that occludes this channel in the absence of ligand (Figure 1) (Bolam & van den Berg, 2018; Noinaj et al., 2010). The 22 β-strand barrel domain spanning the outer membrane can be identified via homology to the Pfam domain PF00593: TonB-dependent receptor (El-Gebali et al., 2019). The plug domain is housed inside the barrel and can be identified via homology to the Pfam domain PF07715: TonB-dependent receptor plug domain (El-Gebali et al., 2019). By searching the *B. theta* VPI-5482 genome for proteins with either of these Pfams, we identified 126 potential TBDTs (search conducted July 2020). By comparing our Pfam characterization to previous annotations conducted by the CAZy database and compiled within PULDB (www.cazy.org/PULDB), we realized that the barrel domain is not always reliably identified (Terrapon et al., 2018). Manual inspection of proteins predicted to contain the Pfam plug domain but not the Pfam barrel domain suggested that most of these proteins are likely TBDTs and do contain regions homologous to the barrel domain but with substantial changes limiting the domain prediction. However, we did identify four proteins (BT1139- truncated gene, BT4200, BT4201, and BT4324) with predicted Pfam plug domains that appear divergent from characterized TBDTs and likely have another function. These proteins are also not predicted to be part of a PUL and were removed from this analysis.

That some of the proteins possessing a Pfam TonB-dependent transporter plug domain do not have a readily identifiable Pfam TonB-dependent receptor for the barrel domain may suggest that there are features within the barrel that are common in Bacteroidetes but largely absent in the TBDTs used to define this Pfam domain. Indeed, the TBDTs identified by our search demonstrate that most *B. theta* TBDTs that are devoted to glycan uptake are 200–300 amino acids longer than the well-studied TBDTs involved in iron uptake in Gram-negative pathogens. The recent crystal structures of three TBDTs from Bacteroidetes demonstrate that some of this additional sequence is devoted toward elongated loops at the extracellular face of the barrel (Figure 1c) (Glenwright et al., 2017; Gray et al., 2021; Madej et al., 2020). These loops in part engage with the ligand-binding SusD protein that caps the barrel, as described later.

Additionally, we identified one protein, BT3560, predicted to contain the Pfam barrel domain but not the Pfam plug domain. This could be the result of loss of genomic content or a miscalled start site which is common in the *B. theta* genome annotation but the protein shows low sequence similarity outside of the barrel

domain suggesting this is not the case. However, it is possible that this protein functions as another type of outer membrane transporter. Plug-less versions of *E. coli* FhuA and *B. theta* BT1763 are stable and show functional conductance in single-channel electrophysiology experiments (Glenwright et al., 2017; Mohammad et al., 2011). This suggests that BT3560 could be a functional transporter without a plug domain but is, therefore, unlikely to be TonB-dependent. Thus, this protein has also been removed from our analysis leaving us with 121 predicted TBDTs which closely matches previous analysis (Blanvillain et al., 2007; Schauer et al., 2008). The full list of predicted TBDTs along with their Pfam annotations can be found in Supplementary Table 1 and a summary of this analysis is presented in Table 1.

5 | CLASSICAL TBDT

From our list of 121 predicted TBDTs, only 16 were predicted to contain just the plug and barrel domains based on the current gene annotation (Figure 2a,b). However, this domain prediction is complicated by the recent discovery of a functional shufflon containing multiple TBDTs where recombination can result in the addition of other domains onto the TBDT genes (Porter et al., 2020). The BT1032-BT1053 locus contains recombination sites that allow the 5' end of the *bt1042* gene to be appended onto either the *bt1040* or *bt1046* genes (Porter et al., 2020). Therefore, we have placed both the BT1040 and BT1046 proteins with the BT1042 protein in the Signal Transduction TBDT subtype to be discussed later. The operon structure of BT2260-BT2268 contains a similarly positioned integrase, BT2267, which could facilitate movement of the N-terminal extension from BT2268 to either BT2260 or BT2264 although this has not been experimentally validated and was not seen in the BT2264 structure. Due to this possibility, we have tentatively removed BT2260 and BT2264 from the Classical TBDT group as well. Based on current work, it seems likely that these are novel systems and these types of rearrangements are not widespread but further work is needed to confirm the extent that recombination affects TBDT genes across the Bacteroidetes.

After this analysis for functional shufflons, we were left with 12 Classical TBDT. Seven of these are not predicted to be part of a PUL, are not associated with a SusD protein, and likely transport iron and vitamins including BT2390, which has been characterized as a thiamine transporter (Figure 2b, Table 1) (Costliow & Degnan, 2017). Similarly, five of the 12 TBDTs are predicted to be part of PUL and are associated with a SusD protein including BT0268 which has been shown to be part of an arabinogalactan-responsive PUL (Martens et al., 2011; Schwalm et al., 2016).

The presence of these 12 transporters suggests that additional domains are not required for TBDTs to properly interact with the *B. theta* TonB homologs or otherwise for proper TBDT function. However, even within this small group of proteins, there is a large size distribution (Table 1, Supplementary Table 1). The TBDTs not associated with PUL are smaller, less than 800 amino acids, suggesting they may show a similar structure to other characterized iron, B12, and thiamine transporters. However, the TBDTs associated with PUL and SusD proteins are larger with BT2032 being the largest at 955 amino acids. Larger TBDTs have previously been associated with transport of larger and more complex substrates such as iron capture from plant ferredoxin by *Pectobacterium* FusA (863 amino acids) and from human transferrin by *Neisseria* TbpA (915 amino acids) (Bolam & van den Berg, 2018; Gómez et al., 1998; Grinter et al., 2016). It has been hypothesized that longer outer membrane loops associated with the barrel domain assist in capturing these complex substrates and our observations

TABLE 1 Subclasses of *B. theta* TonB-dependent Transporters (TBDT)

Type of predicted TBDT (Figure 2)	Total number in *B. thetaiotaomicron*	In PUL?	Typical substrates	Protein length range (amino acids)	Examples highlighted in text (genome locus tag)
Classical	12	Yes- 5	Arabinogalactan; unknown	898–955	BT0286
		No- 7	Thiamine, iron, B12	613–799	BT2390
N-terminal extension	90	Yes- 82	Plant polysaccharides	938–1120	SusC (BT3702)
					BT1763
					BT2264*
		No- 8	Ferric iron, B12	700–953	BT0150
					BT1799
Signal Transduction	19	Yes- 19	Host-associated glycans	897–1182	BT0754
					BT1040/42/46
					BT4044
					BT4357
					BT4634
		No- 0			

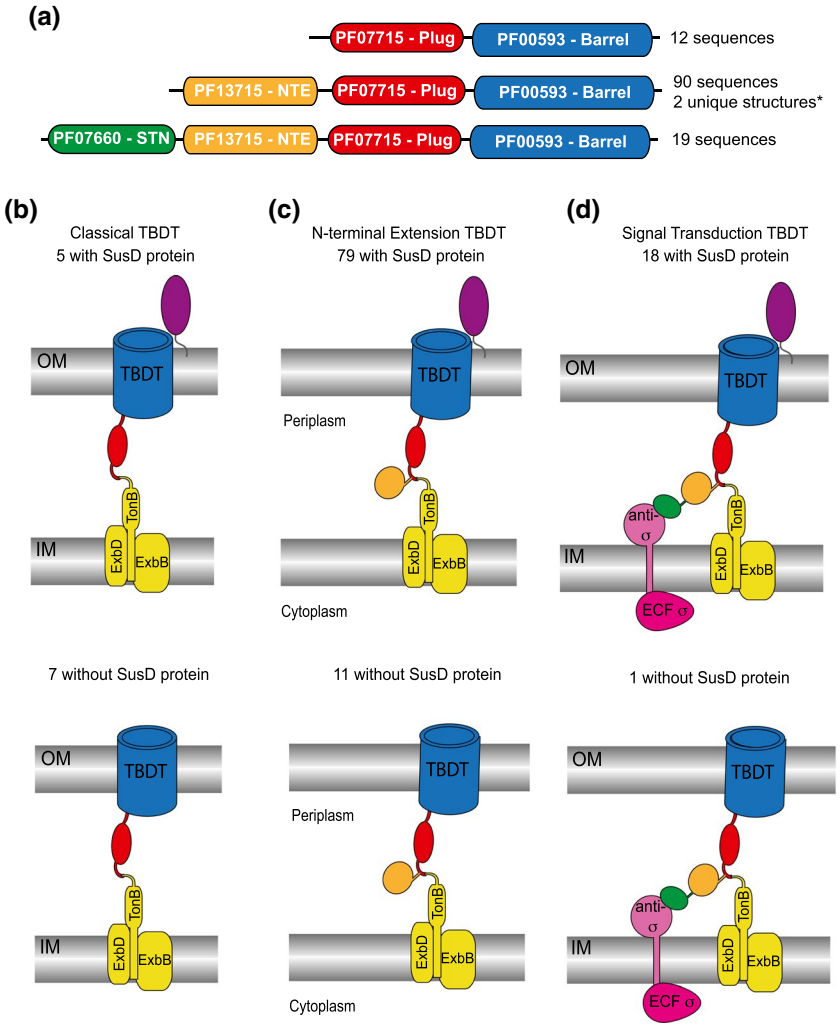

FIGURE 2 Pfam architectures and subclasses of TonB-dependent transporters (TBDTs) from *Bacteroides thetaiotaomicron*. (a) Domain architectures of TonB-dependent transporters identified in *B. theta* using the Pfam 33.1 database (https://pfam.xfam.org/). (b) Classical TBDT composed of the Pfam 07715-plug and Pfam 00593-barrel are found with and without a SusD protein. (c) N-terminal Extension TBDT have the addition of the Pfam 13715 domain and are found with and without SusD proteins. Note that we have tentatively placed BT2264 into the NTE subclass, as described in the text, and the crystal structure of this TBDT has been determined. D. Signal Transduction TBDT interact with anti-sigma factors and are found with and without SusD proteins

suggests that this trend continues in *B. theta* as PUL-associated TBDTs have been characterized to target complex polysaccharides. The association with a SusD protein or other lipoprotein may also require additional length to facilitate the TBDT-lipoprotein-substrate interaction as seen in *Neisseria* TbpA (Bolam & van den Berg, 2018; Gómez et al., 1998).

A surprising feature of the BT2264 and BT1763 structures and subsequent characterization is that these TBDTs are dimers, both in the crystal structures and in size-exclusion experiments (Glenwright et al., 2017; Gray et al., 2021). The structure of the TBDT/SusD complex RagAB from the oral Bacteroidetes *Porphyromonas gingivalis* also revealed this complex to be a dimer, which suggests that we may continue to see this trend among the TBDTs of the Bacteroidetes. A dimeric complex has not been seen in any of the previously characterized TBDTs from Gram-negative bacteria, and its specific functional role in transport within these examples is unknown (Bolam & van den Berg, 2018). It is possible that the additional length of the Bacteroides TBDTs that work with SusD proteins in part contributes to features of the barrel that allow for dimerization. Further characterization of these TBDTs will help to elucidate the mechanism of dimer complex formation.

6 | N-TERMINAL EXTENSION TBDT

Of the 121 likely TBDTs identified in *B. theta*, the remaining 109 TBDTs not classified as classical TBDT are predicted to include a PF13715: carboxypeptidase D regulatory-like domain in addition to the barrel and plug domain. This domain is found N-terminal to the plug domain and the TonB box and is, therefore, often referred to as an N-terminal extension (orange in Figure 2a,c). This domain, also annotated as a DUF4480 domain, is found in the well-characterized TBDTs from *B. theta* including SusC and BT1763 (Glenwright et al., 2017). Despite the prevalence of the N-terminal extension domain, the function of this domain is unknown. Recent characterization of the BT1763 transporter showed that this domain is essential for proper function of the transporter as *B. theta* is not able to grow on the cognate substrate levan when this domain is removed from BT1763 (Gray et al., 2021). Structural characterization of this domain revealed a small, well-structured Ig-like fold (Gray et al., 2021). It has been hypothesized that this domain might be important for TBDT pairing to TonB. This is an appealing proposal as the TonB box of the TBDTs is found between the plug domain and this PF13715 domain, putting the PF13715 domain in optimal position for interacting with TonB or the ExbBD inner

membrane complex (Figure 2a,c). Additionally, *B. theta* encodes at least 10 TonB homologs and the specificity of interactions between specific TBDTs and TonB homologs is not known (Bolam & van den Berg, 2018; Xu, 2003). Interestingly, two TonB homologs, BT3192 and BT4460, are predicted to contain PF13715 domains in addition to the TonB domain further suggesting that this domain may play a role in TBDT-TonB pairing.

There are 90 of the 121 predicted TBDT composed of just the PF13715-CarboxypepD_reg-like, PF07715- plug, and PF00593-TonB-dependent receptor domains (Figure 2a,c). Due to the lack of functional characterization of this domain we have termed transporters with this domain architecture as N-terminal extension (NTE) TBDT. As previously noted, BT2260, BT2264, and BT2268 represented a special case where BT2260 and BT2264 may gain the NTE domain from the *bt2268* gene through recombination although this has not been experimentally confirmed.

This domain architecture is found in transporters associated with PUL as well as TBDTs not predicted to be found within PUL including a transporter likely to be involved in ferric iron transport (BT0150) and one predicted to be involved in B12 transport (BT1799) (Table 1). Like the classical TBDTs, there is a large size distribution among the TBDTs that include an NTE, though TBDTs that are not encoded within a PUL (i.e., without a downstream *susD* gene) are shorter in length (Table 1). TBDTs not associated with PUL are 700–953 amino acids long while TBDTs predicted within PUL are 938–1120 amino acids long (Table 1, Supplementary Table 1). As suggested for the classical TBDT, substrate complexity seems to be associated with TBDT amino acid length.

Interestingly, three NTE TBDTs associated with PUL are not associated with a SusD protein. Two of these proteins, BT3016 and BT3633, are at least 40 amino acids smaller than all other PUL-associated TBDT and are slightly smaller than the largest non-PUL TBDTs. Both transporters are associated with PUL that lack predicted carbohydrate-active enzymes suggesting that they may have a novel function including the potential ability of BT3016/3633 to capture substrates without a SusD protein. Alternatively, a third PUL-associated TBDT without an associated SusD protein, BT4168, is very large at 1,050 amino acids and is predicted to target the complex glycan rhamnogalacturonan I. Further characterization of these three unique transporters will shed light on if these PUL-associated TBDTs can indeed function without a SusD protein and how this activity may be related to TBDT length.

7 | SIGNAL TRANSDUCTION TBDT

The final 19 TBDT found in *B. theta* share a similar domain architecture with the NTE TBDTs but contain an additional N-terminal domain, PF07660-Secretin and TonB N-terminus short domain (Figure 2a,d). This domain is often referred to as a STN domain (Secretin and TonB N-terminus domain) and has also been referred to as an N-terminal extension because it is always found N-terminal to the TonB box. This domain has been characterized in several TBDT outside of the Bacteroidetes including *E. coli* FecA, *Serratia marcescens* HasR, and *Pseudomonas aeruginosa* FoxA. In *B. theta*, this domain is always found N-terminal to a PF13715 domain although this has not been the case for the transporters characterized from other organisms. This domain in the *Pseudomonas aeruginosa* TBDT FoxA is well characterized and removal of this domain did not impact FoxA-TonB binding (Josts et al., 2019). The structures of the STN domain of FoxA, FecA, and HasR have shown this to be a small globular domain made up of two α-helices and five β-sheets (Garcia-Herrero & Vogel, 2005; Josts et al., 2019; Malki et al., 2014). Further characterization is needed to confirm conservation of this structure in the Bacteroidetes. As noted previously, three STN TBDT, BT1040, BT1042, and BT1046, represent a special case where both the PF07660 (STN) and PF13715 (NTE) domains are seen only in the *bt1042* gene in the deposited genome sequence but can be appended on either the *bt1040* or *bt1046* genes through recombination (Porter et al., 2020). This is thought to be a novel feature of this PUL, but further study is needed to fully understand the affect of recombination on the movement of these domains and could expand this category of TBDTs.

The STN domain has been shown to function as a signaling domain important for interaction between the TBDT and the associated anti-sigma factor involved in transcriptional regulation of the transporter as shown in Figure 2d (Malki et al., 2014). Because of this role in signaling we have termed this group of transporters Signal Transduction TBDT. This function has been confirmed in five of these 19 *B. theta* TBDTs through a yeast two-hybrid screen that confirmed interaction between the STN domain and the transporter's associated anti-sigma factor (Martens et al., 2009b). Interestingly, despite this signaling domain being found in transporters with a wide range of substrates in other bacteria, in *B. theta* the Signal Transduction TBDT are found only within predicted PUL. Fourteen of the 19 PUL containing Signal Transduction TBDT are predicted to target host glycans and the remaining five do not have predicted substrates. Transporters with predicted substrates include BT0754 which has been characterized to target sulfated host glycans, and BT1040, BT1042, BT1046, and BT4404 which target complex N-glycans (Benjdia et al., 2011; Briliūtė et al., 2019). Additionally, BT4357 and BT4634 are transcriptionally activated in the presence of *O*-glycans and glycosaminoglycans (Pudlo et al., 2015). Host-glycan-associated PUL are generally repressed in the presence of other polysaccharides and this unique interaction between the STN domain of the TBDT and anti-sigma factor may contribute to this important level of transcriptional regulation in *B. theta* (Pudlo et al., 2015; Rogers et al., 2013). Strikingly, despite the complex nature of host glycans, one Signal Transduction TBDT, BT2172, within a predicted host-glycan-targeting PUL is not associated with a SusD protein and is 200 amino acids shorter than all other Signal Transduction TBDT. This PUL displays a unique gene arrangement in addition to the lack of a SusD protein and may represent a novel functioning PUL and TBDT.

8 | ROLE OF THE SusD PROTEINS IN BACTEROIDETES

About 101 of the TBDTs in *B. theta* are encoded as part of a TBDT-*susD* pair, which is a key genetic marker for the identification of the Bacteroidetes PUL (Terrapon et al., 2015; Xu et al., 2007). The SusD protein encoded within the starch utilization system locus of *B. theta* was the first in this protein family to be functionally characterized and validated as binding starch (Koropatkin et al., 2008; Shipman et al., 2000). Early work on the *B. theta* Sus provided direct evidence of an interaction between the SusC and SusD proteins, which has been validated in the recent crystal structures from homologous systems such as the levan BT1763/2 and the peptide-targeting BT2264/3 complexes (Glenwright et al., 2017; Gray et al., 2021; Madej et al., 2020; Shipman et al., 2000).

It is noteworthy here that unlike the TBDTs that are ubiquitous across Gram-negative Bacteria, SusD proteins have not been characterized or described outside of the context of Bacteroidetes PUL. Bioinformatically, the SusD proteins are identified within four Pfams (Figure 3) (El-Gebali et al., 2019). The two largest families as of September 2020 are PF07980: SusD_RagB with 19,002 sequences and PF14322: SusD-like_3 with 18,457 sequences. There are 29 different contexts in which the SusD_RagB family has been reported within the Pfam database, though the majority of the sequences (17,935) are annotated as having the C-terminal half (~200–250 amino acids) falling within the SusD_RagB family and the N-terminal portion of the sequence belonging to the SusD-like_3 family (Figure 3a). Related to these families are PF12741: SusD-like and PF12771: SusD-like_2, which have significantly fewer family members. The predominant architectures that SusD-related Pfam proteins have been assigned is displayed in Figure 3a, along with the current number of sequences as reported by the Pfam database (analysis performed September 2020) (El-Gebali et al., 2019).

Currently, 43 unique protein structures have been reported in the PDB that fall within one of the four Pfam SusD groupings and are validated or likely to be SusD proteins in Bacteroidetes. Regardless of their membership within discrete Pfams, all of these proteins share distinct architectural features. SusD proteins are typically 450–650 amino acids and contain eight tetratricopeptide repeat (TPR) domains that form a right-handed superhelix that scaffolds the rest of the structure (Figure 3b) (Bolam & Koropatkin, 2012). These TPR domains dominate the N-terminus of the structure while the

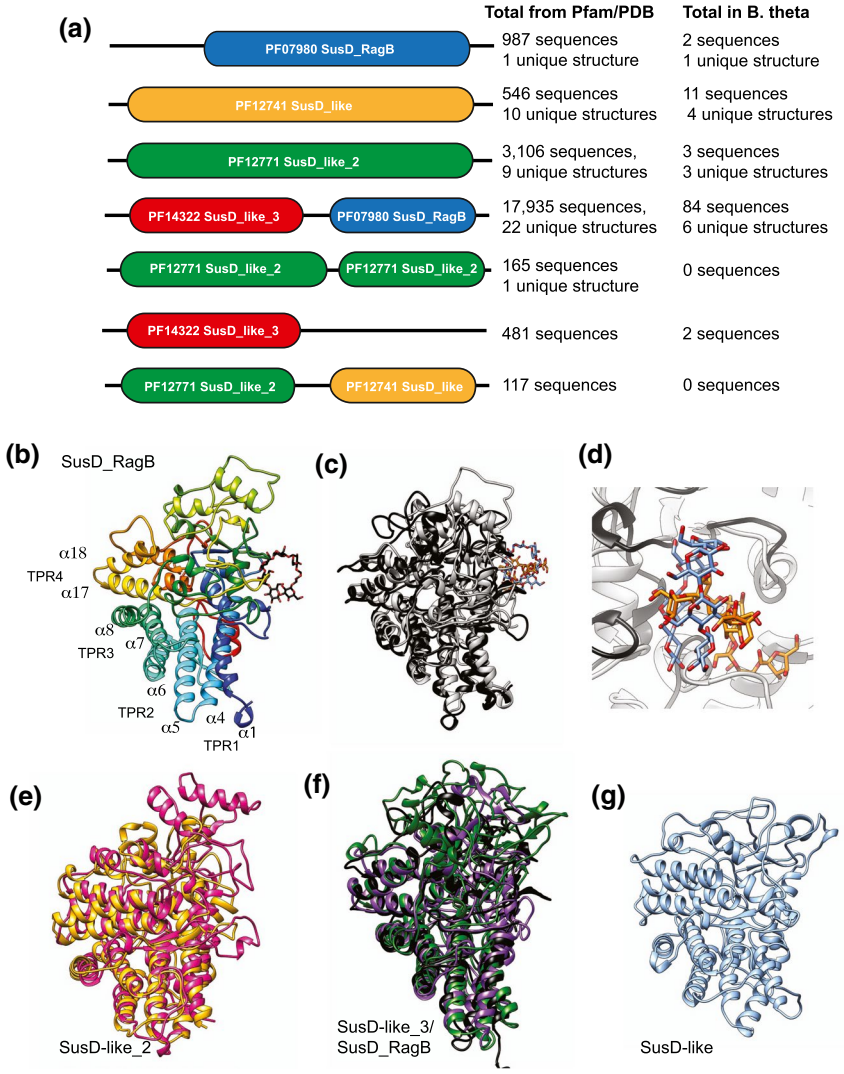

FIGURE 3 Pfam architectures and structure for SusD proteins. (a) Dominant architectures of proteins with the four families for SusD within the Pfam 33.1 database (https://pfam.xfam.org/). (b) Representative SusD_RagB Pfam member *B. theta* SusD (551 residues, PDB 3CK9). Bound maltoheptaose is displayed in blue and red sticks. (c) Overlay of *B. theta* SusD (551 residues, PDB 3CK9, gray, maltoheptaose as blue and red sticks) with *B. theta* BT1762 (570 residues, PDB 5LX8, black, fructooligosaccharide as yellow and red sticks) (d) Close-up of the glycan-binding pockets of SusD and BT1762 from panel C to better highlight the conservation of binding site location (e). Two representative members of the SusD-like_2 Pfam, a metagenomic SusD homolog (560 residues, PDB 6DK2, pink) and *B. theta* BT2263 (480 residues, PDB 5FQ4, yellow). (f) Three representative proteins of the architecture SusD-like 3/SusD_RagB Pfam, *B. theta* BT4246 (642 residues, PDB 5CK0, green), *B. theta* BT1762 (570 residues, PDB 5LX8, black), *B. theta* BT1439 (493 residues, PDB 3SNX, purple) (g) Representative of the SusD-like Pfam, *B. theta* BT3984 (537 residues, PDB 3CGH). Note that panels C,E,F, and G display the proteins in the same orientation as SusD from panel B

C-terminal portion of the structure is more variable and houses the ligand-binding region (Figure 3b,d). Thus far determined SusD structures complexed with ligand reveal conservation of the ligand-binding location (Figure 3b-d) (Glenwright et al., 2017; Gray et al., 2021; Koropatkin et al., 2008; Larsbrink et al., 2016; Tamura et al., 2019; Tauzin et al., 2016a). How the Pfam designation matches with the presumed or known ligand preferences of the protein is unknown and beyond the scope of this review. Even within the largest SusD sequence subtype (SusD-like_3/SusD-RagB architecture) in *B. theta*, the length of the protein varies substantially, with determined crystal structures deviating by ~ 150 amino acids and most of this variation is ascribed to the C-terminal portion and not the TPR domains (Figure 3f). Moreover, within this SusD type in *B. theta* the target glycans include both host and dietary polysaccharides, and the precise features of these glycans recognized by the SusD are not known. From a structural perspective, there are no obvious differences in functional domains appended to or within the structures of currently determined SusD structures. For all thus far, the predicted ligand-binding face of these proteins resides opposite of the TPR domains and is the most variable portion of these structures (Figure 3b-g). A list of the SusD proteins from *B. theta* with determined structures is summarized in Table 2 and underscores that thus far we cannot link presumed substrate to SusD Pfam type.

Three recent structures of SusC/D transporters have revealed that the SusD protein sits like a lid over the TBDT, with its ligand-binding face toward the barrel interior (Figure 1c) (Glenwright et al., 2017; Gray et al., 2021; Madej et al., 2020). An extensive network of hydrogen-bonding interactions covering an interface area of ~3,800 Å^2 stabilizes this complex. During the transport cycle, the SusD lid is predicted to open and shut over the SusC TBDT in a pedal bin mechanism that is well supported both by the structure and molecular dynamics simulations of this interaction (Glenwright et al., 2017; Gray et al., 2021). In many studies, the binding affinity of the isolated SusD for its target glycan is relatively weak (i.e., K_d ~ 10^{-4} – 10^{-5} M) (Koropatkin et al., 2008; Tamura et al., 2019; Tauzin et al., 2016a). It is likely that synergy between the TBDT and SusD during the transport cycle enhances binding affinity, as the recent structure of the TBDT-SusD transporter for β2,6-fructans demonstrates that substrate binding spans both proteins (Gray et al., 2021). In several studies of *Bacteroides* PUL that target polysaccharide, a knockout of the *susD* gene eliminates growth on the target glycan (Cho & Salyers, 2001; Koropatkin et al., 2008; Sonnenburg et al., 2010; Tamura et al., 2019; Tauzin et al., 2016a, 2016b). However, in some cases replacement of the wild-type *susD* allele with a site-directed mutant that cannot bind the target glycan restores growth on the polysaccharide (Cameron et al., 2014; Tauzin et al., 2016b). Moreover, a recent investigation of the PUL from *Bacteroides uniformis* that targets β1,3 glucans demonstrated that the isolated SusD protein does not bind glycan (Déjean et al., 2020). However, deliberate mutation of the *susD* gene to abolish protein binding to glycan eliminates cell growth on the target glycan in some instances, as seen with the β2,6 fructan PUL of

TABLE 2 *B. theta* SusD crystal structures reported in the PDB (September 2020)

Locus tag	Substrate[a]	Pfam architecture	# amino acids	PDB[b]	Reference[c]
BT3701 (SusD)	Starch	SusD_RagB	551	3CK7	(Koropatkin et al., 2008)
BT3984	High mannose N-glycans	SusD-like	537	3CGH	(Cuskin et al., 2015)
BT3752	Mucin O-glycans	SusD-like	521	3SGH	
BT1281	Mucin O-glycans	SusD-like	531	4MRU	
BT1043	Complex N-glycans	SusD-like	546	3EHN	(Briliūtė et al., 2019; Koropatkin et al., 2009)
BT2263	Peptide/protein	SusD-like_2	498	5FQ4	(Glenwright et al., 2017)
BT2259	Unknown	SusD-like_2	488	4Q69	
BT2033	Unknown host glycan	SusD-like_2	520	3FDH	
BT4659	Heparin/heparan sulfate	SusD-like_3/SusD_RagB	557	3IHV	(Cartmell et al., 2017)
BT4246	Mucin O-glycan	SusD-like_3/SusD_RagB	642	5CJZ	
BT2365	Unknown host glycan	SusD-like_3/SusD_RagB	497	3MCX	
BT1762	Levan	SusD-like_3/SusD_RagB	570	5T3R	(Glenwright et al., 2017; Sonnenburg et al., 2010)
BT1439	Unknown	SusD-like_3/SusD_RagB	493	3SNX	
BT0273	Arabinogalactan	SusD-like_3/SusD_RagB	503	3IV0	(Cartmell et al., 2018)
BT0269	Arabinogalactan	SusD-like_3/SusD_RagB	512	3HDX	(Cartmell et al., 2018)

[a]Most substrate predictions are derived from the Martens, 2008 and the Martens, Lowe 2011 studies demonstrating transcriptional activation of the *susD* genes during *B. theta* growth on different substrates (Martens et al., 2008, 2011). In some instances, substrate binding was pursued in further studies, as listed under References. This information is also summarized in PULDB (http://www.cazy.org/PULDB/) (Terrapon et al., 2018).

[b]A single PDB accession is given as an example for each structure. In some cases, more than structure was obtained of the protein.

[c]Listed references include structure or functional studies of individual proteins or PUL.

B. theta (Gray et al., 2021) and the mixed linkage β-glucan PUL of *B. ovatus* (Tamura et al., 2019). Together, these data support a critical role for the SusD protein as part of the import cycle. A notable exception is the NanOU TBDT and SusD complex that targets sialic acid in *Bacteroides fragilis* and *Tannerella forsythia* (Phansopa et al., 2014). When expressed in *E. coli*, the NanOU complex could complement a deficiency in sialic acid uptake in a TonB-dependent manner. Here, sialic acid uptake occurred when only the TBDT NanO was expressed, though uptake was maximally efficient when expressed with the SusD homolog NanU.

How the size range of the different SusD proteins influences potential interactions with its cognate TBDT is unknown. Unlike the TBDTs that are appended with individual discrete domains, the SusD size differences cannot be readily attributed to distinct features. The 43 unique SusD structures we report here range in size from 441 to 626 amino acids (Figure 3, Table 2), and the size difference appears more distributed across the structure, instead of as distinct domains.

9 | CONCLUSIONS AND PERSPECTIVES

Characterization of TonB-dependent transporters and their cognate SusD lipoproteins is essential for fully understanding nutrient utilization by the Bacteroidetes. This is especially important for understanding glycan foraging by these bacteria in the human gut, oral cavity, and environment as this is how these bacteria establish their niche in these ecosystems. In this review, we present three different domain structures of TBDTs within the model human gut symbiont *Bacteroides thetaiotaomicron* and explore the cognate SusD lipoproteins associated with these proteins to better understand the diversity of these transporters.

There are many questions that remain about the function of the TBDTs within the Bacteroidetes, and several were laid out in detail within the excellent review by Bolam and van den Berg (2018), and therefore, we will not expand upon these here. This included the role of the dimerization for the PUL-encoded TBDTs that function with a SusD protein, how glycan binding is mediated between the TBDT plug domain and SusD protein, and the specific role of the PF13715 N-terminal extension (NTE). Here, we delineate another "flavor" of SusC/D protein pairs that includes not only the NTE, but also the signal transduction domain, which provides another mechanism for control over the function of the transporter. Additionally, we note that further investigation is needed to confirm the role of functional shufflons in altering the domains associated with these transporters as seen in the BT1032-BT1053 locus (Porter et al., 2020).

One aspect of PUL architecture that we did not review is the fact that some include more than one predicted TBDT/SusD pair. Some of these are predicted based upon putative operon structure, but for others many of the individual proteins encoded within the PUL have been functionally characterized including those that target arabinogalactan, complex N-glycans, and rhamnogalacturonan II (Briliūtė et al., 2019; Martens et al., 2011; Ndeh et al., 2017). TBDT/SusD pairs may be co-regulated within the same contiguous PUL (Luis et al., 2018), but other PUL genes are co-regulated despite separation within the genome (Briliūtė et al., 2019; Ndeh et al., 2020). What is unknown is if SusD proteins can pair with other TBDT proteins within the same PUL, besides the one that is encoded immediately upstream of the *susD* gene. Based on the large protein–protein interface between the TBDT/SusD proteins, non-cognate pairing seems unlikely, but has not been explored. Moreover, it is not known whether heterodimers could potentially form from TBDTs encoded within the same PUL.

Moving forward, an unexplored area of TBDT function in the Bacteroidetes is how the transporters pair with the TonB/ExbB/ExbD inner membrane complex. *B. theta* has ~10 TonB homologs encoded within its genome and similar numbers are found in other sequenced human gut species (Bolam & van den Berg, 2018; Xu, 2003). Whether there is discrete pairing between these TBDTs and TonBs, redundant pairing, or some combination of specific and redundant pairing is unknown. The unique sequences, structures, and mechanisms of the Bacteroidetes TBDTs represent a novel type of TonB-dependent transporter and further characterization will elucidate how these transporters contribute to nutrient uptake.

ACKNOWLEDGMENTS

Research in Ann Arbor was supported by the National Institutes of Health (NIH R01 GM118475 to N.M.K. and Diversity Supplement to R.M.P.). L.M.M. was matched with the project and supported through the National Summer Undergraduate Research Project. None of the funders had any role in study design, data collection and interpretation, or the decision to submit the work for publication. The authors would also like to thank the publicly available resources provided by the Pfam database (https://pfam.xfam.org/) and the CAZy PULDB (www.cazy.org/PULDB), as referenced within the text. Finally, we would like to acknowledge an anonymous reviewer and Eric Martens for their assistance in correctly annotating transporters that may be found in shufflons.

AUTHOR CONTRIBUTIONS

L.M.M. conducted data analysis and interpretation. R.M.P. and N.M.K. designed the review, made major contributions to data acquisition, analysis, and interpretation, and wrote the manuscript.

ORCID

Rebecca M. Pollet https://orcid.org/0000-0003-2950-6374
Nicole M. Koropatkin https://orcid.org/0000-0002-2459-3336

REFERENCES

Auchtung, T.A., Fofanova, T.Y., Stewart, C.J., Nash, A.K., Wong, M.C., Gesell, J.R., et al. (2018) Investigating colonization of the healthy adult gastrointestinal tract by fungi. *mSphere*, *3*(2), e00092-18. https://doi.org/10.1128/mSphere.00092-18

Ausland, C., Zheng, J., Yi, H., Yang, B., Li, T., Feng, X., et al. (2021) dbCAN-PUL: A database of experimentally characterized CAZyme gene clusters and their substrates. *Nucleic Acids Research*, *49*, D523–D528. https://doi.org/10.1093/nar/gkaa742

Backhed, F. (2005) Host-bacterial mutualism in the human intestine. *Science*, *307*, 1915–1920. https://doi.org/10.1126/science.1104816

Benjdia, A., Martens, E.C., Gordon, J.I. & Berteau, O. (2011) 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*. *Journal of Biological Chemistry*, *286*, 25973–25982. https://doi.org/10.1074/jbc.M111.228841

Blanvillain, S., Meyer, D., Boulanger, A., Lautier, M., Guynet, C., Denancé, N., et al. (2007) Plant carbohydrate scavenging through TonB-dependent receptors: A feature shared by phytopathogenic and aquatic bacteria. *PLoS One*, *2*, e224. https://doi.org/10.1371/journal.pone.0000224

Bolam, D.N. & Koropatkin, N.M. (2012) Glycan recognition by the Bacteroidetes Sus-like systems. *Current Opinion in Structural Biology*, *22*, 563–569. https://doi.org/10.1016/j.sbi.2012.06.006

Bolam, D.N. & van den Berg, B. (2018) TonB-dependent transport by the gut microbiota: Novel aspects of an old problem. *Current Opinion in Structural Biology*, *51*, 35–43. https://doi.org/10.1016/j.sbi.2018.03.001

Briliūtė, J., Urbanowicz, P.A., Luis, A.S., Baslé, A., Paterson, N., Rebello, O., et al. (2019) Complex N-glycan breakdown by gut Bacteroides involves an extensive enzymatic apparatus encoded by multiple co-regulated genetic loci. *Nature Microbiology*, *4*, 1571–1581. https://doi.org/10.1038/s41564-019-0466-x

Brown, H.A. & Koropatkin, N.M. (2020) Host glycan utilization within the Bacteroidetes Sus-like paradigm. *Glycobiology*, cwaa054. Epub ahead of print. https://doi.org/10.1093/glycob/cwaa054

Cameron, E.A., Kwiatkowski, K.J., Lee, B.-H., Hamaker, B.R., Koropatkin, N.M. & Martens, E.C. (2014) Multifunctional nutrient-binding proteins adapt human symbiotic bacteria for glycan competition in the gut by separately promoting enhanced sensing and catalysis. *mBio*, *5*(5), e01441–e1514. https://doi.org/10.1128/mBio.01441-14

Cartmell, A., Lowe, E.C., Baslé, A., Firbank, S.J., Ndeh, D.A., Murray, H., et al (2017) How members of the human gut microbiota overcome the sulfation problem posed by glycosaminoglycans. *Proceedings of the National Academy of Sciences of the United States of America*, *114*, 7037–7042. https://doi.org/10.1073/pnas.1704367114

Cartmell, A., Muñoz-Muñoz, J., Briggs, J.A., Ndeh, D.A., Lowe, E.C., Baslé, A., et al. (2018) A surface endogalactanase in Bacteroides thetaiotaomicron confers keystone status for arabinogalactan degradation. *Nature Microbiology*, *3*, 1314–1326. https://doi.org/10.1038/s41564-018-0258-8

Celia, H., Botos, I., Ni, X., Fox, T., De Val, N., Lloubes, R., et al. (2019) Cryo-EM structure of the bacterial Ton motor subcomplex ExbB–ExbD provides information on structure and stoichiometry. *Communications Biology*, *2*, 358. https://doi.org/10.1038/s42003-019-0604-2

Celia, H., Noinaj, N., Zakharov, S.D., Bordignon, E., Botos, I., Santamaria, M., et al (2016) Structural insight into the role of the Ton complex in energy transduction. *Nature*, *538*, 60–65. https://doi.org/10.1038/nature19757

Cheng, Q., Yu, M.C., Reeves, A.R. & Salyers, A.A. (1995) Identification and characterization of a Bacteroides gene, csuF, which encodes an outer membrane protein that is essential for growth on chondroitin sulfate. *Journal of Bacteriology*, *177*, 3721–3727. https://doi.org/10.1128/JB.177.13.3721-3727.1995

Cho, K.H. & Salyers, A.A. (2001) Biochemical analysis of interactions between outer membrane proteins that contribute to starch utilization by bacteroides thetaiotaomicron. *Journal of Bacteriology*, *183*, 7224–7230. https://doi.org/10.1128/JB.183.24.7224-7230.2001

Costliow, Z.A. & Degnan, P.H. (2017) Thiamine acquisition strategies impact metabolism and competition in the gut microbe bacteroides thetaiotaomicron. *mSystems*, *2*(5), e00116-17. https://doi.org/10.1128/mSystems.00116-17

Cuskin, F., Lowe, E.C., Temple, M.J., Zhu, Y., Cameron, E.A., Pudlo, N.A., et al (2015) Human gut bacteroidetes can utilize yeast mannan through a selfish mechanism. *Nature*, *517*, 165–169. https://doi.org/10.1038/nature13995

Déjean, G., Tamura, K., Cabrera, A., Jain, N., Pudlo, N.A., Pereira, G., et al. (2020) Synergy between cell surface glycosidases and glycan-binding proteins dictates the utilization of specific beta(1,3)-glucans by human gut bacteroides. *mBio*, *11*, e00095-20. Available from: https://doi.org/10.1128/mBio.00095-20

Di Masi, D.R., White, J.C., Schnaitman, C.A. & Bradbeer, C. (1973) Transport of vitamin B12 in *Escherichia coli*: Common receptor sites for vitamin B12 and the E colicins on the outer membrane of the cell envelope. *Journal of Bacteriology*, *115*, 506–513.Available from: https://doi.org/10.1128/JB.115.2.506-513.1973

Ding, T. & Schloss, P.D. (2014) Dynamics and associations of microbial community types across the human body. *Nature*, *509*, 357–360. Available from: https://doi.org/10.1038/nature13178

Domingo Köhler, S., Weber, A., Howard, S.P., Welte, W. & Drescher, M. (2010) The proline-rich domain of TonB possesses an extended polyproline II-like conformation of sufficient length to span the periplasm of Gram-negative bacteria. *Protein Science*, *19*, 625–630.Available from: https://doi.org/10.1002/pro.345

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., et al. (2019) The Pfam protein families database in 2019. *Nucleic Acids Research*, *47*, D427–D432.Available from: https://doi.org/10.1093/nar/gky995

Ferguson, A.D., Hofmann, E., Coulton, J.W., Diederichs, K. & Welte, W. (1998) Siderophore-mediated iron transport: Crystal structure of FhuA with bound lipopolysaccharide. *Science*, *282*, 2215–2220. Available from: https://doi.org/10.1126/science.282.5397.2215

Garcia-Herrero, A. & Vogel, H.J. (2005) Nuclear magnetic resonance solution structure of the periplasmic signalling domain of the TonB-dependent outer membrane transporter FecA from *Escherichia coli*: ECF-signalling domain of TonB-dependent receptors. *Molecular Microbiology*, *58*, 1226–1237.Available from: https://doi.org/10.1111/j.1365-2958.2005.04889.x

Glenwright, A.J., Pothula, K.R., Bhamidimarri, S.P., Chorev, D.S., Baslé, A., Firbank, S.J., et al. (2017) Structural basis for nutrient acquisition by dominant members of the human gut microbiota. *Nature*, *541*, 407–411.Available from: https://doi.org/10.1038/nature20828

Gómez, J.A., Criado, M.T. & Ferreirós, C.M. (1998) Cooperation between the components of the meningococcal transferrin receptor, TbpA and TbpB, in the uptake of transferrin iron by the 37-kDa ferric-binding protein (FbpA). *Research in Microbiology*, *149*, 381–387.Available from: https://doi.org/10.1016/S0923-2508(98)80320-3

Gray, D.A., White, J.B.R., Oluwole, A.O., Rath, P., Glenwright, A.J., Mazur, A., et al. (2021) Insights into SusCD-mediated glycan import by a prominent gut symbiont. *Nature Communications*, *12*, 1–14.Available from: https://doi.org/10.1038/s41467-020-20285-y

Grinter, R., Josts, I., Mosbahi, K., Roszak, A.W., Cogdell, R.J., Bonvin, A.M.J.J., et al. (2016) Structure of the bacterial plant-ferredoxin receptor FusA. *Nature Communications*, *7*, 13308. Available from: https://doi.org/10.1038/ncomms13308

Grondin, J.M., Tamura, K., Déjean, G., Abbott, D.W. & Brumer, H. (2017) Polysaccharide utilization loci: Fueling microbial communities. *Journal of Bacteriology*, *199*, e00860-16. Available from: https://doi.org/10.1128/JB.00860-16

Hantke, K. & Braun, V. (1975) Membrane receptor dependent iron transport in *Escherichia coli*. *FEBS Letters*, *49*, 301–305.Available from: https://doi.org/10.1016/0014-5793(75)80771-X

Higgs, P.I., Larsen, R.A. & Postle, K. (2002) Quantification of known components of the Escherichia coli TonB energy transduction system: TonB, ExbB, ExbD and FepA: TonB, ExbB, ExbD and FepA ratios. *Molecular Microbiology*, *44*, 271–281.Available from: https://doi.org/10.1046/j.1365-2958.2002.02880.x

Josts, I., Veith, K. & Tidow, H. (2019) Ternary structure of the outer membrane transporter FoxA with resolved signalling domain provides insights into TonB-mediated siderophore uptake. *eLife*, *8*, e48528. Available from: https://doi.org/10.7554/eLife.48528

Kadner, R.J. (1990) Vitamin B12 transport in *Escherichia coli*: Energy coupling between membranes. *Molecular Microbiology*, *4*, 2027–2033. Available from: https://doi.org/10.1111/j.1365-2958.1990.tb00562.x

Kadner, R.J., Heller, K., Coulton, J.W. & Braun, V. (1980) Genetic control of hydroxamate-mediated iron uptake in *Escherichia coli*. *Journal of Bacteriology*, *143*, 256–264.Available from: https://doi.org/10.1128/JB.143.1.256-264.1980

Koebnik, R. (2005) TonB-dependent trans-envelope signalling: The exception or the rule? *Trends in Microbiology*, *13*, 343–347.Available from: https://doi.org/10.1016/j.tim.2005.06.005

Koropatkin, N.M., Martens, E.C., Gordon, J.I. & Smith, T.J. (2008) Starch catabolism by a prominent human gut symbiont is directed by the recognition of amylose helices. *Structure*, *16*, 1105–1115.Available from: https://doi.org/10.1016/j.str.2008.03.017

Koropatkin, N., Martens, E.C., Gordon, J.I. & Smith, T.J. (2009) Structure of a SusD homologue, BT1043, involved in mucin *O* -glycan utilization in a prominent human gut symbiont. *Biochemistry*, *48*, 1532–1542.Available from: https://doi.org/10.1021/bi801942a

Lapébie, P., Lombard, V., Drula, E., Terrapon, N. & Henrissat, B. (2019) Bacteroidetes use thousands of enzyme combinations to break down glycans. *Nature Communications*, *10*, 2043. Available from: https://doi.org/10.1038/s41467-019-10068-5

Larsbrink, J., Zhu, Y., Kharade, S.S., Kwiatkowski, K.J., Eijsink, V.G.H., Koropatkin, N.M., et al. (2016) A polysaccharide utilization locus from Flavobacterium johnsoniae enables conversion of recalcitrant chitin. *Biotechnology for Biofuels*, *9*, 260. Available from: https://doi.org/10.1186/s13068-016-0674-z

Locher, K.P., Rees, B., Koebnik, R., Mitschler, A., Moulinier, L., Rosenbusch, J.P., et al. (1998) Transmembrane signaling across the ligand-gated FhuA receptor: Crystal structures of free and ferrichrome-bound states reveal allosteric changes. *Cell*, *95*, 771–778.Available from: https://doi.org/10.1016/s0092-8674(00)81700-6

Luckey, M., Wayne, R. & Neilands, J.B. (1975) Biochemical and biophysical research communications. *Biochemical and Biophysical Research Communications*, *64*, 7.

Luis, A.S., Briggs, J., Zhang, X., Farnell, B., Ndeh, D., Labourel, A., et al. (2018) Dietary pectic glycans are degraded by coordinated enzyme pathways in human colonic bacteroides. *Nature Microbiology*, *3*, 210–219.Available from: https://doi.org/10.1038/s41564-017-0079-1

Luria, S.E. & Delbruck, M. (1943) Mutations of bacteria from virus sensitivity to virus resistance. *Genetics*, *28*, 491–511.

Madej, M., White, J.B.R., Nowakowska, Z., Rawson, S., Scavenius, C., Enghild, J.J., et al. (2020) Structural and functional insights into oligopeptide acquisition by the RagAB transporter from *Porphyromonas gingivalis*. *Nature Microbiology*, *5*(8), 1016–1025.Available from: https://doi.org/10.1038/s41564-020-0716-y

Malki, I., Simenel, C., Wojtowicz, H., Cardoso de Amorim, G., Prochnicka-Chalufour, A., Hoos, S., et al. (2014) Interaction of a partially disordered antisigma factor with its partner, the signaling domain of the TonB-dependent transporter HasR. *PLoS One*, *9*, e89502. Available from: https://doi.org/10.1371/journal.pone.0089502

Martens, E.C., Chiang, H.C. & Gordon, J.I. (2008) Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. *Cell Host & Microbe*, *4*, 447–457.Available from: https://doi.org/10.1016/j.chom.2008.09.007

Martens, E.C., Kelly, A.G., Tauzin, A.S. & Brumer, H. (2014) The devil lies in the details: How variations in polysaccharide fine-structure impact the physiology and evolution of gut microbes. *Journal of Molecular Biology*, *426*, 3851–3865.Available from: https://doi.org/10.1016/j.jmb.2014.06.022

Martens, E.C., Koropatkin, N.M., Smith, T.J. & Gordon, J.I. (2009a) Complex glycan catabolism by the human gut microbiota: The bacteroidetes sus-like paradigm. *Journal of Biological Chemistry*, *284*, 24673–24677.Available from: https://doi.org/10.1074/jbc.R109.022848

Martens, E.C., Lowe, E.C., Chiang, H., Pudlo, N.A., Wu, M., McNulty, N.P., et al. (2011) Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. *PLoS Biology*, *9*, e1001221. Available from: https://doi.org/10.1371/journal.pbio.1001221

Martens, E.C., Roth, R., Heuser, J.E. & Gordon, J.I. (2009b) Coordinate regulation of glycan degradation and polysaccharide capsule biosynthesis by a prominent human gut symbiont. *Journal of Biological Chemistry*, *284*, 18445–18457.Available from: https://doi.org/10.1074/jbc.M109.008094

McNulty, N.P., Wu, M., Erickson, A.R., Pan, C., Erickson, B.K., Martens, E.C., et al. (2013) Effects of diet on resource utilization by a model human gut microbiota containing Bacteroides cellulosilyticus WH2, a symbiont with an extensive glycobiome. *PLoS Biology*, *11*, e1001637. Available from: https://doi.org/10.1371/journal.pbio.1001637

MetaHIT Consortium (additional members), Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., et al. (2011) Enterotypes of the human gut microbiome. *Nature*, *473*, 174–180.Available from: https://doi.org/10.1038/nature09944

Mohammad, M.M., Howard, K.R. & Movileanu, L. (2011) Redesign of a plugged β-barrel membrane protein. *Journal of Biological Chemistry*, *286*, 8000–8013.Available from: https://doi.org/10.1074/jbc.M110.197723

Ndeh, D., Baslé, A., Strahl, H., Yates, E.A., McClurgg, U.L., Henrissat, B., et al. (2020) Metabolism of multiple glycosaminoglycans by Bacteroides thetaiotaomicron is orchestrated by a versatile core genetic locus. *Nature Communications*, *11*. Available from: https://doi.org/10.1038/s41467-020-14509-4

Ndeh, D., Rogowski, A., Cartmell, A., Luis, A.S., Baslé, A., Gray, J., et al. (2017) Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature*, *544*, 65–70.Available from: https://doi.org/10.1038/nature21725

Neugebauer, H., Herrmann, C., Kammer, W., Schwarz, G., Nordheim, A. & Braun, V. (2005) ExbBD-dependent transport of maltodextrins through the novel MalA protein across the outer membrane of *Caulobacter crescentus*. *Journal of Bacteriology*, *187*, 8300–8311. Available from: https://doi.org/10.1128/JB.187.24.8300-8311.2005

Noinaj, N., Guillier, M., Barnard, T.J. & Buchanan, S.K. (2010) TonB-dependent transporters: Regulation, structure, and function. *Annual Review of Microbiology*, *64*, 43–60. https://doi.org/10.1146/annurev.micro.112408.134247

Pawelek, P.D., Croteau, N., Ng-Thow-Hing, C., Khursigara, C.M., Moiseeva, N., Allaire, M., et al. (2006) Structure of TonB in complex with FhuA, *E. coli* outer membrane receptor. *Science*, *312*, 1399–1402.Available from: https://doi.org/10.1126/science.1128057

Phansopa, C., Roy, S., Rafferty, J.B., Douglas, C.W.I., Pandhal, J., Wright, P.C., et al. (2014) Structural and functional characterization of NanU, a novel high-affinity sialic acid-inducible binding protein of oral and gut-dwelling Bacteroidetes species. *The Biochemical Journal*, *458*, 499–511.Available from: https://doi.org/10.1042/BJ20131415

Porter, N.T., Hryckowian, A.J., Merrill, B.D., Fuentes, J.J., Gardner, J.O., Glowacki, R.W.P., et al. (2020) Phase-variable capsular polysaccharides and lipoproteins modify bacteriophage susceptibility in Bacteroides thetaiotaomicron. *Nature Microbiology*, *5*, 1170–1181. Available from: https://doi.org/10.1038/s41564-020-0746-5

Pudlo, N.A., Urs, K., Kumar, S.S., German, J.B., Mills, D.A. & Martens, E.C. (2015) Symbiotic human gut bacteria with variable metabolic priorities for host mucosal glycans. *mBio*, *6*, e01282-15. Available from: https://doi.org/10.1128/mBio.01282-15

Reeves, A.R., D'Elia, J.N., Frias, J. & Salyers, A.A. (1996) A Bacteroides thetaiotaomicron outer membrane protein that is essential for utilization

of maltooligosaccharides and starch. *Journal of Bacteriology*, *178*, 823–830.Available from: https://doi.org/10.1128/JB.178.3.823-830.1996

Reeves, A.R., Wang, G.R. & Salyers, A.A. (1997) Characterization of four outer membrane proteins that play a role in utilization of starch by Bacteroides thetaiotaomicron. *Journal of Bacteriology*, *179*, 643–649. Available from: https://doi.org/10.1128/jb.179.3.643-649.1997

Rogers, T.E., Pudlo, N.A., Koropatkin, N.M., Bell, J.S.K., Moya Balasch, M., Jasker, K., et al. (2013) Dynamic responses of *Bacteroides thetaiotaomicron* during growth on glycan mixtures: *Bacteroides* responses to glycan mixtures. *Molecular Microbiology*, *88*, 876–890.Available from: https://doi.org/10.1111/mmi.12228

Rogowski, A., Briggs, J.A., Mortimer, J.C., Tryfona, T., Terrapon, N., Lowe, E.C., et al. (2015) Glycan complexity dictates microbial resource allocation in the large intestine. *Nature Communications*, *6*, 7481. Available from: https://doi.org/10.1038/ncomms8481

Schauer, K., Rodionov, D.A. & de Reuse, H. (2008) New substrates for TonB-dependent transport: Do we only see the 'tip of the iceberg'? *Trends in Biochemical Sciences*, *33*, 330–338.Available from: https://doi.org/10.1016/j.tibs.2008.04.012

Schwalm, N.D., Townsend, G.E. & Groisman, E.A. (2016) Multiple signals govern utilization of a polysaccharide in the gut bacterium *Bacteroides thetaiotaomicron*. *mBio*, *7*, e01342-16. Available from: https://doi.org/10.1128/mBio.01342-16

Shipman, J.A., Berleman, J.E. & Salyers, A.A. (2000) Characterization of four outer membrane proteins involved in binding starch to the cell surface of bacteroides thetaiotaomicron. *Journal of Bacteriology*, *182*, 5365–5372.Available from: https://doi.org/10.1128/JB.182.19.5365-5372.2000

Shkoporov, A.N., Clooney, A.G., Sutton, T.D.S., Ryan, F.J., Daly, K.M., Nolan, J.A., et al. (2019) The human gut virome is highly diverse, stable, and individual specific. *Cell Host & Microbe*, *26*, 527–541.e5. Available from: https://doi.org/10.1016/j.chom.2019.09.009

Shultis, D.D., Purdy, M.D., Banchs, C.N. & Wiener, M.C. (2006) Outer membrane active transport: Structure of the BtuB:TonB complex. *Science*, *312*, 1396–1399.Available from: https://doi.org/10.1126/science.1127694

Sonnenburg, E.D., Zheng, H., Joglekar, P., Higginbottom, S.K., Firbank, S.J., Bolam, D.N., et al. (2010) Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations. *Cell*, *141*, 1241–1252.Available from: https://doi.org/10.1016/j.cell.2010.05.005

Sverzhinsky, A., Chung, J.W., Deme, J.C., Fabre, L., Levey, K.T., Plesa, M., et al. (2015) Membrane protein complex $ExbB_4$-$ExbD_1$-$TonB_1$ from Escherichia coli demonstrates conformational plasticity. *Journal of Bacteriology*, *197*, 1873–1885.Available from: https://doi.org/10.1128/JB.00069-15

Szmelcman, S. & Hofnung, M. (1975) Maltose transport in *Escherichia coli* K-12: Involvement of the bacteriophage lambda receptor. *Journal of Bacteriology*, *124*, 112–118.Available from: https://doi.org/10.1128/JB.124.1.112-118.1975

Takase, H., Nitanai, H., Hoshino, K. & Otani, T. (2000) Requirement of the *Pseudomonas aeruginosa* tonB Gene For High-Affinity Iron Acquisition And Infection. *Infection and Immunity*, *68*, 4498–4504. Available from: https://doi.org/10.1128/IAI.68.8.4498-4504.2000

Tamura, K., Foley, M.H., Gardill, B.R., Dejean, G., Schnizlein, M., Bahr, C.M.E., et al. (2019) Surface glycan-binding proteins are essential for cereal beta-glucan utilization by the human gut symbiont *Bacteroides ovatus*. *Cellular and Molecular Life Sciences*, *76*, 4319–4340.Available from: https://doi.org/10.1007/s00018-019-03115-3

Tap, J., Mondot, S., Levenez, F., Pelletier, E., Caron, C., Furet, J.-P., et al. (2009) Towards the human intestinal microbiota phylogenetic core. *Environmental Microbiology*, *11*, 2574–2584.Available from: https://doi.org/10.1111/j.1462-2920.2009.01982.x

Tauzin, A.S., Kwiatkowski, K.J., Orlovsky, N.I., Smith, C.J., Creagh, A.L., Haynes, C.A., et al. (2016) Molecular dissection of xyloglucan recognition in a prominent human gut symbiont. *mBio*, *7*, e02134-15. Available from: https://doi.org/10.1128/mBio.02134-15

Tauzin, A.S., Laville, E., Xiao, Y., Nouaille, S., Le Bourgeois, P., Heux, S., et al. (2016b) Functional characterization of a gene locus from an uncultured gut *Bacteroides* conferring xylo-oligosaccharides utilization to *Escherichia coli*: Carbohydrate transporters of gut bacteria. *Molecular Microbiology*, *102*, 579–592.Available from: https://doi.org/10.1111/mmi.13480

Terrapon, N., Lombard, V., Drula, É., Lapébie, P., Al-Masaudi, S., Gilbert, H.J., et al. (2018) PULDB: The expanded database of polysaccharide utilization loci. *Nucleic Acids Research*, *46*, D677–D683.Available from: https://doi.org/10.1093/nar/gkx1022

Terrapon, N., Lombard, V., Gilbert, H.J. & Henrissat, B. (2015) Automatic prediction of polysaccharide utilization loci in bacteroidetes species. *Bioinformatics*, *31*, 647–655.Available from: https://doi.org/10.1093/bioinformatics/btu716

Torres, A.G., Redford, P., Welch, R.A. & Payne, S.M. (2001) TonB-Dependent systems of uropathogenic *Escherichia coli*: Aerobactin and heme transport and tonb are required for virulence in the mouse. *Infection and Immunity*, *69*, 6179–6185.Available from: https://doi.org/10.1128/IAI.69.10.6179-6185.2001

Weakland, D.R., Smith, S.N., Bell, B., Tripathi, A. & Mobley, H.L.T. (2020) The *Serratia marcescens* siderophore serratiochelin is necessary for full virulence during bloodstream infection. *Infection and Immunity*, *88*, e00117-20. Available from: https://doi.org/10.1128/IAI.00117-20

Xu, J. (2003) A genomic view of the human-bacteroides thetaiotaomicron symbiosis. *Science*, *299*, 2074–2076.Available from: https://doi.org/10.1126/science.1080029

Xu, J., Mahowald, M.A., Ley, R.E., Lozupone, C.A., Hamady, M., Martens, E.C., et al. (2007) Evolution of symbiotic bacteria in the distal human intestine. *PLoS Biology*, *5*, e156. Available from: https://doi.org/10.1371/journal.pbio.0050156

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

How to cite this article: Pollet RM, Martin LM, Koropatkin NM. TonB-dependent transporters in the Bacteroidetes: Unique domain structures and potential functions. *Mol Microbiol*. 2021;115:490–501. https://doi.org/10.1111/mmi.14683