DR. NICOLE MARIE KOROPATKIN (Orcid ID : 0000-0002-2459-3336)

Article type   : MicroReview

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 28035

[2] Department of Biology, Alcorn State University, Alcorn, MS 39096

[3] Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, MI 48109, USA

To Whom Correspondence should be addressed: N.M. Koropatkin, nkoropat@umich.edu

Running Title: TonB-Dependent Transporters in *Bacteroides thetaiotaomicron*

22

Summary

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 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.

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 and 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 and 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

54 Bacteroidetes PUL encode genes for the initial capture, degradation, import and complete hydrolysis of

55 a target polysaccharide to its component sugars. Some PUL are highly specific for distinct glycan

56 substructures while others can target a range of structures within a broader glycan class (Martens et al.,

57 2014). For example, *Bacteroides ovatus* deploys multiple PUL for the recognition of different types of

58 xylan and fine differences in structure (i.e. corn glucuronarabinoxylan vs birch glucuronxylan) drive

59 differential PUL activation (Rogowski et al., 2015). Other PUL are less specific and can recognize multiple

60 related glycan structures as seen in *Bacteroides uniformis* in which a single PUL recognizes multiple

61 discrete β1,3-glucan structures (Déjean et al., 2020). The repertoire of PUL within a species influences

62 its metabolic niche and fitness within the host (Backhed, 2005; Martens et al., 2011; McNulty et al.,

63 2013).

64   The hallmark feature of the Bacteroidetes PUL that allows their relatively easy identification

65 from genomic data is the presence of a pair of genes encoding a TonB-dependent transporter (TBDT)

66 and a surface lipoprotein (Martens et al., 2009a; Xu et al., 2007). This pairing is usually called out as a

67 *susC*/*susD* homolog pair after the s̲tarch u̲tilization s̲ystem operon in *Bacteroides thetaiotaomicron* (*B.*

68 *theta*), the human gut bacterium from which this genetic pairing and later function was first discovered

69 by Abigail Salyers and colleagues (Reeves et al., 1997). Early work on this system suggested that *susC*

70 and *susD* encoded a TBDT and surface lipoprotein respectively that play a key role in starch utilization

71 (Reeves et al., 1996; Shipman et al., 2000). Sequencing of the *B. theta* genome revealed at least 101

72 homologous susC/D pairings, many flanked by genes encoding predicted glycoside hydrolases,

73 polysaccharide lyases and other accessory enzymes for carbohydrate degradation (Xu et al., 2007). Over

74 the past two decades and as sequencing of bacterial genomes exponentially increased, the

75 Bacteroidetes PUL for carbohydrate uptake was established, with the SusC/D proteins defining a novel

76 type of TBDT (SusC) that works with a surface lipoprotein (SusD). Since that time many PUL encoded

77 within Bacteroidetes from the human gut, oral cavity and environment have been biochemically and

78 functionally studied, providing a greater appreciation of both the conserved and novel features found

79 among these systems. The Sus paradigm for carbohydrate uptake in the Bacteroidetes and the

80 mechanistic features of TBDTs have been the subject of several excellent and recent reviews (Bolam and

81 van den Berg, 2018; Brown and Koropatkin, 2020; Grondin et al., 2017). However, the variety of TBDTs

82 encoded within gut Bacteroides genomes – for polysaccharide utilization or other nutrient uptake –

83 have not been fully explored. In this review, we detail the structural features of TBDTs across Gram

84 negative bacteria, with an emphasis on the novel predicted features of TBDTs found within the model

85 human gut symbiont *Bacteroides thetaiotaomicron*. A reference catalogue of all annotated

86 Bacteroidetes PUL can be found in the PULDB database maintained by the CAZy (Terrapon et al., 2018,

87 2015) as well as the dbCAN-PUL database (Ausland et al., 2021).

88 **General Features of TonB-Dependent Transporters**

89   TonB-dependent transporters (TBDTs) were first identified in *Escherichia coli* K12 when

90 mutation of the transporter now named FhuA and the associated inner membrane TonB protein caused

91 bacteria to become resistant to infection by bacteriophage T1 (Luria and Delbruck, 1943). Later the role

92 of FhuA and other homologous proteins as outer-membrane transporters was determined (Di Masi et

93 al., 1973; Hantke and Braun, 1975; Luckey et al., 1975; Szmelcman and Hofnung, 1975). Most of these

94 transporters in *E. coli* are involved in the uptake of iron-siderophore complexes and vitamin B12 (Di Masi

95 et al., 1973; Kadner et al., 1980). As homologous transporters have been characterized across Gram

96 negative bacteria, it has become clear that these transporters play a large role in transporting many

97 different macromolecules across the outer membrane that are too large to diffuse via porins. Because of

98 this key role in nutrient uptake, TBDTs are often essential for sensing and adapting to environmental

99 signals and are associated with pathogenicity in bacteria such as *E. coli, Pseudomonas aeruginosa* and

100 *Serratia marcescens*, among others (Takase et al., 2000; Torres et al., 2001; Weakland et al., 2020).

101   TBDTs are powered by their interaction with TonB, an inner membrane-associated protein with

102 a significant soluble portion that spans the periplasm (Figure 1A) (Domingo Köhler et al., 2010). Along

103 with the inner membrane proteins ExbB and ExbD, TonB is essential for the function of TBDTs (Higgs et

104 al., 2002; Sverzhinsky et al., 2015). The inner membrane complex formed by a pentamer of ExbB and a

105 dimer of ExbD harnesses proton motive force that energizes the active transport process through the

106 associated TBDT (Celia et al., 2019). At least one copy of TonB spans the periplasm, interacting with the

107 ExbBD complex via an N-terminal membrane spanning α-helix and the TBDT via a well-ordered C-

108 terminal domain (Celia et al., 2016; Domingo Köhler et al., 2010; Higgs et al., 2002; Sverzhinsky et al.,

109 2015). The final β-strand of this TonB C-terminal domain directly contacts a β-strand of the TBDT called

110 the TonB box (Kadner, 1990; Pawelek et al., 2006; Shultis et al., 2006).

111   Two key domains define the structure of TBDTs, a 22 β-strand barrel that spans the outer

112 membrane (orange) and a plug domain (blue) housed within that barrel (Figure 1B,C) (Ferguson et al.,

113 1998; Locher et al., 1998). The TonB box for pairing to TonB is found immediately N-terminal to the plug

114 domain (Kadner, 1990; Pawelek et al., 2006; Shultis et al., 2006). The general structure of these domains

115 is well defined with more than 50 structures having been determined of at least 15 unique transporters

116 from many different bacteria with and without ligand (Bolam and van den Berg, 2018; Noinaj et al.,

117 2010). For many of the iron and vitamin capturing TBDT systems, the transporter works alone to capture

118   ligand, with the extracellular loops of the barrel and the plug domain providing specific recognition of

119   the substrate (Locher et al., 1998; Noinaj et al., 2010). A notable exception to this is the *Neisseria* TbpAB

120   system for the capture of transferrin iron where both the TBDT TbpA and the surface-exposed

121   lipoprotein TbpB are required for efficient iron transport (Gómez et al., 1998). Despite this structural

122   understanding of classical TBDTs, many details on the function of these transporters are still poorly

123   understood including the rearrangement of the plug domain that is needed to create a channel for the

124   ligand to pass through the transporter. Additionally, as more TBDTs are identified, it has become clear

125   that additional domains are found in these transporters but the impact of these domains on transport is

126   often unclear.

127   **TonB-Dependent Transporters in the Bacteroidetes**

128   Abigail Salyers proposed the presence of TBDTs in *Bacteroides* PUL as early as 1995 (Cheng et al.,

129   1995). However, it was not until ten years later that carbohydrates including sucrose and maltodextrins

130   were confirmed to be transported via TBDTs from *Xanthomonas campestris* and *Caulobacter crescentus*

131   (Blanvillain et al., 2007; Neugebauer et al., 2005). Even then, the number of proteins within this family

132   that would be devoted to carbohydrate transport was under appreciated (Schauer et al., 2008).

133   Additionally, surveys of TBDTs across bacteria have often noted key differences between characterized

134   TBDT largely from Enterobacteria and predicted TBDTs in the Bacteroidetes (Blanvillain et al., 2007;

135   Koebnik, 2005). Today, biochemical studies of *Bacteroides* PUL and the associated TBDTs have begun to

136   give us a better picture of how the TBDTs encoded by the Bacteroidetes share similar structure and

137   function to the classical TBDT while also allowing for novel structures and functions (Glenwright et al.,

138   2017; Gray et al., 2021; Madej et al., 2020). Here we use Pfam domain analysis of the 121 predicted

139   TBDTs in *Bacteroides thetaiotaomicron* to review what is currently known and what work remains to be

140   done to better understand sugar transport and the broader role of TBDTs in the Bacteroidetes.

141   **Structural domains within *B. thetaiotaomicron* TBDTs**

142   The core structure of all TBDTs is the barrel through which the ligand will pass and the plug that

143   occludes this channel in the absence of ligand (Figure 1) (Bolam and van den Berg, 2018; Noinaj et al.,

144   2010). The 22 β-strand barrel domain spanning the outer membrane can be identified via homology to

145   the Pfam domain PF00593: TonB-dependent receptor (El-Gebali et al., 2019). The plug domain is housed

146   inside the barrel and can be identified via homology to the Pfam domain PF07715: TonB-dependent

147   receptor plug domain (El-Gebali et al., 2019). By searching the *B. thetaiotaomicron* VPI-5482 genome for

148   proteins with either of these Pfams, we identified 126 potential TBDTs (search conducted July 2020). By

149   comparing our Pfam characterization to previous annotations conducted by the Carbohydrate-Active

150   Enzymes (CAZy) database and compiled within PULDB (www.cazy.org/PULDB), we realized that the

151   barrel domain is not always reliably identified (Terrapon et al., 2018). Manual inspection of proteins

152   predicted to contain the Pfam plug domain but not the Pfam barrel domain suggested that most of

153   these proteins are likely TBDTs and do contain regions homologous to the barrel domain but with

154   substantial changes limiting the domain prediction. However, we did identify four proteins (BT1139-

155   truncated gene, BT4200, BT4201, and BT4324) with predicted Pfam plug domains that appear divergent

156   from characterized TBDTs and likely have another function. These proteins are also not predicted to be

157   part of a PUL and were removed from this analysis.

158   That some of the proteins possessing a Pfam TonB-dependent transporter plug domain do not

159   have a readily identifiable Pfam TonB-dependent receptor for the barrel domain may suggest that there

160   are features within the barrel that are common in Bacteroidetes but largely absent in the TBDTs used to

161   define this Pfam domain. Indeed, the TBDTs identified by our search demonstrate that most *B. theta*

162   TBDTs that are devoted to glycan uptake are 200-300 amino acids longer than the well-studied TBDTs

163   involved in iron uptake in Gram negative pathogens. The recent crystal structures of three TBDTs from

164   Bacteroidetes demonstrate that some of this additional sequence is devoted towards elongated loops at

165   the extracellular face of the barrel (Figure 1C) (Glenwright et al., 2017; Gray et al., 2021; Madej et al.,

166   2020). These loops in part engage with the ligand-binding SusD protein that caps the barrel, as described

167   later.

168   Additionally, we identified one protein, BT3560, predicted to contain the Pfam barrel domain

169   but not the Pfam plug domain. This could be the result of loss of genomic content or a miscalled start

170   site which is common in the *B. theta* genome annotation but the protein shows low sequence similarity

171   outside of the barrel domain suggesting this is not the case. However, it is possible that this protein

172   functions as another type of outer membrane transporter. Plug-less versions of *E. coli* FhuA and *B. theta*

173   BT1763 are stable and show functional conductance in single-channel electrophysiology experiments

174   (Glenwright et al., 2017; Mohammad et al., 2011). This suggests that BT3560 could be a functional

175   transporter without a plug domain but is therefore unlikely to be TonB-dependent. Thus, this protein

176   has also been removed from our analysis leaving us with 121 predicted TBDTs which closely matches

177   previous analysis (Blanvillain et al., 2007; Schauer et al., 2008). The full list of predicted TBDTs along with

178   their Pfam annotations can be found in Supplemental Table 1 and a summary of this analysis is

179   presented in Table 1.

180   **Classical TBDT**

181   From our list of 121 predicted TBDTs, only 16 were predicted to contain just the plug and barrel

182   domains based on the current gene annotation (Figure 2A, B). However, this domain prediction is

183   complicated by the recent discovery of a functional shufflon containing multiple TBDTs where

184   recombination can result in the addition of other domains onto the TBDT genes (Porter et al., 2020). The

185   BT1032-BT1053 locus contains recombination sites that allow the 5' end of the *bt1042* gene to be

186   appended onto either the *bt1040* or *bt1046* genes (Porter et al., 2020). Therefore, we have placed both

187   the BT1040 and BT1046 proteins with the BT1042 protein in the Signal Transduction TBDT subtype to be

188   discussed later. The operon structure of BT2260-BT2268 contains a similarly positioned integrase,

189   BT2267, which could facilitate movement of the N-terminal extension from BT2268 to either BT2260 or

190   BT2264 although this has not been experimentally validated and was not seen in the BT2264 structure.

191   Due to this possibility, we have tentatively removed BT2260 and BT2264 from the Classical TBDT group

192   as well. Based on current work, it seems likely that these are novel systems and these types of

193   rearrangements are not widespread but further work is needed to confirm the extent that

194   recombination affects TBDT genes across the Bacteroidetes.

195   After this analysis for functional shufflons, we were left with twelve Classical TBDT. Seven of

196   these are not predicted to be part of a PUL, are not associated with a SusD protein, and likely transport

197   iron and vitamins including BT2390 which has been characterized as a thiamine transporter (Figure 2B,

198   Table 1) (Costliow and Degnan, 2017). Similarly, five of the 12 TBDTs are predicted to be part of PUL and

199   are associated with a SusD protein including BT0268 which has been shown to be part of an

200   arabinogalactan-responsive PUL (Martens et al., 2011; Schwalm et al., 2016).

201   The presence of these 12 transporters suggests that additional domains are not required for

202   TBDTs to properly interact with the *B. theta* TonB homologues or otherwise for proper TBDT function.

203   However, even within this small group of proteins, there is a large size distribution (Table 1,

204   Supplemental Table 1). The TBDTs not associated with PUL are smaller, less than 800 amino acids,

205   suggesting they may show a similar structure to other characterized iron, B12, and thiamine

206   transporters. However, the TBDTs associated with PUL and SusD proteins are larger with BT2032 being

207   the largest at 955 amino acids. Larger TBDTs have previously been associated with transport of larger

208   and more complex substrates such as iron capture from plant ferredoxin by *Pectobacterium* FusA (863

209   amino acids) and from human transferrin by *Neisseria* TbpA (915 amino acids) (Bolam and van den Berg,

210   2018; Gómez et al., 1998; Grinter et al., 2016). It has been hypothesized that longer outer membrane

211   loops associated with the barrel domain assist in capturing these complex substrates and our

212   observations suggests that this trend continues in *B. theta* as PUL-associated TBDTs have been

213   characterized to target complex polysaccharides. The association with a SusD protein or other

214   lipoprotein may also require additional length to facilitate the TBDT-lipoprotein-substrate interaction as

215   seen in *Neisseria* TbpA (Bolam and van den Berg, 2018; Gómez et al., 1998).

216   A surprising feature of the BT2264 and BT1762 structures and subsequent characterization is

217   that these TBDTs are dimers, both in the crystal structures and in size-exclusion experiments

218   (Glenwright et al., 2017; Gray et al., 2021). The structure of the TBDT/SusD complex RagAB from the

219   oral Bacteroidetes *Porphyromonas gingivalis* also revealed this complex to be a dimer, which suggests

220   that we may continue to see this trend among the TBDTs of the Bacteroidetes. A dimeric complex has

221   not been seen in any of the previously characterized TBDTs from Gram negative bacteria, and its specific

222   functional role in transport within these examples is unknown (Bolam and van den Berg, 2018). It is

223   possible that the additional length of the Bacteroides TBDTs that work with SusD proteins in part

224   contributes to features of the barrel that allow for dimerization. Further characterization of these

225   TBDTs will help elucidate the mechanism of dimer complex formation.

226   **N-terminal Extension TBDT**

227   Of the 121 likely TBDTs identified in *B. theta*, the remaining 109 TBDTs not classified as classical

228   TBDT are predicted to include a PF13715: carboxypeptidase D regulatory-like domain in addition to the

229   barrel and plug domain. This domain is found N-terminal to the plug domain and the TonB box and is

230   therefore often referred to as an N-terminal extension (orange in Figure 2A, C). This domain, also

231   annotated as a DUF4480 domain, is found in the well characterized TBDTs from *B. theta* including SusC

232   and BT1763 (Glenwright et al., 2017). Despite the prevalence of the N-terminal extension domain, the

233   function of this domain is unknown. Recent characterization of the BT1763 transporter showed that this

234   domain is essential for proper function of the transporter as *B. theta* is not able to grow on the cognate

235   substrate levan when this domain is removed from BT1763 (Gray et al., 2021). Structural

236   characterization of this domain revealed a small, well-structured Ig-like fold (Gray et al., 2021). It has

237   been hypothesized that this domain might be important for TBDT pairing to TonB. This is an appealing

238   proposal as the TonB box of the TBDTs is found between the plug domain and this PF13715 domain,

239   putting the PF13715 domain in optimal position for interacting with TonB or the ExbBD inner membrane

240   complex (Figure 2A,C). Additionally, *B. theta* encodes at least 10 TonB homologs and the specificity of

241   interactions between specific TBDTs and TonB homologs is not known (Bolam and van den Berg, 2018;

242   Xu, 2003). Interestingly, two TonB homologs, BT3192 and BT4460, are predicted to contain PF13715

243   domains in addition to the TonB domain further suggesting that this domain may play a role in TBDT-

244   TonB pairing.

245	 There are 90 of the 121 predicted TBDT composed of just the PF13715-CarboxypepD_reg-like,

246	PF07715- plug, and PF00593-TonB-dependent receptor domains (Figure 2A,C). Due to the lack of

247	functional characterization of this domain we have termed transporters with this domain architecture as

248	N-terminal extension (NTE) TBDT. As previously noted, BT2260, BT2264, and BT2268 represented a

249	special case where BT2260 and BT2264 may gain the NTE domain from the *bt2268* gene through

250	recombination although this has not been experimentally confirmed.

251	 This domain architecture is found in transporters associated with PUL as well as TBDTs not

252	predicted to be found within PUL including a transporter likely to be involved in ferric iron transport

253	(BT0150) and one predicted to be involved in B12 transport (BT1799) (Table 1). Like the classical TBDTs,

254	there is a large size distribution among the TBDTs that include an NTE, though TBDTs that are not

255	encoded within a PUL (i.e. without a downstream *susD* gene) are shorter in length (Table 1) . TBDTs not

256	associated with PUL are 700-953 amino acids long while TBDTs predicted within PUL are 938-1120

257	amino acids long (Table 1, Supplemental Table 1). As suggested for the classical TBDT, substrate

258	complexity seems to be associated with TBDT amino acid length.

259	 Interestingly, three NTE TBDTs associated with PUL are not associated with a SusD protein. Two

260	of these proteins, BT3016 and BT3633, are at least 40 amino acids smaller than all other PUL-associated

261	TBDT and are slightly smaller than the largest non-PUL TBDTs. Both transporters are associated with PUL

262	that lack predicted carbohydrate-active enzymes suggesting that they may have a novel function

263	including the potential ability of BT3016/3633 to capture substrates without a SusD protein.

264	Alternatively, a third PUL-associated TBDT without an associated SusD protein, BT4168, is very large at

265	1050 amino acids and is predicted to target the complex glycan rhamnogalacturonan I. Further

266	characterization of these three unique transporters will shed light on if these PUL-associated TBDTs can

267	indeed function without a SusD protein and how this activity may be related to TBDT length.

268	**Signal-Transduction TBDT**

269	 The final nineteen TBDT found in *B. theta* share a similar domain architecture with the NTE

270	TBDTs but contain an additional N-terminal domain, PF07660- Secretin and TonB N-terminus short

271	domain (Figure 2A, D). This domain is often referred to as a STN domain (<u>S</u>ecretin and <u>T</u>onB <u>N</u>-terminus

272	domain) and has also been referred to as an N-terminal extension because it is always found N-terminal

273	to the TonB box. This domain has been characterized in several TBDT outside of the Bacteroidetes

274	including *E. coli* FecA, *Serratia marcescens* HasR, and *Pseudomonas aeruginosa* FoxA. In *B. theta* this

275	domain is always found N-terminal to a PF13715 domain although this has not been the case for the

276	transporters characterized from other organisms. This domain in the *Pseudomonas aeruginosa* TBDT

277   FoxA is well characterized and removal of this domain did not impact FoxA-TonB binding (Josts et al.,

278   2019). The structures of the STN domain of FoxA, FecA, and HasR have shown this to be a small globular

279   domain made up of two α-helices and five β-sheets (Garcia-Herrero and Vogel, 2005; Josts et al., 2019;

280   Malki et al., 2014). Further characterization is needed to confirm conservation of this structure in the

281   Bacteroidetes. As noted previously, three STN TBDT, BT1040, BT1042 and BT1046, represent a special

282   case where both the PF07660 (STN) and PF13715 (NTE) domains are seen only on in the *bt1042* gene in

283   the deposited genome sequence but can be appended on either the *bt1040* or *bt1046* genes through

284   recombination (Porter et al., 2020). This is thought to be a novel feature to this PUL, but further study is

285   needed to fully understand the affect of recombination on the movement of these domains and could

286   expand this category of TBDTs.

287   The STN domain has been shown to function as a signaling domain important for interaction

288   between the TBDT and the associated anti-sigma factor involved in transcriptional regulation of the

289   transporter as shown in Figure 2D (Malki et al., 2014). Because of this role in signaling we have termed

290   this group of transporters Signal Transduction TBDT. This function has been confirmed in five of these 19

291   *B. theta* TBDTs through a yeast-two hybrid screen that confirmed interaction between the STN domain

292   and the transporter's associated anti-sigma factor (Martens et al., 2009b). Interestingly, despite this

293   signaling domain being found in transporters with a wide range of substrates in other bacteria, in *B.*

294   *theta* the Signal Transduction TBDT are found only within predicted PUL. Fourteen of the nineteen PUL

295   containing Signal Transduction TBDT are predicted to target host glycans and the remaining six do not

296   have predicted substrates. This includes BT0754 which has been characterized to target sulfated host

297   glycans, and BT1040, BT1042, BT1046, and BT4404 which target complex N-glycans (Benjdia et al., 2011;

298   Briliūtė et al., 2019). Additionally, BT4357 and BT4634 are transcriptionally activated in the presence of

299   *O*-glycans and glycosaminoglycans (Pudlo et al., 2015). Host-glycan associated PUL are generally

300   repressed in the presence of other polysaccharides and this unique interaction between the STN domain

301   of the TBDT and anti-sigma factor may contribute to this important level of transcriptional regulation in

302   *B. theta* (Pudlo et al., 2015; Rogers et al., 2013). Strikingly, despite the complex nature of host glycans,

303   one Signal Transduction TBDT, BT2172, within a predicted host-glycan-targeting PUL is not associated

304   with a SusD protein and is 200 amino acids shorter than all other Signal Transduction TBDT. This PUL

305   displays a unique gene arrangement in addition to the lack of a SusD protein and may represent a novel

306   functioning PUL and TBDT.

307   **Role of the SusD proteins in Bacteroidetes**

308   101 of the TBDTs in *B. theta* are encoded as part of a TBDT-*susD* pair, which is a key genetic

309   marker for the identification of the Bacteroidetes PUL (Terrapon et al., 2015; Xu et al., 2007). The SusD

310   protein encoded within the starch utilization system locus of *B. theta* was the first in this protein family

311   to be functionally characterized and validated as binding starch (Koropatkin et al., 2008; Shipman et al.,

312   2000). Early work on the *B. theta* Sus provided direct evidence of an interaction between the SusC and

313   SusD proteins, which has been validated in the recent crystal structures from homologous systems such

314   as the levan BT1763/2 and the peptide-targeting BT2264/3 complexes (Glenwright et al., 2017; Gray et

315   al., 2021; Madej et al., 2020; Shipman et al., 2000).

316   It is noteworthy here that unlike the TBDTs that are ubiquitous across Gram negative Bacteria,

317   SusD proteins have not been characterized or described outside of the context of Bacteroidetes PUL.

318   Bioinformatically the SusD proteins are identified within four Pfams (Figure 3) (El-Gebali et al., 2019).

319   The two largest families as of September 2020 are PF07980: SusD_RagB with 19002 sequences and

320   PF14322: SusD-like_3 with 18457 sequences. There are 29 different contexts in which the SusD_RagB

321   family has been reported within the Pfam database, though the majority of the sequences (17935) are

322   annotated as having the C-terminal half (~200-250 amino acids) falling within the SusD_RagB family and

323   the N-terminal portion of the sequence belonging to the SusD-like_3 family (Figure 3A). Related to these

324   families are PF12741: SusD-like and PF12771: SusD-like_2, which have significantly fewer family

325   members. The predominant architectures that SusD-related Pfam proteins have been assigned is

326   displayed in Figure 3A, along with the current number of sequences as reported by the Pfam database

327   (analysis performed September 2020) (El-Gebali et al., 2019).

328   Currently 43 unique protein structures have been reported in the PDB that fall within one of the

329   four Pfam SusD groupings and are validated or likely to be SusD proteins in Bacteroidetes. Regardless of

330   their membership within discrete Pfams, all of these proteins share distinct architectural features. SusD

331   proteins are typically 450-650 amino acids and contain eight tetratricopeptide repeat (TPR) domains

332   that form a right-handed superhelix that scaffolds the rest of the structure (Figure 3B) (Bolam and

333   Koropatkin, 2012). These TPR domains dominate the N-terminus of the structure while the C-terminal

334   portion of the structure is more variable and houses the ligand binding region (Figure 3B,D). Thus far

335   determined SusD structures complexed with ligand reveal conservation of the ligand-binding location

336   (Figure 3B, C,D) (Glenwright et al., 2017; Gray et al., 2021; Koropatkin et al., 2008; Larsbrink et al., 2016;

337   Tamura et al., 2019; Tauzin et al., 2016a). How the Pfam designation matches with the presumed or

338   known ligand preferences of the protein is unknown and beyond the scope of this review. Even within

339   the largest SusD sequence subtype (SusD-like_3/SusD-RagB architecture) in *B. theta*, the length of the

340   protein varies substantially, with determined crystal structures deviating by ~150 amino acids and most

341   of this variation is ascribed to the C-terminal portion and not the TPR domains (Figure 3F). Moreover,

342   within this SusD type in *B. theta* the target glycans include both host and dietary polysaccharides, and

343   the precise features of these glycans recognized by the SusD are not known. From a structural

344   perspective, there are no obvious differences in functional domains appended to or within the

345   structures of currently determined SusD structures. For all thus far, the predicted ligand-binding face of

346   these proteins resides opposite of the TPR domains and is the most variable portion of these structures

347   (Figure 3B-G). A list of the SusD proteins from *B. theta* with determined structures is summarized in

348   Table 2 and underscores that thus far we cannot link presumed substrate to SusD Pfam type.

349   Three recent structures of SusC/D transporters have revealed that the SusD protein sits like a lid

350   over the TBDT, with its ligand-binding face towards the barrel interior (Figure 1C) (Glenwright et al.,

351   2017; Gray et al., 2021; Madej et al., 2020). An extensive network of hydrogen-bonding interactions

352   covering an interface area of ~3800 Å^2 stabilizes this complex. During the transport cycle, the SusD lid is

353   predicted to open and shut over the SusC TBDT in a pedal-bin mechanism that is well-supported both by

354   the structure and molecular dynamics simulations of this interaction (Glenwright et al., 2017; Gray et al.,

355   2021). In many studies, the binding affinity of the isolated SusD for its target glycan is relatively weak

356   (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

357   synergy between the TBDT and SusD during the transport cycle enhances binding affinity, as the recent

358   structure of the TBDT-SusD transporter for β2,6-fructans demonstrates that substrate binding spans

359   both proteins (Gray et al., 2021). In several studies of *Bacteroides* PUL that target polysaccharide, a

360   knockout of the *susD* gene eliminates growth on the target glycan (Cho and Salyers, 2001; Koropatkin et

361   al., 2008; Sonnenburg et al., 2010; Tamura et al., 2019; Tauzin et al., 2016a, 2016b). However, in some

362   cases replacement of the wild-type *susD* allele with a site-directed mutant that cannot bind the target

363   glycan restores growth on the polysaccharide (Cameron et al., 2014; Tauzin et al., 2016b). Moreover, a

364   recent investigation of the PUL from *Bacteroides uniformis* that targets β1,3 glucans demonstrated that

365   the isolated SusD protein does not bind glycan (Déjean et al., 2020). However, deliberate mutation of

366   the *susD* gene to abolish protein binding to glycan eliminates cell growth on the target glycan in some

367   instances, as seen with the β2,6 fructan PUL of *B. theta* (Gray et al., 2021) and the mixed linkage β-

368   glucan PUL of *B. ovatus* (Tamura et al., 2019). Together, these data support a critical role for the SusD

369   protein as part of the import cycle. A notable exception is the NanOU TBDT and SusD complex that

370   targets sialic acid in *Bacteroides fragilis* and *Tannerella forsythia* (Phansopa et al., 2014). When

371   expressed in *E. coli*, the NanOU complex could complement a deficiency in sialic acid uptake in a TonB-

372   dependent manner. Here sialic acid uptake occurred when only the TBDT NanO was expressed, though

373   uptake was maximally efficient when expressed with the SusD homolog NanU.

374   How the size range of the different SusD proteins influences potential interactions with its

375   cognate TBDT is unknown. Unlike the TBDTs that are appended with individual discrete domains, the

376   SusD size differences can't be readily attributed to distinct features. The 43 unique SusD structures we

377   report here range in size from 441 to 626 amino acids (Figure 3, Table 2), and the size difference appears

378   more distributed across the structure, instead of as distinct domains.

379   **Conclusions and Perspectives**

380   Characterization of TonB-dependent transporters and their cognate SusD lipoproteins is

381   essential for fully understanding nutrient utilization by the Bacteroidetes. This is especially important for

382   understanding glycan foraging by these bacteria in the human gut, oral cavity, and environment as this is

383   how these bacteria establish their niche in these ecosystems. In this review we present three different

384   domain structures of TBDTs within the model human gut symbiont *Bacteroides thetaiotaomicron* and

385   explore the cognate SusD lipoproteins associated with these proteins to better understand the diversity

386   of these transporters.

387   There are many questions that remain about the function of the TBDTs within the Bacteroidetes,

388   and several were laid out in detail within the excellent review by Bolam and van den Berg in 2018 and

389   therefore we will not expand upon these here. This included the role of the dimerization for the PUL-

390   encoded TBDTs that function with a SusD protein, how glycan-binding is mediated between the TBDT

391   plug domain and SusD protein, and the specific role of the PF13715 N-terminal extension (NTE). Here we

392   delineate another "flavor" of SusC/D protein pairs that includes not only the NTE but also the signal

393   transduction domain, which provides another mechanism for control over the function of the

394   transporter. Additionally, we note that further investigation is needed to confirm the role of functional

395   shufflons in altering the domains associated with these transporters as seen in the BT1032-BT1053 locus

396   (Porter et al., 2020).

397   One aspect of PUL architecture that we did not review is the fact that some include more than one

398   predicted TBDT/SusD pair. Some of these are predicted based upon putative operon structure, but for

399   others many of the individual proteins encoded within the PUL have been functionally characterized

400   including those that target arabinogalactan, complex N-glycans, and rhamnogalacturonan II (Briliūtė et

401   al., 2019; Martens et al., 2011; Ndeh et al., 2017). TBDT/SusD pairs may be co-regulated within the same

402   contiguous PUL (Luis et al., 2018), but other PUL genes are coregulated despite separation within the

403   genome (Briliūtė et al., 2019; Ndeh et al., 2020). What is unknown is if SusD proteins can pair with other

404   TBDT proteins within the same PUL, besides the one that is encoded immediately upstream of the *susD*

405   gene. Based on the large protein-protein interface between the TBDT/SusD proteins, non-cognate

406   pairing seems unlikely, but has not been explored. Moreover, it is not known whether heterodimers

407   could potentially form from TBDTs encoded within the same PUL.

408   Moving forward, an unexplored area of TBDT function in the Bacteroidetes is how the transporters

409   pair with the TonB/ExbB/ExbD inner membrane complex. *B. theta* has ~10 TonB homologs encoded

410   within its genome and similar numbers are found in other sequenced human gut species (Bolam and van

411   den Berg, 2018; Xu, 2003). Whether there is discrete pairing between these TBDTs and TonBs,

412   redundant pairing, or some combination of specific and redundant pairing is unknown. The unique

413   sequences, structures, and mechanisms of the Bacteroidetes TBDTs represent a novel type of TonB-

414   dependent transporter and further characterization will elucidate how these transporters contribute to

415   nutrient uptake.

416   **Acknowledgements**

425   **Author Contributions**

426   L.M.M. conducted data analysis and interpretation. R.M.P. and N.M.K. designed the review,

427   made major contributions to data acquisition, analysis, and interpretation, and wrote the manuscript.

428   **References**

429   Auchtung, T.A., Fofanova, T.Y., Stewart, C.J., Nash, A.K., Wong, M.C., Gesell, J.R., Auchtung, J.M., Ajami,

430   N.J., Petrosino, J.F., 2018. Investigating Colonization of the Healthy Adult Gastrointestinal Tract

431   by Fungi. mSphere 3, e00092-18, /msphere/3/2/mSphere092-18.atom.

432   https://doi.org/10.1128/mSphere.00092-18

433   Ausland, C., Zheng, J., Yi, H., Yang, B., Li, T., Feng, X., Zheng, B., Yin, Y., 2021. dbCAN-PUL: a database of
434   experimentally characterized CAZyme gene clusters and their substrates. Nucleic Acids Res. 49,
435   D523–D528. https://doi.org/10.1093/nar/gkaa742

436   Backhed, F., 2005. Host-Bacterial Mutualism in the Human Intestine. Science 307, 1915–1920.
437   https://doi.org/10.1126/science.1104816

438   Benjdia, A., Martens, E.C., Gordon, J.I., Berteau, O., 2011. Sulfatases and a Radical *S* -Adenosyl-l-
439   methionine (AdoMet) Enzyme Are Key for Mucosal Foraging and Fitness of the Prominent
440   Human Gut Symbiont, *Bacteroides thetaiotaomicron*. J. Biol. Chem. 286, 25973–25982.
441   https://doi.org/10.1074/jbc.M111.228841

442   Blanvillain, S., Meyer, D., Boulanger, A., Lautier, M., Guynet, C., Denancé, N., Vasse, J., Lauber, E., Arlat,
443   M., 2007. Plant Carbohydrate Scavenging through TonB-Dependent Receptors: A Feature Shared
444   by Phytopathogenic and Aquatic Bacteria. PLoS ONE 2, e224.
445   https://doi.org/10.1371/journal.pone.0000224

446   Bolam, D.N., Koropatkin, N.M., 2012. Glycan recognition by the Bacteroidetes Sus-like systems. Curr.
447   Opin. Struct. Biol. 22, 563–569. https://doi.org/10.1016/j.sbi.2012.06.006

448   Bolam, D.N., van den Berg, B., 2018. TonB-dependent transport by the gut microbiota: novel aspects of
449   an old problem. Curr. Opin. Struct. Biol. 51, 35–43. https://doi.org/10.1016/j.sbi.2018.03.001

450   Briliūtė, J., Urbanowicz, P.A., Luis, A.S., Baslé, A., Paterson, N., Rebello, O., Hendel, J., Ndeh, D.A., Lowe,
451   E.C., Martens, E.C., Spencer, D.I.R., Bolam, D.N., Crouch, L.I., 2019. Complex N-glycan breakdown
452   by gut Bacteroides involves an extensive enzymatic apparatus encoded by multiple co-regulated
453   genetic loci. Nat. Microbiol. 4, 1571–1581. https://doi.org/10.1038/s41564-019-0466-x

454   Brown, H.A., Koropatkin, N.M., 2020. Host glycan utilization within the Bacteroidetes Sus-like paradigm.
455   Glycobiology cwaa054. https://doi.org/10.1093/glycob/cwaa054

456   Cameron, E.A., Kwiatkowski, K.J., Lee, B.-H., Hamaker, B.R., Koropatkin, N.M., Martens, E.C., 2014.
457   Multifunctional Nutrient-Binding Proteins Adapt Human Symbiotic Bacteria for Glycan
458   Competition in the Gut by Separately Promoting Enhanced Sensing and Catalysis. mBio 5,
459   e01441-14. https://doi.org/10.1128/mBio.01441-14

460   Cartmell, A., Lowe, E.C., Baslé, A., Firbank, S.J., Ndeh, D.A., Murray, H., Terrapon, N., Lombard, V.,
461   Henrissat, B., Turnbull, J.E., Czjzek, M., Gilbert, H.J., Bolam, D.N., 2017. How members of the
462   human gut microbiota overcome the sulfation problem posed by glycosaminoglycans. Proc. Natl.
463   Acad. Sci. 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., Terrapon, N., Stott, K., Heunis, T., Gray, J., Yu, L., Dupree, P., Fernandes, P.Z., Shah, S., Williams, S.J., Labourel, A., Trost, M., Henrissat, B., Gilbert, H.J., 2018. A surface endogalactanase in Bacteroides thetaiotaomicron confers keystone status for arabinogalactan degradation. Nat. Microbiol. 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., Jiang, J., Buchanan, S.K., 2019. Cryo-EM structure of the bacterial Ton motor subcomplex ExbB–ExbD provides information on structure and stoichiometry. Commun. Biol. 2, 358. https://doi.org/10.1038/s42003-019-0604-2

Celia, H., Noinaj, N., Zakharov, S.D., Bordignon, E., Botos, I., Santamaria, M., Barnard, T.J., Cramer, W.A., Lloubes, R., Buchanan, S.K., 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. J. Bacteriol. 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 byBacteroides thetaiotaomicron. J. Bacteriol. 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, mSystems.00116-17, 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., Porter, N.T., Urs, K., Thompson, A.J., Cartmell, A., Rogowski, A., Hamilton, B.S., Chen, R., Tolbert, T.J., Piens, K., Bracke, D., Vervecken, W., Hakki, Z., Speciale, G., Munōz-Munōz, J.L., Day, A., Peña, M.J., McLean, R., Suits, M.D., Boraston, A.B., Atherly, T., Ziemer, C.J., Williams, S.J., Davies, G.J., Abbott, D.W., Martens, E.C., Gilbert, H.J., 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., Viborg, A.H., Van Petegem, F., Martens, E.C., Brumer, H., 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, /mbio/11/2/mBio.00095-20.atom. https://doi.org/10.1128/mBio.00095-20

495   Di Masi, D.R., White, J.C., Schnaitman, C.A., Bradbeer, C., 1973. Transport of Vitamin B12 in Escherichia
496   coli: Common Receptor Sites for Vitamin B12 and the E Colicins on the Outer Membrane of the
497   Cell Envelope. J. Bacteriol. 115, 506–513. https://doi.org/10.1128/JB.115.2.506-513.1973

498   Ding, T., Schloss, P.D., 2014. Dynamics and associations of microbial community types across the human
499   body. Nature 509, 357–360. https://doi.org/10.1038/nature13178

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

503   El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson, L.J.,
504   Salazar, G.A., Smart, A., Sonnhammer, E.L.L., Hirsh, L., Paladin, L., Piovesan, D., Tosatto, S.C.E.,
505   Finn, R.D., 2019. The Pfam protein families database in 2019. Nucleic Acids Res. 47, D427–D432.
506   https://doi.org/10.1093/nar/gky995

507   Ferguson, A.D., Hofmann, E., Coulton, J.W., Diederichs, K., Welte, W., 1998. Siderophore-Mediated Iron
508   Transport: Crystal Structure of FhuA with Bound Lipopolysaccharide. Science 282, 2215–2220.
509   https://doi.org/10.1126/science.282.5397.2215

510   Garcia-Herrero, A., Vogel, H.J., 2005. Nuclear magnetic resonance solution structure of the periplasmic
511   signalling domain of the TonB-dependent outer membrane transporter FecA from Escherichia
512   coli: ECF-signalling domain of TonB-dependent receptors. Mol. Microbiol. 58, 1226–1237.
513   https://doi.org/10.1111/j.1365-2958.2005.04889.x

514   Glenwright, A.J., Pothula, K.R., Bhamidimarri, S.P., Chorev, D.S., Baslé, A., Firbank, S.J., Zheng, H.,
515   Robinson, C.V., Winterhalter, M., Kleinekathöfer, U., Bolam, D.N., van den Berg, B., 2017.
516   Structural basis for nutrient acquisition by dominant members of the human gut microbiota.
517   Nature 541, 407–411. https://doi.org/10.1038/nature20828

518   Gómez, J.A., Criado, M.T., Ferreirós, C.M., 1998. Cooperation between the components of the
519   meningococcal transferrin receptor, TbpA and TbpB, in the uptake of transferrin iron by the 37-
520   kDa ferric-binding protein (FbpA). Res. Microbiol. 149, 381–387. https://doi.org/10.1016/S0923-
521   2508(98)80320-3

522   Gray, D.A., White, J.B.R., Oluwole, A.O., Rath, P., Glenwright, A.J., Mazur, A., Zahn, M., Baslé, A.,
523   Morland, C., Evans, S.L., Cartmell, A., Robinson, C.V., Hiller, S., Ranson, N.A., Bolam, D.N., Berg,
524   B. van den, 2021. Insights into SusCD-mediated glycan import by a prominent gut symbiont. Nat.
525   Commun. 12, 1–14. https://doi.org/10.1038/s41467-020-20285-y

526 Grinter, R., Josts, I., Mosbahi, K., Roszak, A.W., Cogdell, R.J., Bonvin, A.M.J.J., Milner, J.J., Kelly, S.M.,

527   Byron, O., Smith, B.O., Walker, D., 2016. Structure of the bacterial plant-ferredoxin receptor

528   FusA. Nat. Commun. 7, 13308. https://doi.org/10.1038/ncomms13308

529 Grondin, J.M., Tamura, K., Déjean, G., Abbott, D.W., Brumer, H., 2017. Polysaccharide Utilization Loci:

530   Fueling Microbial Communities. J. Bacteriol. 199, e00860-16, e00860-16.

531   https://doi.org/10.1128/JB.00860-16

532 Hantke, K., Braun, V., 1975. Membrane receptor dependent iron transport in *Escherichia coli*. FEBS Lett.

533   49, 301–305. https://doi.org/10.1016/0014-5793(75)80771-X

534 Higgs, P.I., Larsen, R.A., Postle, K., 2002. Quantification of known components of the Escherichia coli

535   TonB energy transduction system: TonB, ExbB, ExbD and FepA: TonB, ExbB, ExbD and FepA

536   ratios. Mol. Microbiol. 44, 271–281. https://doi.org/10.1046/j.1365-2958.2002.02880.x

537 Josts, I., Veith, K., Tidow, H., 2019. Ternary structure of the outer membrane transporter FoxA with

538   resolved signalling domain provides insights into TonB-mediated siderophore uptake. eLife 8,

539   e48528. https://doi.org/10.7554/eLife.48528

540 Kadner, R.J., 1990. Vitamin B12 transport in Escherichia coli: energy coupling between membranes. Mol.

541   Microbiol. 4, 2027–2033. https://doi.org/10.1111/j.1365-2958.1990.tb00562.x

542 Kadner, R.J., Heller, K., Coulton, J.W., Braun, V., 1980. Genetic Control of Hydroxamate-Mediated Iron

543   Uptake in Escherichia coli. J. Bacteriol. 143, 256–264. https://doi.org/10.1128/JB.143.1.256-

544   264.1980

545 Koebnik, R., 2005. TonB-dependent trans-envelope signalling: the exception or the rule? Trends

546   Microbiol. 13, 343–347. https://doi.org/10.1016/j.tim.2005.06.005

547 Koropatkin, N., Martens, E.C., Gordon, J.I., Smith, T.J., 2009. Structure of a SusD Homologue, BT1043,

548   Involved in Mucin *O*-Glycan Utilization in a Prominent Human Gut Symbiont [†,‡]. Biochemistry

549   48, 1532–1542. https://doi.org/10.1021/bi801942a

550 Koropatkin, N.M., Martens, E.C., Gordon, J.I., Smith, T.J., 2008. Starch Catabolism by a Prominent Human

551   Gut Symbiont Is Directed by the Recognition of Amylose Helices. Structure 16, 1105–1115.

552   https://doi.org/10.1016/j.str.2008.03.017

553 Lapébie, P., Lombard, V., Drula, E., Terrapon, N., Henrissat, B., 2019. Bacteroidetes use thousands of

554   enzyme combinations to break down glycans. Nat. Commun. 10, 2043.

555   https://doi.org/10.1038/s41467-019-10068-5

556 Larsbrink, J., Zhu, Y., Kharade, S.S., Kwiatkowski, K.J., Eijsink, V.G.H., Koropatkin, N.M., McBride, M.J.,

557   Pope, P.B., 2016. A polysaccharide utilization locus from Flavobacterium johnsoniae enables

558   conversion of recalcitrant chitin. Biotechnol. Biofuels 9, 260. https://doi.org/10.1186/s13068-

559   016-0674-z

560   Locher, K.P., Rees, B., Koebnik, R., Mitschler, A., Moulinier, L., Rosenbusch, J.P., Moras, D., 1998.

561   Transmembrane Signaling across the Ligand-Gated FhuA Receptor: Crystal Structures of Free

562   and Ferrichrome-Bound States Reveal Allosteric Changes. Cell 95, 771–778.

563   https://doi.org/10.1016/s0092-8674(00)81700-6

564   Luckey, M., Wayne, R., Neilands, J.B., 1975. BIOCHEMICAL AND BIOPHYSICAL RESEARCH

565   COMMUNICATIONS. Biochem. Biophys. Res. Commun. 64, 7.

566   Luis, A.S., Briggs, J., Zhang, X., Farnell, B., Ndeh, D., Labourel, A., Baslé, A., Cartmell, A., Terrapon, N.,

567   Stott, K., Lowe, E.C., McLean, R., Shearer, K., Schückel, J., Venditto, I., Ralet, M.-C., Henrissat, B.,

568   Martens, E.C., Mosimann, S.C., Abbott, D.W., Gilbert, H.J., 2018. Dietary pectic glycans are

569   degraded by coordinated enzyme pathways in human colonic Bacteroides. Nat. Microbiol. 3,

570   210–219. https://doi.org/10.1038/s41564-017-0079-1

571   Luria, S.E., Delbruck, M., 1943. Mutations of Bacteria from Virus Sensitivity to Virus Resistance. Genetics

572   28, 491–511.

573   Madej, M., White, J.B.R., Nowakowska, Z., Rawson, S., Scavenius, C., Enghild, J.J., Bereta, G.P., Pothula,

574   K., Kleinekathoefer, U., Baslé, A., Ranson, N.A., Potempa, J., van den Berg, B., 2020. Structural

575   and functional insights into oligopeptide acquisition by the RagAB transporter from

576   Porphyromonas gingivalis. Nat. Microbiol. https://doi.org/10.1038/s41564-020-0716-y

577   Malki, I., Simenel, C., Wojtowicz, H., Cardoso de Amorim, G., Prochnicka-Chalufour, A., Hoos, S., Raynal,

578   B., England, P., Chaffotte, A., Delepierre, M., Delepelaire, P., Izadi-Pruneyre, N., 2014.

579   Interaction of a Partially Disordered Antisigma Factor with Its Partner, the Signaling Domain of

580   the TonB-Dependent Transporter HasR. PLoS ONE 9, e89502.

581   https://doi.org/10.1371/journal.pone.0089502

582   Martens, E.C., Chiang, H.C., Gordon, J.I., 2008. Mucosal Glycan Foraging Enhances Fitness and

583   Transmission of a Saccharolytic Human Gut Bacterial Symbiont. Cell Host Microbe 4, 447–457.

584   https://doi.org/10.1016/j.chom.2008.09.007

585   Martens, E.C., Kelly, A.G., Tauzin, A.S., Brumer, H., 2014. The devil lies in the details: how variations in

586   polysaccharide fine-structure impact the physiology and evolution of gut microbes. J. Mol. Biol.

587   426, 3851–3865. https://doi.org/10.1016/j.jmb.2014.06.022

588   Martens, E.C., Koropatkin, N.M., Smith, T.J., Gordon, J.I., 2009a. Complex Glycan Catabolism by the

589   Human Gut Microbiota: The Bacteroidetes Sus-like Paradigm. J. Biol. Chem. 284, 24673–24677.

590   https://doi.org/10.1074/jbc.R109.022848

591   Martens, E.C., Lowe, E.C., Chiang, H., Pudlo, N.A., Wu, M., McNulty, N.P., Abbott, D.W., Henrissat, B.,

592   Gilbert, H.J., Bolam, D.N., Gordon, J.I., 2011. Recognition and Degradation of Plant Cell Wall

593   Polysaccharides by Two Human Gut Symbionts. PLoS Biol. 9, e1001221.

594   https://doi.org/10.1371/journal.pbio.1001221

595   Martens, E.C., Roth, R., Heuser, J.E., Gordon, J.I., 2009b. Coordinate Regulation of Glycan Degradation

596   and Polysaccharide Capsule Biosynthesis by a Prominent Human Gut Symbiont. J. Biol. Chem.

597   284, 18445–18457. https://doi.org/10.1074/jbc.M109.008094

598   McNulty, N.P., Wu, M., Erickson, A.R., Pan, C., Erickson, B.K., Martens, E.C., Pudlo, N.A., Muegge, B.D.,

599   Henrissat, B., Hettich, R.L., Gordon, J.I., 2013. Effects of Diet on Resource Utilization by a Model

600   Human Gut Microbiota Containing Bacteroides cellulosilyticus WH2, a Symbiont with an

601   Extensive Glycobiome. PLoS Biol. 11, e1001637. https://doi.org/10.1371/journal.pbio.1001637

602   MetaHIT Consortium (additional members), Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada,

603   T., Mende, D.R., Fernandes, G.R., Tap, J., Bruls, T., Batto, J.-M., Bertalan, M., Borruel, N.,

604   Casellas, F., Fernandez, L., Gautier, L., Hansen, T., Hattori, M., Hayashi, T., Kleerebezem, M.,

605   Kurokawa, K., Leclerc, M., Levenez, F., Manichanh, C., Nielsen, H.B., Nielsen, T., Pons, N.,

606   Poulain, J., Qin, J., Sicheritz-Ponten, T., Tims, S., Torrents, D., Ugarte, E., Zoetendal, E.G., Wang,

607   J., Guarner, F., Pedersen, O., de Vos, W.M., Brunak, S., Doré, J., Weissenbach, J., Ehrlich, S.D.,

608   Bork, P., 2011. Enterotypes of the human gut microbiome. Nature 473, 174–180.

609   https://doi.org/10.1038/nature09944

610   Mohammad, M.M., Howard, K.R., Movileanu, L., 2011. Redesign of a Plugged β-Barrel Membrane

611   Protein. J. Biol. Chem. 286, 8000–8013. https://doi.org/10.1074/jbc.M110.197723

612   Ndeh, D., Baslé, A., Strahl, H., Yates, E.A., McClurgg, U.L., Henrissat, B., Terrapon, N., Cartmell, A., 2020.

613   Metabolism of multiple glycosaminoglycans by Bacteroides thetaiotaomicron is orchestrated by

614   a versatile core genetic locus. Nat. Commun. 11. https://doi.org/10.1038/s41467-020-14509-4

615   Ndeh, D., Rogowski, A., Cartmell, A., Luis, A.S., Baslé, A., Gray, J., Venditto, I., Briggs, J., Zhang, X.,

616   Labourel, A., Terrapon, N., Buffetto, F., Nepogodiev, S., Xiao, Y., Field, R.A., Zhu, Y., O'Neill, M.A.,

617   Urbanowicz, B.R., York, W.S., Davies, G.J., Abbott, D.W., Ralet, M.-C., Martens, E.C., Henrissat,

618   B., Gilbert, H.J., 2017. Complex pectin metabolism by gut bacteria reveals novel catalytic

619   functions. Nature 544, 65–70. https://doi.org/10.1038/nature21725

620   Neugebauer, H., Herrmann, C., Kammer, W., Schwarz, G., Nordheim, A., Braun, V., 2005. ExbBD-

621   Dependent Transport of Maltodextrins through the Novel MalA Protein across the Outer

622   Membrane of Caulobacter crescentus. J. Bacteriol. 187, 8300–8311.

623   https://doi.org/10.1128/JB.187.24.8300-8311.2005

624   Noinaj, N., Guillier, M., Barnard, T.J., Buchanan, S.K., 2010. TonB-Dependent Transporters: Regulation,

625   Structure, and Function. Annu. Rev. Microbiol. 64, 43–60.

626   https://doi.org/10.1146/annurev.micro.112408.134247

627   Pawelek, P.D., Croteau, N., Ng-Thow-Hing, C., Khursigara, C.M., Moiseeva, N., Allaire, M., Coulton, J.W.,

628   2006. Structure of TonB in Complex with FhuA, E. coli Outer Membrane Receptor. Science 312,

629   1399–1402. https://doi.org/10.1126/science.1128057

630   Phansopa, C., Roy, S., Rafferty, J.B., Douglas, C.W.I., Pandhal, J., Wright, P.C., Kelly, D.J., Stafford, G.P.,

631   2014. Structural and functional characterization of NanU, a novel high-affinity sialic acid-

632   inducible binding protein of oral and gut-dwelling Bacteroidetes species. Biochem. J. 458, 499–

633   511. https://doi.org/10.1042/BJ20131415

634   Porter, N.T., Hryckowian, A.J., Merrill, B.D., Fuentes, J.J., Gardner, J.O., Glowacki, R.W.P., Singh, S.,

635   Crawford, R.D., Snitkin, E.S., Sonnenburg, J.L., Martens, E.C., 2020. Phase-variable capsular

636   polysaccharides and lipoproteins modify bacteriophage susceptibility in Bacteroides

637   thetaiotaomicron. Nat. Microbiol. 5, 1170–1181. https://doi.org/10.1038/s41564-020-0746-5

638   Pudlo, N.A., Urs, K., Kumar, S.S., German, J.B., Mills, D.A., Martens, E.C., 2015. Symbiotic Human Gut

639   Bacteria with Variable Metabolic Priorities for Host Mucosal Glycans. mBio 6, e01282-15.

640   https://doi.org/10.1128/mBio.01282-15

641   Reeves, A.R., D'Elia, J.N., Frias, J., Salyers, A.A., 1996. A Bacteroides thetaiotaomicron outer membrane

642   protein that is essential for utilization of maltooligosaccharides and starch. J. Bacteriol. 178,

643   823–830. https://doi.org/10.1128/JB.178.3.823-830.1996

644   Reeves, A.R., Wang, G.R., Salyers, A.A., 1997. Characterization of four outer membrane proteins that

645   play a role in utilization of starch by Bacteroides thetaiotaomicron. J. Bacteriol. 179, 643–649.

646   https://doi.org/10.1128/jb.179.3.643-649.1997

647   Rogers, T.E., Pudlo, N.A., Koropatkin, N.M., Bell, J.S.K., Moya Balasch, M., Jasker, K., Martens, E.C., 2013.

648   Dynamic responses of *B acteroides thetaiotaomicron* during growth on glycan mixtures:

649   *Bacteroides* responses to glycan mixtures. Mol. Microbiol. 88, 876–890.

650   https://doi.org/10.1111/mmi.12228

651 Rogowski, A., Briggs, J.A., Mortimer, J.C., Tryfona, T., Terrapon, N., Lowe, E.C., Baslé, A., Morland, C.,

652   Day, A.M., Zheng, H., Rogers, T.E., Thompson, P., Hawkins, A.R., Yadav, M.P., Henrissat, B.,

653   Martens, E.C., Dupree, P., Gilbert, H.J., Bolam, D.N., 2015. Glycan complexity dictates microbial

654   resource allocation in the large intestine. Nat. Commun. 6, 7481.

655   https://doi.org/10.1038/ncomms8481

656 Schauer, K., Rodionov, D.A., de Reuse, H., 2008. New substrates for TonB-dependent transport: do we

657   only see the 'tip of the iceberg'? Trends Biochem. Sci. 33, 330–338.

658   https://doi.org/10.1016/j.tibs.2008.04.012

659 Schwalm, N.D., Townsend, G.E., Groisman, E.A., 2016. Multiple Signals Govern Utilization of a

660   Polysaccharide in the Gut Bacterium *Bacteroides thetaiotaomicron*. mBio 7, e01342-16,

661   /mbio/7/5/e01342-16.atom. https://doi.org/10.1128/mBio.01342-16

662 Shipman, J.A., Berleman, J.E., Salyers, A.A., 2000. Characterization of Four Outer Membrane Proteins

663   Involved in Binding Starch to the Cell Surface ofBacteroides thetaiotaomicron. J. Bacteriol. 182,

664   5365–5372. https://doi.org/10.1128/JB.182.19.5365-5372.2000

665 Shkoporov, A.N., Clooney, A.G., Sutton, T.D.S., Ryan, F.J., Daly, K.M., Nolan, J.A., McDonnell, S.A.,

666   Khokhlova, E.V., Draper, L.A., Forde, A., Guerin, E., Velayudhan, V., Ross, R.P., Hill, C., 2019. The

667   Human Gut Virome Is Highly Diverse, Stable, and Individual Specific. Cell Host Microbe 26, 527-

668   541 e5.

669 Shultis, D.D., Purdy, M.D., Banchs, C.N., Wiener, M.C., 2006. Outer Membrane Active Transport:

670   Structure of the BtuB:TonB Complex. Science 312, 1396–1399.

671   https://doi.org/10.1126/science.1127694

672 Sonnenburg, E.D., Zheng, H., Joglekar, P., Higginbottom, S.K., Firbank, S.J., Bolam, D.N., Sonnenburg, J.L.,

673   2010. Specificity of Polysaccharide Use in Intestinal Bacteroides Species Determines Diet-

674   Induced Microbiota Alterations. Cell 141, 1241–1252. https://doi.org/10.1016/j.cell.2010.05.005

675 Sverzhinsky, A., Chung, J.W., Deme, J.C., Fabre, L., Levey, K.T., Plesa, M., Carter, D.M., Lypaczewski, P.,

676   Coulton, J.W., 2015. Membrane Protein Complex ExbB$_4$-ExbD$_1$-TonB$_1$ from Escherichia coli

677   Demonstrates Conformational Plasticity. J. Bacteriol. 197, 1873–1885.

678   https://doi.org/10.1128/JB.00069-15

679 Szmelcman, S., Hofnung, M., 1975. Maltose transport in Escherichia coli K-12: involvement of the

680   bacteriophage lambda receptor. J. Bacteriol. 124, 112–118.

681   https://doi.org/10.1128/JB.124.1.112-118.1975

682   Takase, H., Nitanai, H., Hoshino, K., Otani, T., 2000. Requirement of the Pseudomonas aeruginosa tonB
683   Gene for High-Affinity Iron Acquisition and Infection. Infect. Immun. 68, 4498–4504.
684   https://doi.org/10.1128/IAI.68.8.4498-4504.2000

685   Tamura, K., Foley, M.H., Gardill, B.R., Dejean, G., Schnizlein, M., Bahr, C.M.E., Louise Creagh, A., van
686   Petegem, F., Koropatkin, N.M., Brumer, H., 2019. Surface glycan-binding proteins are essential
687   for cereal beta-glucan utilization by the human gut symbiont Bacteroides ovatus. Cell. Mol. Life
688   Sci. 76, 4319–4340. https://doi.org/10.1007/s00018-019-03115-3

689   Tap, J., Mondot, S., Levenez, F., Pelletier, E., Caron, C., Furet, J.-P., Ugarte, E., Muñoz-Tamayo, R., Paslier,
690   D.L.E., Nalin, R., Dore, J., Leclerc, M., 2009. Towards the human intestinal microbiota
691   phylogenetic core. Environ. Microbiol. 11, 2574–2584. https://doi.org/10.1111/j.1462-
692   2920.2009.01982.x

693   Tauzin, A.S., Kwiatkowski, K.J., Orlovsky, N.I., Smith, C.J., Creagh, A.L., Haynes, C.A., Wawrzak, Z.,
694   Brumer, H., Koropatkin, N.M., 2016a. Molecular Dissection of Xyloglucan Recognition in a
695   Prominent Human Gut Symbiont. mBio 7, e02134-15, /mbio/7/2/e02134-15.atom.
696   https://doi.org/10.1128/mBio.02134-15

697   Tauzin, A.S., Laville, E., Xiao, Y., Nouaille, S., Le Bourgeois, P., Heux, S., Portais, J.-C., Monsan, P.,
698   Martens, E.C., Potocki-Veronese, G., Bordes, F., 2016b. Functional characterization of a gene
699   locus from an uncultured gut *Bacteroides* conferring xylo-oligosaccharides utilization to
700   *Escherichia coli*: Carbohydrate transporters of gut bacteria. Mol. Microbiol. 102, 579–592.
701   https://doi.org/10.1111/mmi.13480

702   Terrapon, N., Lombard, V., Drula, É., Lapébie, P., Al-Masaudi, S., Gilbert, H.J., Henrissat, B., 2018. PULDB:
703   the expanded database of Polysaccharide Utilization Loci. Nucleic Acids Res. 46, D677–D683.
704   https://doi.org/10.1093/nar/gkx1022

705   Terrapon, N., Lombard, V., Gilbert, H.J., Henrissat, B., 2015. Automatic prediction of polysaccharide
706   utilization loci in Bacteroidetes species. Bioinformatics 31, 647–655.
707   https://doi.org/10.1093/bioinformatics/btu716

708   Torres, A.G., Redford, P., Welch, R.A., Payne, S.M., 2001. TonB-Dependent Systems of Uropathogenic
709   Escherichia coli: Aerobactin and Heme Transport and TonB Are Required for Virulence in the
710   Mouse. Infect. Immun. 69, 6179–6185. https://doi.org/10.1128/IAI.69.10.6179-6185.2001

711   Weakland, D.R., Smith, S.N., Bell, B., Tripathi, A., Mobley, H.L.T., 2020. The *Serratia marcescens*
712   Siderophore Serratiochelin Is Necessary for Full Virulence during Bloodstream Infection. Infect.
713   Immun. 88, e00117-20, /iai/88/8/IAI.00117-20.atom. https://doi.org/10.1128/IAI.00117-20

714 Xu, J., 2003. A Genomic View of the Human-Bacteroides thetaiotaomicron Symbiosis. Science 299, 2074–

715 2076. https://doi.org/10.1126/science.1080029

716 Xu, J., Mahowald, M.A., Ley, R.E., Lozupone, C.A., Hamady, M., Martens, E.C., Henrissat, B., Coutinho,

717 P.M., Minx, P., Latreille, P., Cordum, H., Van Brunt, A., Kim, K., Fulton, R.S., Fulton, L.A., Clifton,

718 S.W., Wilson, R.K., Knight, R.D., Gordon, J.I., 2007. Evolution of Symbiotic Bacteria in the Distal

719 Human Intestine. PLoS Biol. 5, e156. https://doi.org/10.1371/journal.pbio.0050156

720

721

722 **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
NTE	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			

723

724

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

Locus tag	Substrate[1]	Pfam architecture	# amino acids	PDB[2]	Reference[3]
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)

726 [1] Most substrate predictions are derived from the Martens, 2008 and the Martens, Lowe 2011 studies

727 demonstrating transcriptional activation of the *susD* genes during *B. theta* growth on different

728 substrates (Martens et al., 2011, 2008). In some instances, substrate binding was pursued in further

729 studies, as listed under References. This information is also summarized in PULdb

730 (http://www.cazy.org/PULDB/) (Terrapon et al., 2018). [2]A single PDB accession is given as an example

731 for each structure. In some cases, more than structure was obtained of the protein. [3]Listed references

732 include structure or functional studies of individual proteins or PUL.

733

734 **Figure Legends**

735 **Figure 1: TonB-dependent transporter structure** A. Classic architecture of the TonB-dependent

736 transporters found in Gram negative bacteria including pairing to the TonB/ExbB/ExbD complex. The

737 barrel domain of the TBDT is displayed in orange and the plug domain in dark blue. B. Structure of the *E.*

738 *coli FhuA* TBDT with bound ferrichrome ligand (PDB 1BY5) (Locher et al., 1998). The barrel domain is

739 displayed in orange and the plug domain is colored dark blue. The ferrichrome ligand is displayed in

740 black/red/blue spheres. C. Structure of the *B. theta* BT2261-2264 SusCD complex. The TBDT (SusC-like)

741 protein BT2264 is displayed as in panel B with an orange barrel and dark blue plug. The SusD protein

742 BT2263 is displayed in blue and associated PUL-encoded lipoproteins BT2261 and BT2262 are displayed

743 in red and green respectively. Note that only one half of the BT2261-2264 complex is displayed, as a

744 dimeric complex has been observed via crystallography and size exclusion (Glenwright et al., 2017).

745

746 **Figure 2: Pfam architectures and subclasses of TonB-dependent transporters (TBDT) from *Bacteroides***

747 ***thetaiotaomicron*.** A. Domain architectures of TonB-dependent transporters identified in *B.*

748 *thetaiotaomicron* using the Pfam 33.1 database (https://pfam.xfam.org/). B. Classical TBDT composed of

749 the Pfam 07715-plug and Pfam 00593-barrel are found with and without a SusD protein. C. N-terminal

750 Extension TBDT have the addition of the Pfam 13715 domain and are found with and without SusD

751 proteins. *Note that we have tentatively placed BT2264 into the NTE subclass, as described in the text,

752 and the crystal structure of this TBDT has been determined. D. Signal Transduction TBDT interact with

753 anti-sigma factors and are found with and without SusD proteins.
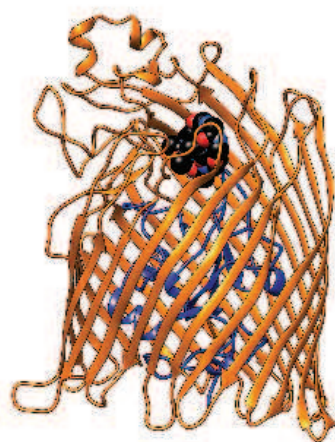
754

755 **Figure 3: Pfam architectures and structure for SusD proteins.** A. Dominant architectures proteins with

756 the four families for SusD within the Pfam 33.1 database (https://pfam.xfam.org/). B. Representative

757 SusD_RagB Pfam member *B. theta* SusD (551 residues, PDB 3CK9). Bound maltoheptaose is displayed in

758 blue and red sticks. C. Overlay of *B. theta* SusD (551 residues, PDB 3CK9, gray, maltoheptaose as blue

759 and red sticks) with *B. theta* BT1762 (570 residues, PDB 5LX8, black, fructooligosaccharide as yellow and

760 red sticks) D. Close-up of the glycan binding pockets of SusD and BT1762 from panel C to better highlight

761 the conservation of binding site location E. Two representative members of the SusD-like_2 Pfam, a

762 metagenomic SusD homolog (560 residues, PDB 6DK2, pink) and *B. theta* BT2263 (480 residues, PDB

763 5FQ4, yellow). F. Three representative proteins of the architecture SusD-like3/SusD_RagB Pfam, *B. theta*

764 BT4246 (642 residues, PDB 5CK0, green), *B. theta* BT1762 (570 residues, PDB 5LX8, black), *B. theta*

765 BT1439 (493 residues, PDB 3SNX, purple) G. Representative of the SusD-like Pfam, *B. theta* BT3984 (537

766 residues, PDB 3CGH). Note that panels B,C,E,F and G display the proteins in the same orientation as

767 SusD from panel B.

A. Gram-negative TBDT
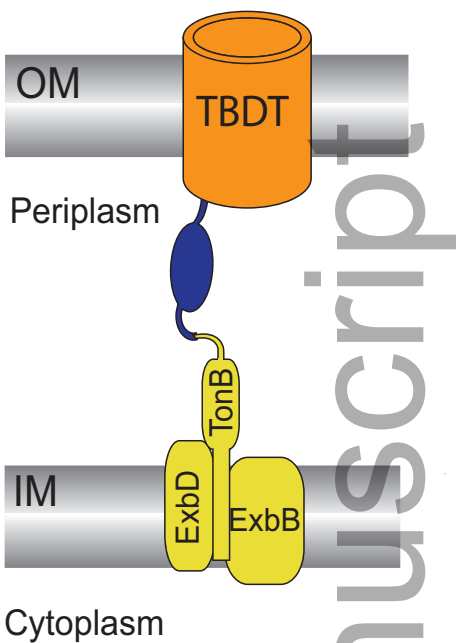
B. *E. coli* FhuA
    with ferrichrome
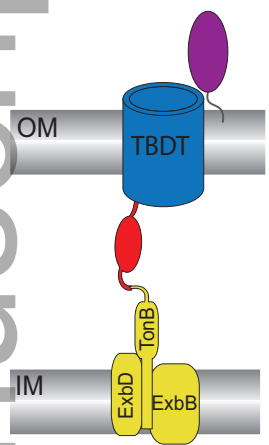   PDB 1BY5

C. *B. theta* BT2261-2264 complex
    PDB 5FQ8

OM

TBDT

Periplasm

TonB

IM

ExbD

ExbB

Cytoplasm
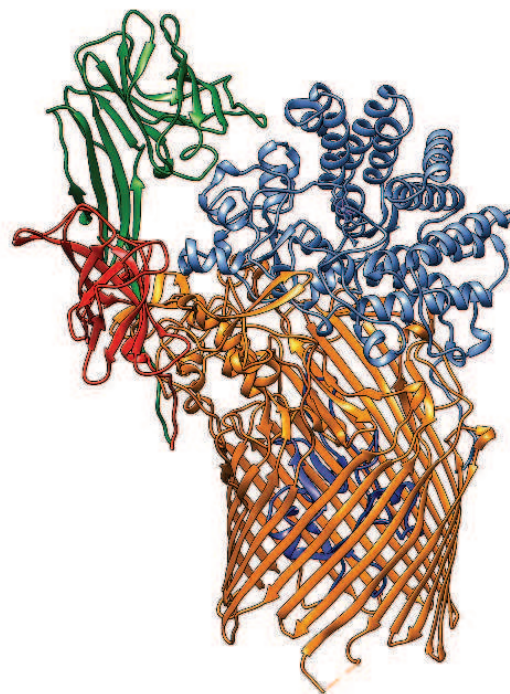
Figure 1 - Pollet, Martin & Koropatkin

mmi_14683_f1.eps
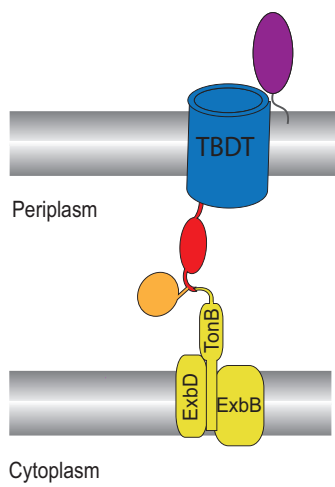
A.

PF07715 - Plug | PF00593 - Barrel — 12 sequences

PF13715 - NTE | PF07715 - Plug | PF00593 - Barrel — 90 sequences
2 unique structures*

PF07660 - STN | PF13715 - NTE | PF07715 - Plug | PF00593 - Barrel — 19 sequences
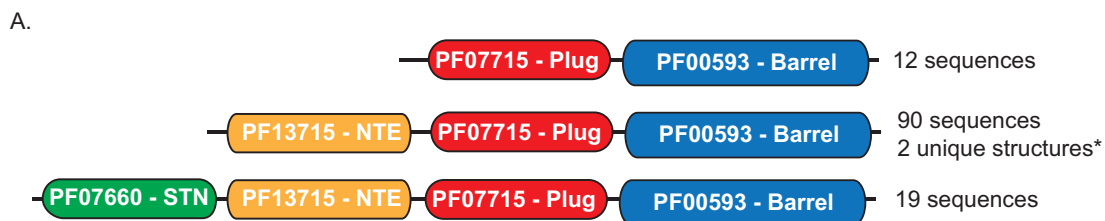
B.
Classical TBDT
5 with SusD protein

OM
TBDT
TonB
IM
ExbD ExbB

7 without SusD protein

OM
TBDT
TonB
IM
ExbD ExbB
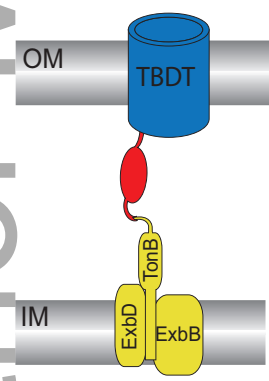
C.
N-terminal Extension TBDT
79 with SusD protein

Periplasm
OM
TBDT
TonB
IM
ExbD ExbB
Cytoplasm

11 without SusD protein

Periplasm
OM
TBDT
TonB
IM
ExbD ExbB
Cytoplasm
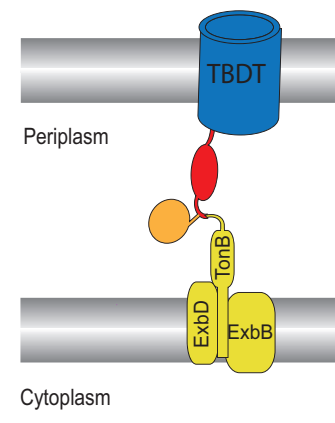
D.
Signal Transduction TBDT
18 with SusD protein

OM
TBDT
anti-σ
TonB
IM
ExbD ExbB
ECF σ

1 without SusD protein

OM
TBDT
anti-σ
TonB
IM
ExbD ExbB
ECF σ
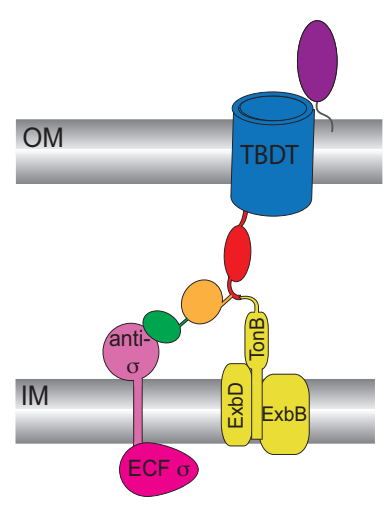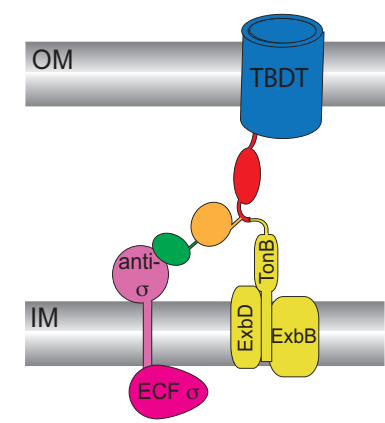
Figure 2 - Pollet, Martin & Koropatkin

mmi_14683_f2.eps

	Total from Pfam/PDB	**Total in B. theta**
PF07980 SusD_RagB	987 sequences 1 unique structure	2 sequences 1 unique structure
PF12741 SusD_like	546 sequences 10 unique structures	11 sequences 4 unique structures
PF12771 SusD_like_2	3,106 sequences, 9 unique structures	3 sequences 3 unique structures
PF14322 SusD_like_3 / PF07980 SusD_RagB	17,935 sequences, 22 unique structures	84 sequences 6 unique structures
PF12771 SusD_like_2 / PF12771 SusD_like_2	165 sequences 1 unique structure	0 sequences
PF14322 SusD_like_3	481 sequences	2 sequences
PF12771 SusD_like_2 / PF12741 SusD_like	117 sequences	0 sequences

B. SusD_RagB

α18
TPR4
α17
α8
TPR3   α7
α6
TPR2
α5   α4
α1
TPR1

C.

D.

E.

F.

SusD-like_3/

G.