Ascertaining the biochemical function of an essential pectin methylesterase in the gut microbe Bacteroides thetaiotaomicron

Duan, Cheng-Jie; Baslé, Arnaud; Liberato, Marcelo Visona; Gray, Joseph; Nepogodiev, Sergey A.; Field, Robert A.

Published in:
Journal of Biological Chemistry

DOI:
10.1074/jbc.RA120.014974

Publication date:
2020

Document Version
Publisher's PDF, also known as Version of record

Ascertaining the biochemical function of an essential pectin methylesterase in the gut microbe *Bacteroides thetaiotaomicron*

Received for publication, July 6, 2020, and in revised form, October 18, 2020 Published, Papers in Press, October 23, 2020, DOI 10.1074/jbc.RA120.014974

Cheng-Jie Duan[1], Arnaud Baslé[2], Marcelo Visona Liberato[3], Joseph Gray[2], Sergey A. Nepogodiev[4], Robert A. Field[5], Nathalie Juge[6], and Didier Ndeh[6,*]

From the [1]State Key Laboratory for Conservation and Utilization of Subtropical Agro-bioresources, College of Life Science and Technology, Guangxi University, Nanning, Guangxi, China, the [2]Institute for Cell and Molecular Biosciences, Newcastle University, Newcastle upon Tyne, United Kingdom, the [3]Laboratório Nacional de Ciência e Tecnologia do Bioetanol, Centro Nacional de Pesquisa em Energia e Materiais, Campinas, Brazil, the [4]Department of Biological Chemistry, John Innes Centre, Norwich, United Kingdom, the [5]Department of Chemistry and Manchester Institute of Biotechnology, University of Manchester, Manchester, United Kingdom, and the [6]Quadram Institute Bioscience, Norwich, United Kingdom

Edited by Chris Whitfield

Pectins are a major dietary nutrient source for the human gut microbiota. The prominent gut microbe *Bacteroides thetaiotaomicron* was recently shown to encode the founding member (BT1017) of a new family of pectin methylesterases essential for the metabolism of the complex pectin rhamnogalacturonan-II (RG-II). However, biochemical and structural knowledge of this family is lacking. Here, we showed that BT1017 is critical for the metabolism of an RG-II–derived oligosaccharide ΔBT1017oligoB generated by a BT1017 deletion mutant (ΔBT1017) during growth on carbohydrate extract from apple juice. Structural analyses of ΔBT1017oligoB using a combination of enzymatic, mass spectrometric, and NMR approaches revealed that it is a bimethylated nonaoligosaccharide (GlcA-β1,4-(2-*O*-Me-Xyl-α1,3)-Fuc-α1,4-(GalA-β1,3)-Rha-α1,3-Api-β1,2-(Araf-α1,3)-(GalA-α1,4)-GalA) containing components of the RG-II backbone and its side chains. We showed that the catalytic module of BT1017 adopts an α/β-hydrolase fold, consisting of a central twisted 10-stranded β-sheet sandwiched by several α-helices. This constitutes a new fold for pectin methylesterases, which are predominantly right-handed β-helical proteins. Bioinformatic analyses revealed that the family is dominated by sequences from prominent genera of the human gut microbiota, including *Bacteroides* and *Prevotella*. Our re-sults not only highlight the critical role played by this family of enzymes in pectin metabolism but also provide new insights into the molecular basis of the adaptation of *B. thetaiotaomicron* to the human gut.

The human large intestine is home to a large microbial community termed the human gut microbiota (HGM), which has substantial impact on the health and physiology of its host. Pectins, which are a major component of plant-based diets, have been shown to exert a significant selective pressure on HGM species (1–3) and hence have great potential as tools to manipulate the HGM. Pectins are defined as D-galacturonic acid–containing plant cell wall polysaccharides. The pectin macrostructure consists of three major polysaccharides: rhamnogalacturonan-I (RG-I), rhamnogalacturonan-II (RG-II), and homogalacturonan (4). Of these, RG-II is the most complex, consisting of several heterogenous side chains (A, B, C, D, E, and F), which are linked to a backbone of D-galacturonic acid (GalA) residues (Fig. 1*A*) (5). In total, RG-II contains at least 22 distinct glycosidic linkages and 13 different monosaccharides (Fig. 1*A*). The structure of RG-II is highly conserved; however, there is some variation in RG-II between plant species particularly at the termini of side chain B and in the methylation pattern of side chain A, as described previously (6, 7).

Bacteroides thetaiotaomicron is a prominent member of the HGM, equipped with a large repertoire of carbohydrate-active enzymes (CAZymes) and considered as a generalist being able to forage on a wide range of dietary or host glycans (for a review see Ref. 4). *B. thetaiotaomicron* has the ability to cleave 21 of the 22 glycosidic linkages in RG-II (except that in the disaccharide 2-*O*-Me-Xyl-α1,3-Fuc) (5). Among the *B. thetaiotaomicron* repertoire, several founding members of novel CAZyme families were characterized including a pectin methylesterase (PME) BT1017. BT1017 was shown to remove the 6-*O*-methyl decoration of GalA in the homogalacturonan backbone of RG-II, therefore playing a critical role in enabling access to the rest of the RG-II structure by other RG-II–degrading enzymes (5). Currently more than 18 carbohydrate esterase families have been assigned, according to the CAZyme database (8); this topic was recently reviewed by Nakamura *et al.* (9). Of the 18 families, CE8 is the only family that contains PMEs. BT1017, however, displays no sequence similarity to CE8 esterases, and hence the structural basis for its catalytic function is unknown. When cultured in media containing extensively purified apple RG-II as a sole carbon source, a *B. thetaiotaomicron* genetic mutant lacking the BT1017 enzyme (ΔBT1017) produces a pentasaccharide Rha-α1,3-Api-β1,2-(Araf-α1,3)-(6-*O*-Me-GalA-α1,4)-GalA here referred to as ΔBT1017oligoA (Fig. 1*A*) (5). The complete degradation of ΔBT1017oligoA requires five enzymes (BT1017, BT1018, BT1021, BT1012, and BT1001) collectively referred to here as A5 (Fig. 1*B*). BT1017 cleaves the 6-*O*-methylester linkage from the backbone GalA; BT1018 (α-galacturonidase) cleaves the glycosidic linkage between the two backbone GalA residues; BT1021 (α-arabinofuranosidase) cleaves the linkage between Araf (chain F) and the reducing end GalA; BT1012

This article contains supporting information.

* For correspondence: Didier Ndeh, dndeh001@dundee.ac.uk.
Present address for Didier Ndeh: Division of Molecular Microbiology, School of Life Sciences, University of Dundee, Dundee, United Kingdom.

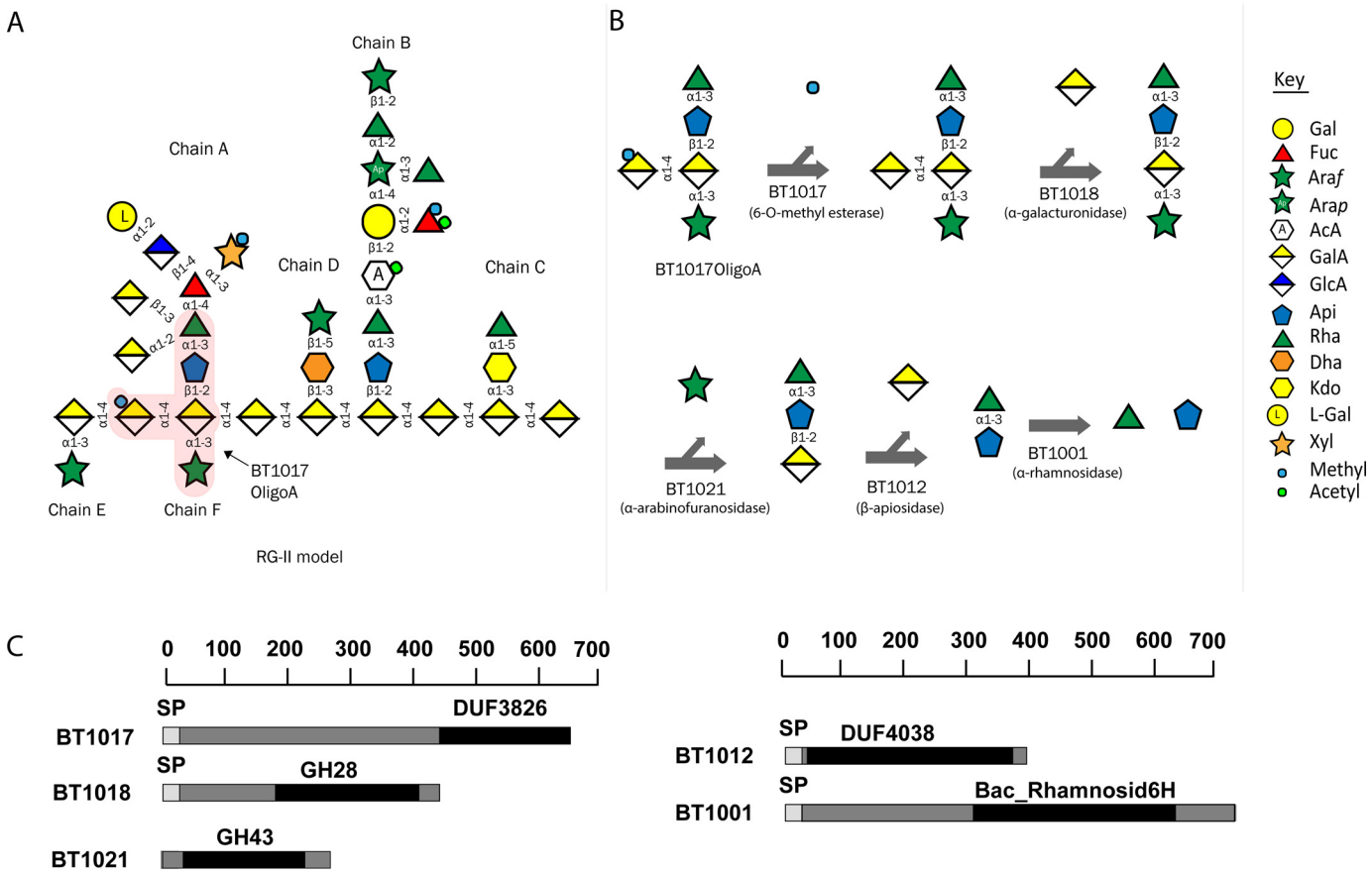


Figure 1. Structure and degradation of pectin RG-II. *A*, structural model of RG-II showing various side chains and ΔBT1017oligoA (*pink shading*). *B*, pathway showing the five enzymes (A5) required for complete ΔBT1017oligoA degradation in the human gut microbe *B. thetaiotaomicron*. *Gal*, galactose; *Fuc*, fucose; *Araf*, arabinofuranose; *Arap*, arabinopyranose; *AcA*, aceric acid; *GalA*, galacturonic acid; *GlcA*, glucuronic acid; *Api*, apiose; *Rha*, rhamnose; *Dha*, 3-deoxy-D-lyxo-2-heptulosaric acid; *Kdo*, 3-deoxy-D-manno-2-octulosonic acid; *L-Gal*, L-galactose; *Xyl*, xylose; *Methyl*, methyl group; *Acetyl*, acetyl group. *C*, modular architecture of BT1017 and various proteins involved in ΔBT1017oligoA degradation.

(β-apiosidase) cleaves the linkage between Api and the reducing end GalA; and finally BT1001 (α-rhamnosidase) cleaves the linkage between Rha and Api in the Rha–Api disaccharide (Fig. 1*B*).

In the present study, we showed that the *B. thetaiotaomicron* mutant ΔBT1017, when cultured in carbohydrate extract from apple juice (CEAJ) as a sole carbon source, generates a second oligosaccharide (hereafter referred to as ΔBT1017oligoB). Our structural analyses revealed that ΔBT1017oligoB is a dimethylated nonasaccharide containing components of the RG-II backbone and its side chains. We characterized the kinetic properties of BT1017, showing that the enzyme has a low turnover against apple RG-II, ΔBT1017oligoA, and ΔBT1017oligoB and hence may represent a rate-limiting step during RG-II metabolism. We revealed that BT1017 is a serine esterase with an α/β-hydrolase fold and hence has not evolved from the progenitor protein that gave rise to the CE8 family of PMEs, which are predominantly comprised of right-handed β-helices.

Results

Characterization of ΔBT1017oligoB

B. thetaiotaomicron ΔBT1017 deletion mutant was cultured on CEAJ for 48 h to stationary phase ($A_{600\ nm}$ ~1.0), and TLC was first used to analyze the culture supernatants. The data showed that ΔBT1017 generates two oligosaccharides, defined as ΔBT1017oligoA and ΔBT1017oligoB (Fig. 2*A*). Both sugars were purified by size-exclusion chromatography and treated independently with a mixture of recombinant A5 enzymes including BT1017, BT1018 (α-galacturonidase), BT1021 (α-arabinofuranosidase), BT1012 (β-apiosidase), and BT1001 (α-rhamnosidase), which target specific glycosidic linkages in RG-II (5). Unless otherwise stated, all the recombinant RG-II–degrading enzymes mentioned in this text were the same constructs used by Ndeh *et al.* (5) and lack the N-terminal signal peptide (SP) regions (Fig. 1*C*). The products of the enzymatic treatment were then analyzed by TLC and HPLC (Fig. 2, *B* and *C*). Digestion of ΔBT1017oligoA yielded GalA, Araf, Rha and Api indicating that the molecule is the methylated pentasaccharide (Rha-α1,3-Api-β1,2-(Ara-α1,3)(6-*O*-Me-GalA-α1,4)-GalA, which was used to demonstrate the site of action of the PME in Ndeh *et al.* (5). The digestion of ΔBT1017oligoB, on the other hand, was incomplete, yielding only two monosaccharides (Araf and GalA) and a third product of unknown identity (Fig. 2, *B* and *C*). The release of GalA and Araf from ΔBT1017oligoB by ΔBT1017oligoA-specific enzymes BT1018 (α-galacturonidase) and BT1021 (α-arabinofuranosidase) suggests that ΔBT1017oligoB contains the backbone GalA and the side-chain F Araf sugars characteristic of ΔBT1017oligoA (Fig. 1, *A* and *B*).

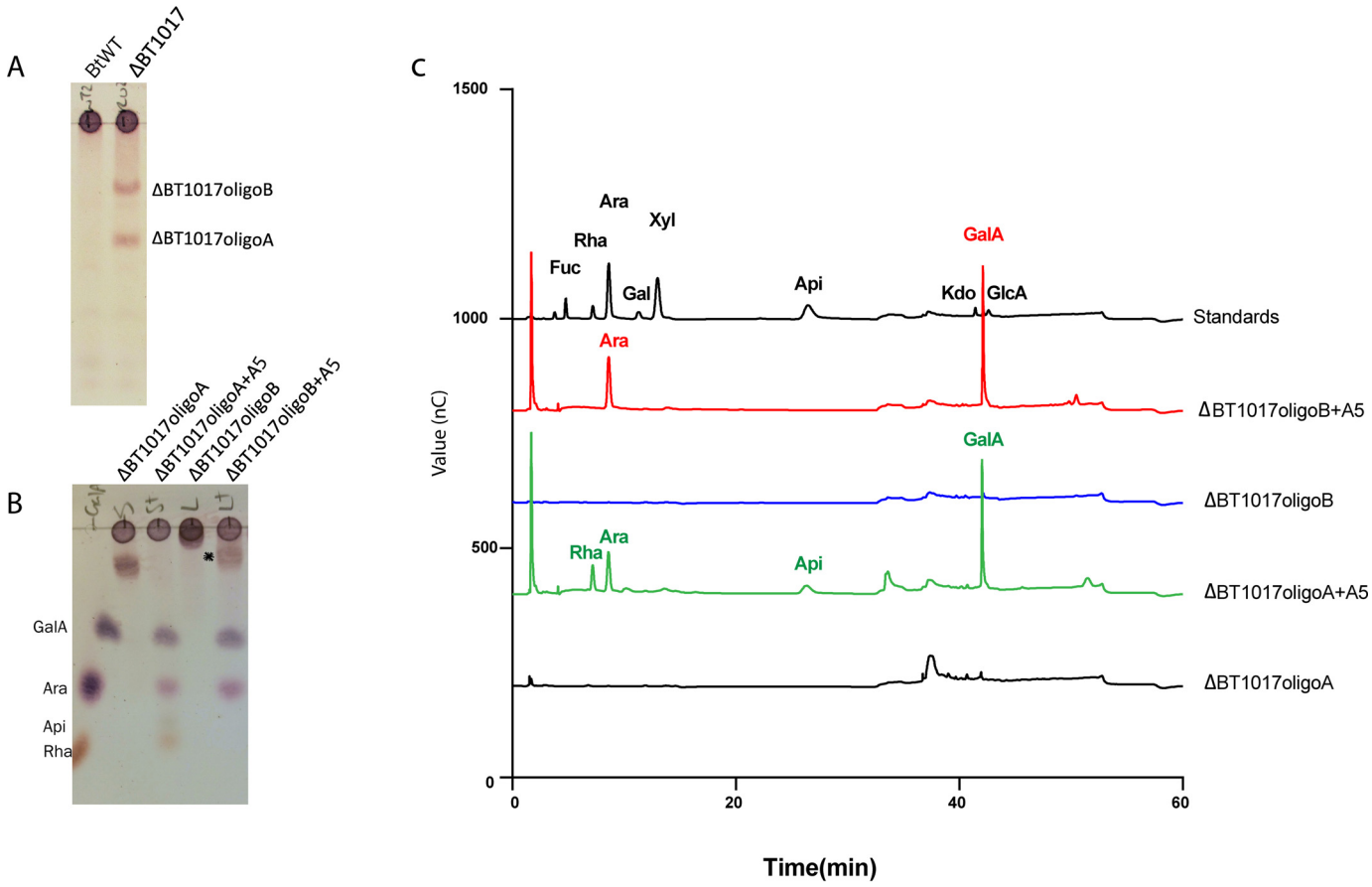


Figure 2. *B. thetaiotaomicron* deletion mutant ΔBT1017 generates two oligosaccharides, ΔBT1017oligoA and ΔBT1017oligoB, during growth on CEAJ. *A*, TLC analyses of culture supernatants from *B. thetaiotaomicron* WT (*BtWT*) and ΔBT1017 cells post-growth on CEAJ. *B*, digestion of ΔBT1017oligoA and B with a mixture of ΔBT1017oligoA-degrading enzymes (A5) containing BT1017, BT1018, BT1021, BT1012, and BT1001. Each substrate was treated with A5, and the reaction was stopped and analyzed by TLC. The product marked with an *asterisk* is the leftover of ΔBT1017oligoB after digestion with the mixture of A5 enzymes. *C*, HPLC analyses of samples in ΔBT1017oligoB.

To determine the full structure of ΔBT1017oligoB, a combination of MS, enzymatic and NMR analyses were performed. First, MS data revealed that ΔBT1017oligoB has a protonated molecular mass ($[M + H]^+$) of 1453.44 Da (Fig. 3*A*). When treated with BT1017, the mass of ΔBT1017oligoB decreased by 28.03 Da (Fig. 3*B*), corresponding to the loss of two methyl groups. This suggests that ΔBT1017oligoB contains two ester-linked methyl groups that were hydrolyzed by the BT1017 PME. Second, when WT *B. thetaiotaomicron* was cultured on ΔBT1017oligoB, the bacterium accumulated the disaccharide 2-*O*-Me-Xyl-α1,3-Fuc, which is unique to side chain A of RG-II but not present in ΔBT1017oligoA (Fig. 3*C*). The sugar Api was also detected. These results demonstrate that ΔBT1017oligoB contains components of ΔBT1017oligoA and additional sugars from RG-II side chain A. Last, ΔBT1017oligoB was shown to be susceptible to attack by the β-D-glucuronidase enzyme BT0996, which released GlcA (Fig. 3*D*). Because this required pretreatment with BT1017, this result suggests that at least one of the methyl decorations in ΔBT1017oligoB sterically hinders the activity of BT0996. Release of free GlcA is also an indication that ΔBT1017oligoB lacks the terminal L-Gal residue, which is α1,2-linked to GlcA at the nonreducing end of chain A (Fig. 1*A*).

Based on the above features of ΔBT1017oligoB (protonated mass of 1452.44 Da, presence of 2-*O*-Me-Xyl-α1,3-Fuc, GalA, Ara, Api, and GlcA (highlighted in Fig. S1*A*) and the absence of terminal L-Gal), two possible structures of de–methyl-esterified BT1017oligoB (ΔBT1017oligoB-2Me) were deduced from the known structure of RG-II (Fig. S1, *B* and *C*). These include ΔBT1017oligoB-2Me-α (GlcA-β1,4-(2-*O*-Me-Xyl-α1,3)-Fuc-α1,4-(GalA-α1,2)-Rha-α1,3-Api-β1,2-(Ara*f*-α1,3)-(GalA-α1,4)-GalA) and ΔBT1017oligoB-2Me-β (GlcA-β1,4-(2-*O*-Me-Xyl-α1,3)-Fuc-α1,4-(GalA-β1,3)-Rha-α1,3-Api-β1,2-(Ara*f*-α1,3)-(GalA-α1,4)-GalA).

Both sugars differ by the presence of either α1,2- or β1,3-linked GalA (underlined). To determine which of them corresponded to ΔBT1017oligoB-2Me, enzymes targeting all linkages in the predicted sugars (ΔBT1017oligoB-2Me-α and ΔBT1017oligoB-2Me-β) were used to sequentially digest ΔBT1017oligoB. The first set of recombinant enzymes collectively referred to here as B5 enzymes include BT1017, BT1018, BT1021, BT0996, and BT1012. These together should cleave the two ester groups, the backbone α1,3-linked GalA, the side chain F α1,3-linked Ara*f*, the side chain A β1,4-linked GlcA, and the reducing end/backbone GalA residue, respectively (Fig. S1, *D* and *E*), to generate two possible pentasaccharide structures: MXFGRA-α (2-*O*-Me-Xyl-α1,3-Fuc-α1,4-(GalA-α1,2)-Rha-α1,3-Api) and MXFGRA-β (2-*O*-Me-Xyl-α1,3-Fuc-α1,4-(GalA-β1,3)-Rha-α1,3-Api), differing by the presence or

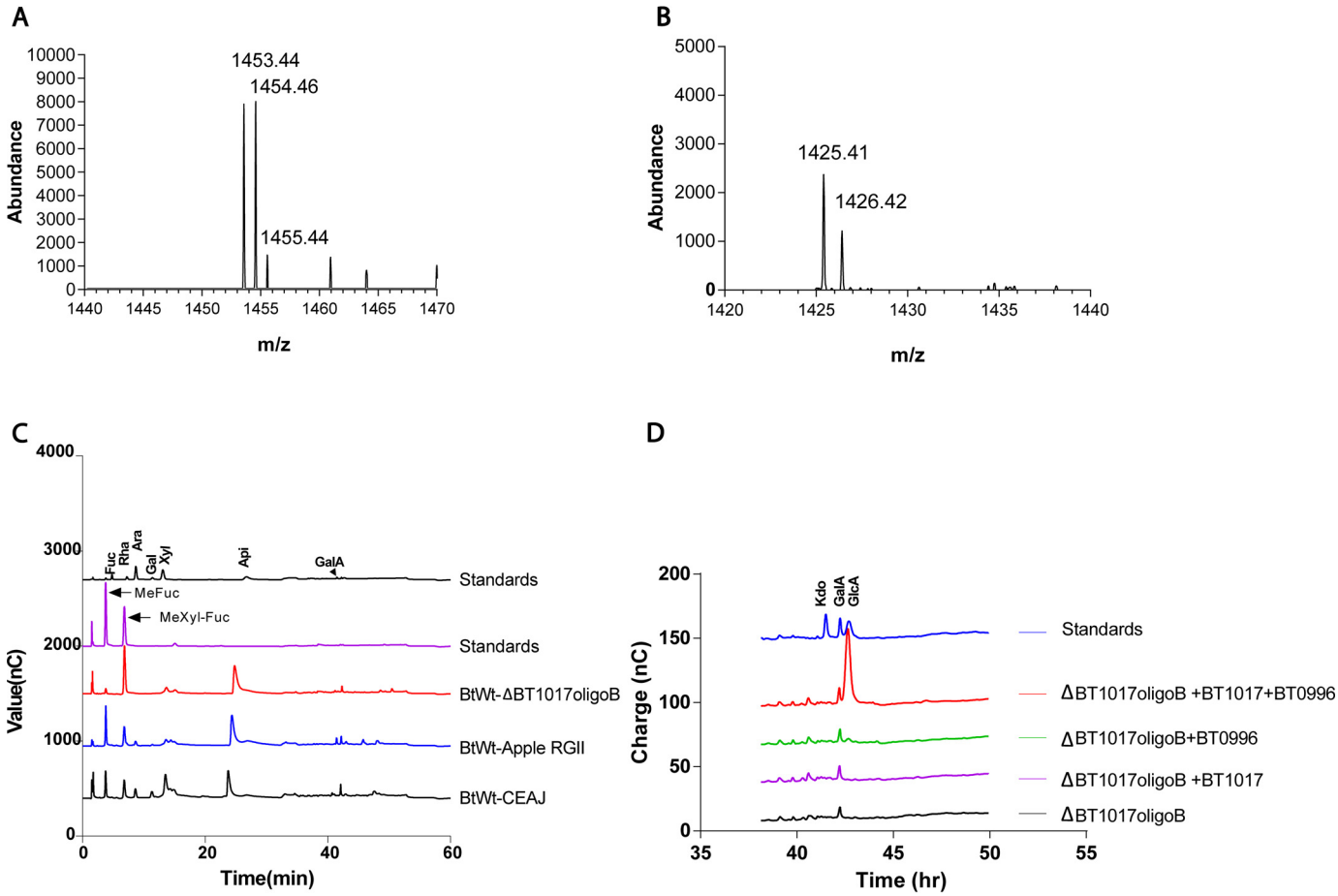
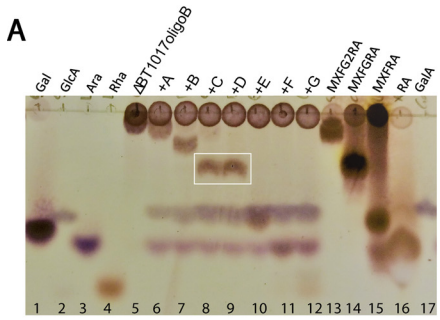


Figure 3. Analyses of ΔBT1017oligoB by MS and HPLC. *A*, MS of purified ΔBT1017oligoB. Masses presented are for ΔBT1017oligoB and fragments plus hydrogen ions $[H]^+$. *B*, MS of purified ΔBT1017oligoB after digestion with BT1017. *C*, HPLC analyses of spent media following growth of *B. thetaiotaomicron* ΔBT1017 on ΔBT1017oligoB, apple RG-II, and CEAJ MeFuc: 2-*O*-Me-Fuc, Mexyl-Fuc: 2-*O*-Me-Xyl-α1,3-Fuc. *D*, HPLC data showing activity of BT0996 and BT1017 on ΔBT1017oligoB.

absence of either α1,2- or β1,3-linked GalA (underlined). Digestion of ΔBT1017oligoB-2Me with a mixture of recombinant B5 enzymes (BT1017, BT1018, BT1021, BT0996, and BT1012) and subsequent analyses by TLC revealed the generation of a product that migrates to a similar extent as the sugar standard MXFGRA-β (Fig. 4*A*, *lane 8* in *white rectangle*). However, it was also possible that the product corresponded to MXFGRA-α because of the significant structural similarity to MXFGRA-β. As a result, it was referred to as MXFGRA-x. Both α1,2- and β1,3-GalA linkages in MXFGRA-α and MXFGRA-β have been shown to be specifically targeted by the enzymes BT0997 (α-galacturonidase) and BT0992 (β-galacturonidase), respectively (5) (Fig. S1, *D* and *E*); hence to determine whether the product contained α- or β-linked GalA residues, each of these enzymes (BT0992 and BT0997) was used to further digest MXFGRA-x. TLC analyses of the reaction showed that MXFGRA-x was digested by BT0992 but not by BT0997 (Fig. 4*A*, *lanes 9* and *10*, respectively), indicating that the exposed GalA residue in the product was β1,3-linked to Rha and that the pentasaccharide was MXFGRA-β. This was also confirmed by 2D HSQC NMR analyses of ΔBT1017oligoB, which detected ^{1}H and ^{13}C HSQC anomeric signals (δ_H 4.67 and δ_C 104.1) of β-GalA. The NMR analyses also revealed H1/C1 signals of all other carbohydrate residues in the anomeric region of the spectrum. These include signals for GalA-α1-4, Ara*f*-α1-3, Api-β1-2, Rha-α1-3, Fuc-α1-4, 2-*O*-Me-Xyl-α1-3, and GlcA-β1,4 (Fig. 4*B*), which were assigned by comparison with published data (10). Two weaker cross-peaks could be assigned to the anomeric center of the reducing-end GalA residue in the backbone of ΔBT1017oligoB. The full monosaccharide composition of ΔBT1017oligoB-2Me was confirmed by treatment of the sugar with a combination of A5 enzymes together with BT0996, BT0992, and other RG-II–degrading enzymes BT1002 (α-L-fucosidase) and BT1001 (α-L-rhamnosidase) and analyses of the digested sample by HPLC. The results showed that the enzymes degraded the sugar into all its constituent monosaccharides GlcA, GalA Rha Api Ara*f*, and the disaccharide 2-*O*-Me-Xyl-α1,3-Fuc (Fig. 4*C*). A model showing the cleavage sites of various enzymes on ΔBT1017oligoB is shown in Fig. 4*D*.

Activity of BT1017 and kinetic parameters

Full-length BT1017 (BT1017-FL) is a 73.7-kDa protein consisting of a SP (positions 1–20), a sequence of unknown function (positions 20–400), or central module (CM) and a domain of unknown function DUF3826 (positions 400–600) (Figs. 1 and 5). To determine the site of the esterase activity in BT1017,



ID	Composition of enzymatic mixture								
A	BT1017	BT1018	BT1021						
B	BT1017	BT1018	BT1021	BT0996					
C	BT1017	BT1018	BT1021	BT0996	BT1012				
D	BT1017	BT1018	BT1021	BT0996	BT1012	BT0997			
E	BT1017	BT1018	BT1021	BT0996	BT1012	BT0997	BT0992		
F	BT1017	BT1018	BT1021	BT0996	BT1012	BT0997	BT0992	BT1002	
G	BT1017	BT1018	BT1021	BT0996	BT1012	BT0997	BT0992	BT1002	BT1001

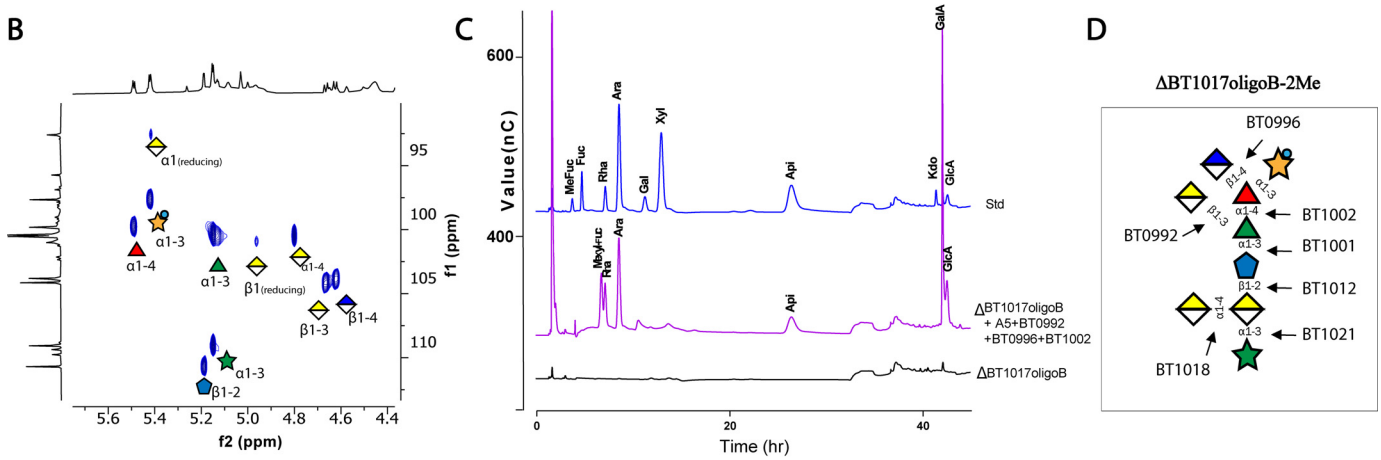


Figure 4. Structural characterization of ΔBT1017oligoB. *A*, TLC analysis ΔBT1017oligoB after digestion with diverse combinations of RG-II–degrading enzymes. The *bands* in the *white rectangle* correspond to MXFGRA-x later confirmed to be MXFGRA-β (2-*O*-Me-Xyl-α1,3-Fuc-α1,4-(GalA-β1,3)-Rha-α1,3-Api). Other complex sugar standards include MXFG2RA (2-*O*-Me-Xyl-α1,3-Fuc-α1,4-(GalA-α1,2)(GalA-β1,3)-Rha-α1,3-Api), MXFRA (2-*O*-Me-Xyl-α1,3-Fuc-α1,4-Rha-α1,3-Api), and RA (Rha-α1,3-Api). *B*, HSQC NMR of anomeric region of ΔBT1017oligoB. *C*, HPLC analyses of ΔBT1017oligoB after complete hydrolysis with a mixture of ΔBT1017oligoB-degrading enzymes. *D*, model of ΔBT1017oligoB-2Me showing the cleavage sites of various ΔBT1017oligoB-degrading enzymes.

various recombinant fragments of the protein namely BT1017-SP (71.7 kDa; lacking the signal peptide), BT1017-CM (49.4 kDa; corresponding to CM), and BT1017-DUF3826 (23.5 kDa; corresponding to the DUF3826 domain) (Fig. 5) were expressed and tested against apple RG-II using a coupled spectrophotometric enzyme assay that measures released methanol (as described earlier (11)). Only BT1017-SP and BT1017-CM showed activity against the substrate (Table 1), indicating that BT1017-CM comprises the catalytic site. The optimal temperature and pH for BT1017 activity were determined to be 37 °C and 8.5, respectively (Fig. S2). To further explore the specificity of the enzyme, BT1017-SP was tested against various methyl- and acetyl-esterified substrates. The enzyme was active against ΔBT1017oligoA, ΔBT1017oligoB, 6-*O*-methyl galacturonic (Me-GalA), and 6-*O*-methyl glucuronic acid (Me-GlcA) but not methylpropionate, methylbutyrate, ethylpropionate ethylbutyrate, acetylated potato RG-I, and 4-nitrophenyl-acetate (Table 1). BT1017 thus appears to have a preference for methylated hexose sugars. The K_m and k_{cat} of BT1017-SP and BT1017-CM against various substrates are reported in Table 1, showing that BT1017-SP has an ~2-fold higher catalytic efficiency compared with BT1017-CM against more complex substrates such as apple pectin. BT1017-SP also showed a higher k_{cat}/K_m toward Me-GalA and ΔBT1017oligoB compared with Me-GlcA and ΔBT1017oligoB; however, the difference was less than 10-fold.

3D structural features of BT1017

BT1017-SP could not be crystallized. BT1017-CM, however, generated crystals in the space group C2221 with one molecule in the asymmetric unit. The structure of the enzyme was determined by single-wavelength anomalous dispersion phasing and refined to 1.9 Å with an R_{factor} of 20.04% and an R_{free} of 24.09%. The final model (PDB entry 6GOC) of BT1017-CM consists of residues 20–462 of BT1017. The stereochemical quality of the model was assessed by validation tools in Coot and MolProbity (12, 13). The ratios of preferred and allowed regions from the Ramachandran plot are 96.61 and 3.17%, respectively. The statistics for data collection and refinement are summarized in