Structure of an ABC Transport Solute-Binding Protein that Facilitates Growth on the Cereal Fiber arabinoxylan in the Human Gut Symbiont *Eubacterium rectale*

Rebecca M. Duvall, BS, and Nicole M. Koropatkin, PhD[1]
Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, MI 48109
[1] To whom correspondence should be addressed: nkoropat@umich.edu

Eubacterium rectale, a member of the Firmicutes phylum species, is a human gut symbiont that produces butyrate, a short chain fatty acid with anti-inflammatory and anti-tumorigenic properties in the intestinal tract. Butyrate is a by-product of carbohydrate fermentation by *E. rectale*, yet little is known about carbohydrate uptake in this bacterium. Here we propose that arabinoxylan, a major plant cell wall hemicellulose, is scavenged by *E. rectale* via an ATP-binding cassette (ABC) transporter system. We describe a genetic locus that has a xylan-binding protein (XBP) that presumably works as part of an ABC transporter to transport arabinoxylo-oligosaccharides into the cell. We have structurally characterized the XBP revealing that it likely targets longer arabinoxylo-oligosaccharides, which may be advantageous in the gut environment.

Introduction

Tens of trillions of microorganisms live in and on humans. These bacteria coexist with the host to form a vast and complex microbial ecosystem in the digestive system known as the gut microbiota[1]. The gut microbiota impacts host health in a variety of ways, including but not limited to vitamin production, immune development, and pathogen resistance[2,3]. The gut microbiota additionally carries out biochemical functions not encoded within the human genome, such as the digestion of complex dietary carbohydrates[4,5]. Most dietary glycans, with the exception of lactose, sucrose and starch, cannot be degraded by human enzymes and thus become food for gut bacteria that ferment these to host-absorbable short chain fatty acids. Thus the microbiota is critically important for the liberation of carbohydrate nutrition for the host.

Many of the metabolic pathways of the gut microbiota have not yet been structurally or biochemically categorized because of the complexity, diversity, and variation of the microbiota between and within individuals[6]. However, 16S rRNA sequencing and metagenomic analyses of the human gut microbiome have revealed the dominant bacterial Phyla as Bacteroidetes and Firmicutes, which have differing capacities for carbohydrate metabolism[7]. The symbiont *Eubacterium rectale* belongs to the Firmicutes phylum and is a Gram-positive, rod-shaped bacterium that is significantly more abundant in otherwise healthy individuals compared to patients with inflammatory bowel disease or colorectal cancer[8]. As a by-product of carbohydrate fermentation *E. rectale* produces butyrate, a short chain fatty acid (SCFA) that has anti-tumorigenic and anti-inflammatory properties because it can regulate gene expression in different host cells[9,10]. A mechanistic understanding of glycan degradation by *E.*

rectale would allow us to manipulate community composition towards improved human health. For example, understanding the specific carbohydrates targeted by this microbe would allow us to identify foods that promote beneficial microbial growth in the gut (called prebiotics), to increase the population of *E. rectale* and perhaps butyrate levels.

Arabinoxylan, a soluble component of the plant cell wall, is a polysaccharide that is utilized by *E. rectale*. Arabinoxylan is a major element of dietary fiber in cereal grains and is thus a common polysaccharide in the average diet. Its structure is made up of xylose and arabinose monosaccharides[11]. Due to its complex nature, breaking down arabinoxylan requires a concerted enzymatic effort that has not yet been structurally or biochemically well-characterized for *E. rectale.* However, the degradation of arabinoxylan has been characterized in other organisms, for example *Bacteroides ovatus*. This organism contains 21 enzymes that partake in the breakdown of arabinoxylan, a process that begins at the cell surface, and fragments of the fiber (oligosaccharides) are completely depolymerized once imported into the cell[12].

We and others have identified a genetic locus in the *E. rectale* DSM 17629 strain that encodes for a glycoside hydrolase family 10 (GH10) enzyme as well as an ATP-binding cassette (ABC) transporter[13]. GH10 enzymes typically hydrolyze (arabino)xylan polysaccharides, breaking the fiber into fragments that can be imported into the bacterial cell. ABC transporters are found across almost every life form. They typically consist of a solute-binding protein that binds ligands with high specificity and affinity, a membrane transporter through which the ligand passes, and an ATPase that supplies energy to the system and facilitates the conformational change in the transporter that allows the ligand to enter the cell[14]. In the case of carbohydrate degradation, these systems typically recognize mono-, di- or oligosaccharides.

Two organisms that use ABC transporters to harvest arabinoxylo-oligosaccharides are *Bifidobacterium animalis* and *Caldanaerobius polysaccharolyticus*, belonging to the Actinobacteria and Firmicutes Phyla, respectively[12,15]. Both organisms have xylan-binding proteins (XBPs) that have been structurally and biochemically characterized. We expect the ABC transporter encoded proximal to the GH10 gene in *E. rectale* DSM 17629 will have similar characteristics to these two species regarding the uptake of arabinoxylan via an ABC transporter system. To test this hypothesis, we are structurally charactering the XBP and seek to determine what aspect of arabinoxylan structure is recognized. Our data demonstrate that the *E. rectale* DSM 17629 XBP has a canonical solute-binding protein structure that likely captures longer arabinoxylo-oligosaccharides for import into the cell.

Materials & Methods

Protein production and purification

The gene fragment corresponding to Eur_20790 (encoding residues 29 – 556) was amplified via PCR and ligated into the pETite Nhis plasmid (Lucigen, Madison, WI) according to the manufacturer's instructions. This construct included an N-terminal 6-His tag for Ni affinity purification, as well as a tobacco etch virus (TEV) protease cleavage site to remove the tag after purification.

The plasmid containing the Eur_20790 gene was transformed into Rosetta (DE3) pLysS *E. coli* cells for protein expression. For structure determination, selenomethionine-substituted Eur_20790 protein was obtained by culturing the Rosetta (DE3) pLysS cells in M9 minimal media supplemented with kanamycin (30 µg/mL) and chloramphenicol (20 µg/mL) and grown at 37°C for 16 hours. This overnight culture was used to inoculate a 2 L baffled flask containing 1 L of Molecular Dimensions SelenoMet Premade Medium supplemented with 50 mls of the recommended sterile nutrient mix, chloramphenicol, and kanamycin. Cultures were grown at 37°C to an A_{600nm}≈0.45 before adjusting the temperature to 20°C, and supplementing each flask with 100 mgs each of L-lysine, L-threonine, L-phenylalanine, and 50 mgs each of L-leucine, L-isoleucine, L-valine, and L-selenomethionine. After 20 additional minutes of growth, the cells were induced with 0.5 mM IPTG, and cultures were grown for an additional 48 hours.

To purify the selenomethionine-substituted protein, cells were thawed and lysed via sonication in His-Buffer (25 mM NaH_2PO_4, 500 mM NaCl, 20 mM imidazole pH 7.5) and purified via immobilized Nickel affinity chromatography (His-Trap, GE Healthcare) using a gradient of 20-300 mM imidazole, according to the manufacturer's instructions. The His-tag was removed by incubation with TEV protease (1:100 molar ratio relative to protein) at room temperature for 2 h, then overnight at 4°C while dialyzing against His Buffer. The cleaved protein was then re-purified via Nickel affinity chromatography to remove undigested target protein, the cleaved His-tag, and His-tagged TEV protease. Purified Eur_20790 protein was dialyzed against 20 mM HEPES, 100 mM NaCl (pH 7.0) prior to crystallization, and concentrated using Vivaspin 15 (10 000 MWCO) centrifugal concentrators (Vivaproducts, Inc.).

Crystallization and data collection

All X-ray diffraction data for both native and selenomethionine-substituted protein crystals were collected at the Life Science Consortium (LS-CAT) at the Advance Photon Source at Argonne National Laboratory in Argonne, IL. Selenomethionine-substituted crystals of Eur_20790 were grown via hanging drop vapor diffusion against 500 µL of 0.1 M MES/Imidazole buffer (pH 6.5), 0.15 M $MgCl_2$, 0.15 M $CaCl_2$, 6.75% v/v 2-methyl-2,4-pentanediol, 6.75% w/v polyethylene glycol (PEG) 1000, and 6.75% w/v PEG 3350. All crystals were flash-frozen prior to data collection by briefly soaking in a solution of 80% mother liquor / 20% ethylene glycol. Data were processed and scaled in HKL2000 and Scalepack[16]. Within the Phenix suite of programs, AutoSol was used to locate the positions of the selenium atoms in the selenomethionine data set, and the structure was automatically built in Autobuild[17,18]. Successive rounds of manual model building and refinement in Coot and Phenix respectively were utilized to build a 1.9 Å model of the selenomethionine-substituted protein[17,19]. We are working to complete the structure of the protein.

Results and Discussion

XBP structural characterization

XBP adopts a typical solute-binding protein fold (Figure 1A). It consists of a bi-lobed structure made of two α-helical domains connected by a flexible hinge region. Notably, the structure shows a platform of aromatic residues in the cleft between the two domains. This platform is a signature of protein-carbohydrate interactions and we presume that like other solute-binding proteins, the ligand binds in between the two domains. This would cause a conformational change mediated by the flexible hinge region causing the two domains to clamp down as a result. This mechanism of ligand binding is expected of a solute-binding protein.

XBP and homologues

Conclusions

The gut microbiota comprises a vast and diverse collection of microorganisms. The microbiota plays many roles in human health, including the digestion and utilization of carbohydrates. Of the microorganisms that participate in carbohydrate degradation, our research focused on *Eubacterium rectale* due to its health-promoting production of butyrate and its abundance in the gut. We are specifically interested in how this microbe recognizes common dietary fibers such as arabinoxylan. The xylan-binding protein (XBP) of *E. rectale* DSM 17629 was predicted to be a solute-binding protein of an ABC transporter in a putative arabinoxylan utilization operon. Our research found that the structure of the protein XBP adopts a typical solute-binding protein and shows evidence for carbohydrate-protein interaction in what we expect is the ligand-binding site. We also found that the closest homologue to XBP is a solute-binding protein. Based upon its extended carbohydrate-binding platform, we expect that XBP will bind to oligosaccharides longer than xylotriose, perhaps as they are liberated from arabinoxylans by the GH10 encoded close to this transporter in the genome (Figure 2).

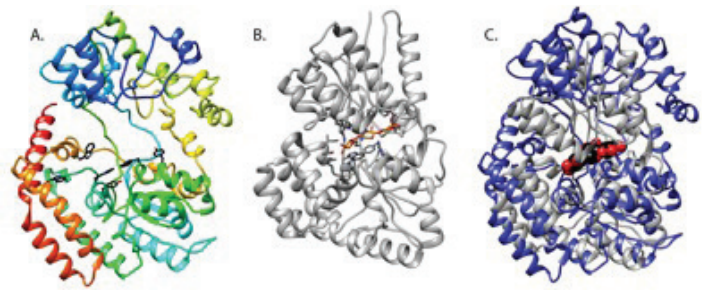

Figure 1: Comparison of the E. rectale XBP and the C. polysaccharolyticus xylotriose-binding protein. A. Ribbon diagram of the E. rectale XBP, colored blue to red from N- to C-terminus. Hydrophobic residues within the cleft between the two domains are displayed as black sticks. B. Ribbon diagram of the C. polysaccharolyticus xylotriose-binding protein (PDB id 4G68), with xylotriose in orange and yellow sticks and residues within 4Å of the sugar as grey, blue and red sticks. C. Overlay of the structures in panels A and B, with XBP in blue and the xylotriose-binding protein in grey. The xylotriose is displayed as black and red spheres.

The next step in our investigation is to refine our structural characterization of XBP and determine its precise glycan-binding profile using isothermal titration calorimetry. This technique involves the titration of a glycan into a concentrated solution of protein. The heat released from this interaction is measured and can be plotted to determine the dissociation constant (K_d) for the pairing. In addition we plan to discern activity of the GH10 enzyme encoded within the same operon as XBP in order to determine the oligosaccharide size released by this enzyme to see if it correlates

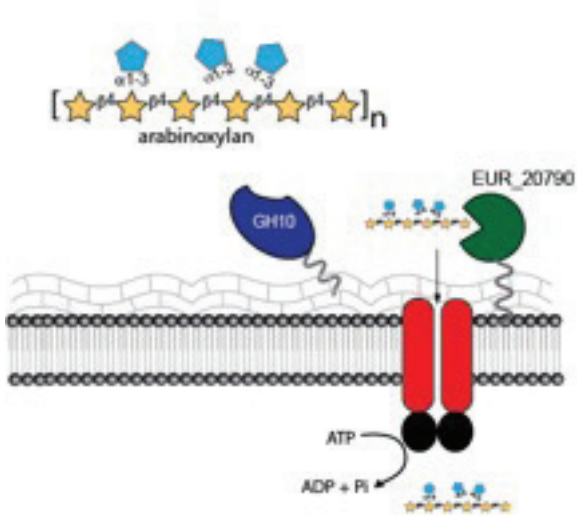

Figure 2: Model of arabinoxylan utilization in E. rectale DSM 17629. Arabinoxylan is a cereal fiber comprised of a xylose monosaccharide backbone (yellow stars) with arabinose monosaccharides (blue pentagon) decorations. The GH10 enzyme presumably hydrolyzes the longer polymer into shorted fragments that are bound by the XBP (Eur_20790) and imported into the cell.

with the size product recognized by the XBP. This work will provide a better understanding of the cereal fiber arabinoxylan degradation pathway in *E. rectale* and puts us closer to developing prebiotics to manipulate the gut microbiota for improved colonic health.

References

1. El Kaoutari, A., Armougom, F., Gordon, J. I., Raoult, D., & Henrissat, B. (2013). The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nature reviews Microbiology, 11(7),* 497-504.
2. Kinross, J. M., Darzi, A. W., & Nicholson, J. K. (2011). *Gut microbiome-host interactions in health and disease. Genome medicine, 3(3),* 1.
3. Kau, A. L., Ahern, P. P., Griffin, N. W., Goodman, A. L., & Gordon, J. I. (2011). *Human nutrition, the gut microbiome and the immune system. Nature, 474(7351*), 327-336.
4. Egert, M., De Graaf, A. A., Smidt, H., De Vos, W. M., & Venema, K. (2006). *Beyond diversity: functional microbiomics of the human colon. Trends in microbiology, 14(2),* 86-91.
5. Ley, R. E., Turnbaugh, P. J., Klein, S., & Gordon, J. I. (2006). *Microbial ecology: human gut microbes associated with obesity. Nature, 444(7122),* 1022-1023.
6. Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K., & Knight, R. (2012). *Diversity, stability and resilience of the human gut microbiota. Nature, 489(7415),* 220-230.
7. Gill, S. R., Pop, M., DeBoy, R. T., Eckburg, P. B., Turnbaugh, P. J., Samuel, B. S., Fraser-Ligett, C. M., & Nelson, K. E. (2006). *Metagenomic analysis of the human distal gut microbiome. science, 312(5778*), 1355-1359.
8. Wang, T., Cai, G., Qiu, Y., Fei, N., Zhang, M., Pang, X., Sanjun, C., & Zhao, L. (2012). Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. *The ISME journal, 6(2),* 320-329.
9. Barcenilla, A., Pryde, S. E., Martin, J. C., Duncan, S. H., Stewart, C. S., Henderson, C., & Flint, H. J. (2000). Phylogenetic relationships of butyrate-producing bacteria from the human gut. *Applied and environmental microbiology, 66(4),* 1654-1661.
10. Louis, P., & Flint, H. J. (2009). Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. *FEMS microbiology letters, 294(1),* 1-8.
11. McKee, L. S., Peña, M. J., Rogowski, A., Jackson, A., Lewis, R. J., York, W. S., Gilbert, H. J., & Marles-Wright, J. (2012). Introducing endo-xylanase activity into an exo-acting arabinofuranosidase that targets side chains. *Proceedings of the National Academy of Sciences, 109(17),* 6537-6542.
12. Han, Y., Agarwal, V., Dodd, D., Kim, J., Bae, B., Mackie, R. I., Nair, S. K., & Cann, I. K. (2012). Biochemical and structural insights into xylan utilization by the thermophilic bacterium Caldanaerobius polysaccharolyticus. *Journal of Biological Chemistry*, *287*(42), 34946-34960.
13. Sheridan, P. O., Martin, J. C., Lawley, T. D., Browne, H. P., Harris, H. M., Bernalier-Donadille, A., Scott, K. P., & Flint, H. J. (2016). Polysaccharide utilization loci and nutritional specialization in a dominant group of butyrate-producing human colonic Firmicutes. *Microbial Genomics*, *2*(2).
14. Cockburn, D. W., Orlovsky, N. I., Foley, M. H., Kwiatkowski, K. J., Bahr, C. M., Maynard, M., Demeler, B., & Koropatkin, N. M. (2015). Molecular details of a starch utilization pathway in the human gut symbiont Eubacterium rectale. *Molecular microbiology*, *95*(2), 209-230.

15. Ejby, M., Fredslund, F., Vujicic□Zagar, A., Svensson, B., Slotboom, D. J., & Abou Hachem, M. (2013). Structural basis for arabinoxylo□oligosaccharide capture by the probiotic Bifidobacterium animalis subsp. lactis Bl□04.*Molecular microbiology*, *90*(5), 1100-1112.
16. Otwinowski, Z. and W. Minor, *Processing of X-ray Diffraction Data Collected in Oscillation Mode.*, in *Methods in Enzymology*, C.W.J. Carter and R.M. R.M. Sweet, Editors. 1997, Academic Press. p. 307-326.
17. Adams, P.D., et al., *PHENIX: building new software for automated crystallographic structure determination.* Acta crystallographica. Section D, Biological crystallography, 2002. **58**(Pt 11): p. 1948-54.
18. Terwilliger, T.C., et al., *Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard.* Acta Crystallogr D Biol Crystallogr, 2009. **65**(Pt 6): p. 582-601.
19. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallogr D Biol Crystallogr, 2004. **60**(Pt 12 Pt 1): p. 2126-32.
20. Koropatkin, N. M., Cameron, E. A., & Martens, E. C. (2012). How glycan metabolism shapes the human gut microbiota. *Nature reviews Microbiology*,*10*(5), 323-335.