

## Selecting a Single Stereocenter: The Molecular Nuances that Differentiate #-Hexuronidases in the Human Gut Microbiome

Samuel J Pellock, William G. Walton, and Matthew R. Redinbo

*Biochemistry*, **Just Accepted Manuscript** • DOI: 10.1021/acs.biochem.8b01285 • Publication Date (Web): 07 Feb 2019

Downloaded from <http://pubs.acs.org> on February 11, 2019

### Just Accepted

“Just Accepted” manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides “Just Accepted” as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. “Just Accepted” manuscripts appear in full in PDF format accompanied by an HTML abstract. “Just Accepted” manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). “Just Accepted” is an optional service offered to authors. Therefore, the “Just Accepted” Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these “Just Accepted” manuscripts.



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20 **Selecting a Single Stereocenter: The Molecular Nuances that Differentiate  $\beta$ -Hexuronidases in**  
21 **the Human Gut Microbiome**  
22  
23  
24

25 Samuel J. Pellock<sup>1,\*,#</sup>, William G. Walton<sup>1,#</sup>, and Matthew R. Redinbo<sup>1,2,3,4,\*</sup>

26  
27 Department of Chemistry<sup>1</sup>, Department of Biochemistry and Biophysics<sup>2</sup>, Department of Microbiology  
28 and Immunology<sup>3</sup>, and Integrated Program in Biological and Genome Sciences<sup>4</sup>, University of North  
29 Carolina at Chapel Hill  
30

31  
32 \*Corresponding Authors: [pellock@email.unc.edu](mailto:pellock@email.unc.edu), [redinbo@unc.edu](mailto:redinbo@unc.edu)  
33

34 #Both authors contributed equally to the results of this work  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**ABSTRACT**

The human gut microbiome is a ripe space for the discovery of new proteins and novel functions. Many genes in the gut microbiome encode glycoside hydrolases that help bacteria scavenge sugars present in the human gut. Glycoside hydrolase family 2 (GH2) is one group of sugar-scavenging proteins, which include  $\beta$ -glucuronidases (GUS) and  $\beta$ -galacturonidases (GalAses), enzymes that cleave the sugar conjugates of the epimers glucuronate and galacturonate. Here we structurally and functionally characterize a GH2 GalAse and a hybrid GUS/GalAse, which reveal the molecular details that enable these GHs to differentiate a single stereocenter. First, we characterized a previously annotated GUS from *Eisenbergiella tayi* and demonstrated that it is, in fact, a GalAse. We determined the crystal structure of this GalAse, identified the key residue that confers GalAse activity, and convert this GalAse into a GUS by mutating a single residue. We performed bioinformatic analysis of 279 putative GUS enzymes from the human gut microbiome and identified 12 additional putative GH2 GalAses, one of which we characterized and confirmed is a GalAse. Lastly, we report the structure of a hybrid GUS/GalAse from *Fusicatenibacter saccharivorans*, which revealed a unique hexamer that positions the N-terminus of adjacent protomers in the aglycone binding site. Taken together, these data reveal a new class of bacterial GalAses in the human gut microbiome and unravel the structural details that differentiate GH2 GUSs and GalAses.

1  
2  
3 The human gut microbiome encodes a wide range of glycoside hydrolases (GHs) that play key  
4 roles in host health and disease. A prominent group of GHs in the gut microbiome is GH family 2 (GH2),  
5 which includes the well-studied  $\beta$ -glucosidases and  $\beta$ -galactosidases, as well as their sugar acid-cleaving  
6 counterparts  $\beta$ -glucuronidases (GUS) and  $\beta$ -galacturonidases (GalAse)<sup>1</sup>. Analogous to glucose and  
7 galactose released by  $\beta$ -galactosidases and  $\beta$ -glucosidases, glucuronate (GlcA) and galacturonate (GalA)  
8 released by bacterial GUSs and GalAses also serve as a source of energy. Bacteria can catabolize GlcA and  
9 GalA via either an isomerase or oxidative catabolic pathway<sup>2</sup>. GlcA and GalA are present in host-derived  
10 and plant-based polysaccharides, respectively, including glycosaminoglycans in animals and pectin from  
11 plants<sup>3,4</sup>. In addition to polysaccharide utilization, bacterial GUS enzymes are important because they play  
12 key roles in reversing mammalian drug metabolism<sup>5</sup>. For example, a variety of drugs are glucuronidated  
13 in the liver, a process by which uridine diphosphate glucuronosyl transferases (UGTs) attach GlcA to  
14 available nucleophilic moieties<sup>6</sup>. Bacterial GUS enzymes can metabolize drug glucuronides generated by  
15 the liver and release active and sometimes toxic drug into the gastrointestinal tract. Indeed, the toxic side  
16 effects of cancer drugs and NSAIDs are alleviated by selective bacterial GUS inhibition<sup>7-10</sup>. Together, the  
17 key roles that gut microbial GUSs and GalAses play in polysaccharide and drug metabolism in the human  
18 gut make them important proteins for structural and functional characterization.

19  
20 We previously screened the human microbiome project (HMP) stool sample metagenomic  
21 database for putative GH2 GUSs<sup>11</sup>. This bioinformatic effort yielded an atlas of 279 putative GUSs, of which  
22 only a handful have been characterized either structurally or functionally to date (**Fig. S1**)<sup>11-13</sup>. During the  
23 characterization of *BuGUS-1* from *Bacteroides uniformis*, we found that, in addition to processing  $\beta$ -  
24 glucuronides, *BuGUS-1* also hydrolyzed  $\beta$ -galacturonides<sup>13</sup>, suggesting that this gut microbial GUS, and  
25 others yet to be identified, may process sugar conjugates beyond glucuronides. Indeed, using novel crystal  
26 structures and site-directed mutagenesis as a guide, here we show that there are 13 GalAses and 10 hybrid  
27 GUS/GalAses in the 279 previously annotated GUSs from the HMP stool sample database. Further kinetic  
28 and structural studies revealed that variations on a shared active site structure allow GUSs and GalAses  
29 to selectively process epimeric substrates.

## 30 **MATERIALS AND METHODS**

### 31 **Gene synthesis, mutagenesis, protein expression, and protein purification**

32 All genes were synthesized by BioBasic and cloned into a pUC57 vector. The genes encoding  
33 *EtGalAse*, *FsGUS*, and *FcGalAse* were then sub-cloned into a pET His6 LIC cloning vector (2Bc-T) for  
34 expression with a C-terminal His tag. *FsGUS* was also cloned into the pLIC-His vector pMCSG7 and  
35 expressed with an N-terminal tag for crystallography. Cloning primers are listed in **Table S1**. All mutants  
36 were generated by standard site-directed mutagenesis and primers utilized for mutagenesis were  
37 purchased from Integrated DNA Technologies (IDT) and are listed in **Table S2**. All cloning and mutagenesis  
38 experiments were confirmed by DNA sequencing by Eton Bioscience Inc. Proteins were purified as  
39 described previously<sup>11</sup>. Briefly, all proteins were expressed via standard IPTG induction with growth  
40 overnight at 18 °C. Cell lysates from expression were clarified and subject to Ni-based affinity  
41 chromatography and size exclusion chromatography. Final protein concentrations were determined with  
42 a ND-1000 spectrophotometer. Proteins were snap frozen in liquid nitrogen and stored at -80 °C.

### 43 **Crystallography**

1  
2  
3 All crystals were grown by sitting drop vapor diffusion in Hampton Research 3-well Midi  
4 Crystallization Plates (Swissci) by an Art Robbins Instruments Crystal Phoenix. All crystals were grown at  
5 20 °C in the following crystallants: ***EtGalAse*** in 1 M NaH<sub>2</sub>PO<sub>4</sub>/K<sub>2</sub>HPO<sub>4</sub>, pH 8.2 and 1 mM GalA, ***EtGalAse***  
6 **bound to GalA** in 0.2 M ammonium sulfate, 0.1 M Bis-Tris:HCl, pH 6.5, 25% (w/v) PEG 3350, 1 mM GalA,  
7 **apo *FsGUS* (N-terminal His tag)** in 0.15 M DL-malic acid, pH 7.0, 20% w/v PEG 3350, and ***FsGUS* (C-terminal**  
8 **His tag) bound to phenyl-thio-β-D-glucuronide (PTG)** in 0.2 M sodium tartrate dibasic, 20% (w/v) PEG  
9 3350, and 1 mM PTG. Resultant crystals were cryoprotected in their corresponding crystallants with 20%  
10 glycerol added. For liganded structures, 1 mM of ligand was also present in the cryoprotectant solution.  
11 Diffraction data were collected for all crystals at 100 K at APS Beamline 23-ID-D. Data were reduced in XDS  
12 and scaled in Aimless<sup>14,15</sup>. The *EtGalAse* structures were solved by molecular replacement in Phenix using  
13 a truncated model of *E. coli* GUS (PDB: 3LPF) as the search model. The *FsGUS* structures were also solved  
14 via molecular replacement in Phenix and *BuGUS-1* (PDB: 6D1N) was used as the search model. Subsequent  
15 refinement and ligand placements were performed in Phenix. Final coordinates and maps were deposited  
16 in the PDB for *EtGalAse* (PDB: 6NCW), *EtGalAse* bound to GalA (PDB: 6NCX), *FsGUS* (PDB: 6NCY), and  
17 *FsGUS* bound to PTG (PDB: 6NCZ) (**Table S3**).

### 21 22 **Kinetic assays for *pNP-GalA* and *pNP-GlcA* processing**

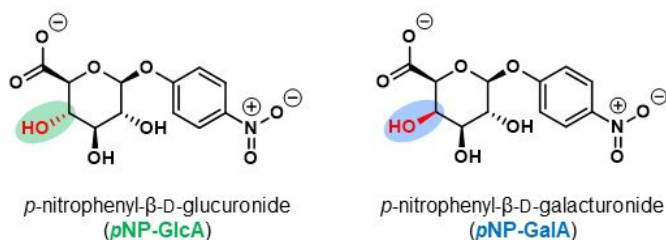
23  
24 Assays to determine the catalytic parameters of GUS and GalAse activity were performed with  
25 the chromogenic substrates *p*-nitrophenyl-β-D-galacturonide (*pNP-GalA*) and *p*-nitrophenyl-β-D-  
26 glucuronide (*pNP-GlcA*). Briefly, assays contained 5 μL enzyme, 5 μL 250 mM NaCl and 250 mM HEPES (pH  
27 6.5 and 7) or 5 μL 250 mM NaCl and 250 mM sodium acetate (pH 5.5), and 40 μL substrate. For reactions  
28 that showed no activity, enzyme concentrations up to 300 nM were tested. For reactions with measurable  
29 activity, enzyme concentrations were as follows: *EtGalAse* WT (5 nM), *EtGalAse* R337A (300 nM), *FcGalAse*  
30 WT (5 nM), *FsGUS* WT (5 nM), *FsGUS* Y377F (20 nM), and *FsGUS* Δ16 (20 nM). Given the lower water  
31 solubility of *pNP-GalA*, all kinetic assays contained 1.6% DMSO final. For pH values greater than or equal  
32 to 7, reactions were monitored continuously at 410 nm and incubated at 37 °C in Costar 96-well black,  
33 flat, clear bottom plates in a BMG lab tech PHERAstar plate reader. For enzyme reactions below pH 6.5,  
34 reactions were quenched with 100 μL of 0.2 M sodium carbonate and measured as described for  
35 continuous monitoring. Initial velocities were determined in MATLAB and Michaelis-Menten fits were  
36 performed in SigmaPlot 13.0.

### 37 38 39 40 **Kinetic assays for SN-38-G processing**

41  
42 SN-38-G processing was measured by a fluorescence-based assay as described previously<sup>13</sup>.  
43 Briefly, assays were performed in Corning 96-well clear, half-area, UV-transparent plates that contained  
44 5 μL enzyme, 5 μL 250 mM NaCl and 250 mM HEPES (pH 7), 5 μL substrate, and 35 μL water. Reactions  
45 were pre-incubated at 37 °C for ten minutes before addition of enzyme. Final DMSO concentrations were  
46 4% for all reactions. Data were fit as described for the *pNP-GlcA* and *pNP-GalA* assays above.

### 47 48 49 **Circular dichroism**

50  
51 Circular dichroism experiments were performed as previously described<sup>13</sup>. In brief, WT and  
52 mutant proteins (2 μM) were prepared in CD buffer (10 mM potassium phosphate, 100 mM potassium  
53 fluoride, pH 7.4) and loaded into a 1-mm cuvette. Both spectra scans and melting temperatures were  
54 performed using a Chirascan-plus instrument (Applied Photophysics Limited).



**Figure 1. Chemical structures of  $\beta$ -nitrophenyl- $\beta$ -D-glucuronide (pNP-GlcA) and  $\beta$ -nitrophenyl- $\beta$ -D-galacturonide (pNP-GalA) with stereochemical difference highlighted in red.**

### Size exclusion chromatography-multi-angle light scattering

Oligomeric states of purified proteins in solution were determined as previously described with slight modifications<sup>13</sup>. Briefly, proteins were analyzed on a Superdex 200 size exclusion column connected to an Agilent FPLC system, Wyatt DAWN HELEOS II multi-angle light scattering instrument and a Trex

refractometer. Elution buffers were the same as employed for size exclusion chromatography (20 mM HEPES, 50 mM NaCl, pH 7.4). All proteins were run at approximately 5 mg/ml. Light scattering and refractive index data were collected and analyzed using Wyatt ASTRA (Ver. 6.1) software.

### Sequence similarity network and genome neighborhood diagram generation

Sequence similarity networks (SSNs) were generated using the 279 putative GUS protein sequences previously identified by a structure- and function-guided search of the HMP stool sample metagenomic database<sup>11</sup>. The 279 sequences were used as input into the FASTA (Option C) section of the web-based Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST)<sup>16</sup> to determine initial alignment scores for the generation of SSNs. Resultant SSNs at multiple different alignment scores were visualized in Cytoscape 3.7.0<sup>17</sup>. Genome neighborhood diagrams (GNDs) were generated with the web-based EFI-Genome Neighborhood Network Tool<sup>16</sup>. UniProt IDs of the proteins of interest were used as input to retrieve the associated GNDs. For protein sequences without a definitive match in either the UniProtKB or NCBI databases, the highest identity match in the UniProtKB database was utilized to generate a GND.

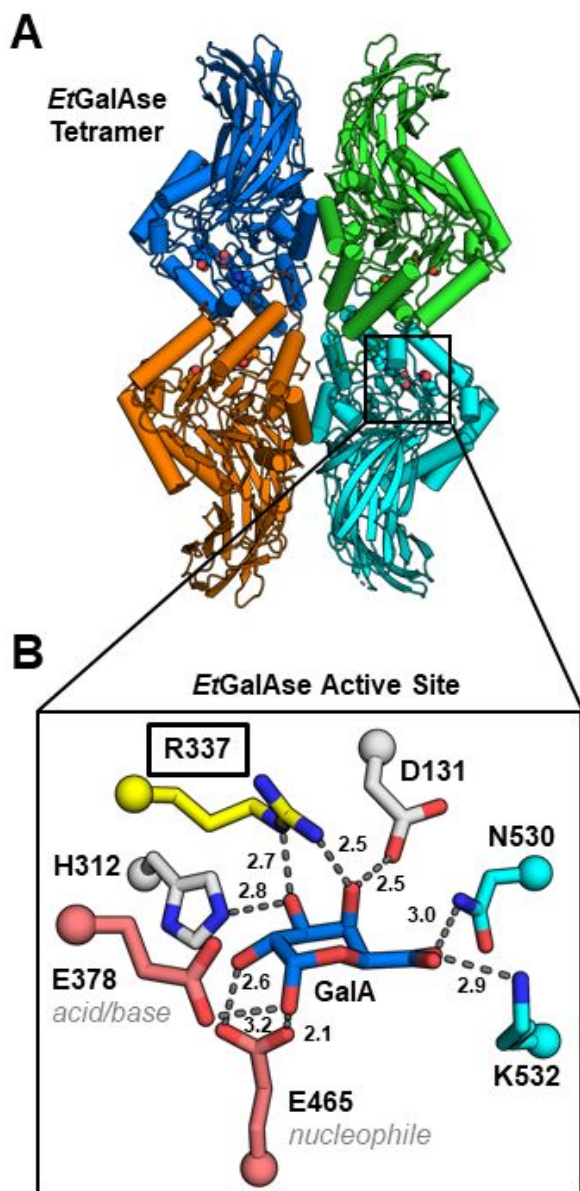
**Table 1. Kinetic parameters of pNP-GlcA and pNP-GalA processing.**

Enzyme	pNP-GlcA			pNP-GalA		
	$k_{cat}$ ( $s^{-1}$ )	$K_M$ ( $\mu M$ )	$k_{cat}/K_M$ ( $s^{-1} M^{-1}$ )	$k_{cat}$ ( $s^{-1}$ )	$K_M$ ( $\mu M$ )	$k_{cat}/K_M$ ( $s^{-1} M^{-1}$ )
<i>EtGalAse</i>	NA	NA	NA	98 $\pm$ 3	160 $\pm$ 6	6.2x10 <sup>5</sup>
<i>EtGalAse</i> R337A	1.0 $\pm$ 0.1	110 $\pm$ 20	5.2x10 <sup>3</sup>	NA	NA	NA
<i>FcGalAse</i>	NA	NA	NA	42 $\pm$ 3	58.8 $\pm$ 0.2	7.1x10 <sup>5</sup>
<i>FcGalAse</i> R335A	NA	NA	NA	NA	NA	NA
<i>FsGUS</i>	15 $\pm$ 1	30 $\pm$ 6	5.0x10 <sup>5</sup>	34 $\pm$ 2	360 $\pm$ 20	9.4x10 <sup>4</sup>
<i>FsGUS</i> Y377F	8.9 $\pm$ 0.4	55 $\pm$ 4	1.6x10 <sup>5</sup>	--	--	<sup>a</sup> 8.9x10 <sup>3</sup>
<i>FsGUS</i> $\Delta$ 16	3.6 $\pm$ 0.6	70 $\pm$ 10	5.1x10 <sup>4</sup>	14 $\pm$ 1	360 $\pm$ 60	3.9x10 <sup>4</sup>

NA = no activity. <sup>a</sup>Catalytic efficiency determined from non-saturating substrate concentrations. Values shown are an average of n = 3 replicates  $\pm$  SD.

## RESULTS

### Putative GUS from *Eisenbergiella tayi* is instead a GH2 GalAse



**Figure 2. Crystal structure of the GH2  $\beta$ -galacturonidase *EtGalAse*.** (A) Quaternary structure of *EtGalAse* shown in cartoon style with active site residues shown as spheres. (B) Active site of *EtGalAse* bound to GalA (blue) with catalytic glutamates shown in deep salmon, NxK motif in cyan, and the unique R337 in yellow. All distances shown are in units of angstroms.

unique feature of the *EtGalAse* active site, however, is Arg-337, which hydrogen bonds to the axial 4-hydroxyl of GalA, the stereocenter that differentiates the epimers GalA and GlcA (Fig. 1 and Fig. 2B). To assess the role of Arg-337 in GalAse function, we mutated Arg-337 to alanine in *EtGalAse*. The equivalent residue in established GUS enzymes is observed to be valine (*EcGUS*, *CpGUS*, *SaGUS*, and *EeGUS*) or isoleucine (*BuGUS-2*)<sup>7,9,11,18</sup>. Remarkably, the R337A mutant of *EtGalAse* both abolished GalAse activity (Fig. S3C) and conferred GUS activity ( $k_{cat} = 1.0 \pm 0.1 \text{ s}^{-1}$ ,  $K_M = 110 \pm 20 \text{ }\mu\text{M}$ ,  $k_{cat}/K_M = 5.2 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$ ) (Table

The vast majority of the previously annotated 279 GUS enzymes in the HMP stool sample database remain uncharacterized (Fig. S1)<sup>11</sup>. Thus, we have performed gene synthesis, expression, and purification of a range of these putative GUSs to determine their structure and function. We expressed and purified the putative GUS from the gut microbe *Eisenbergiella tayi* (UniProtKB ID: A0A1E3AEY6), but it did not process the standard reporter glucuronide *pNP-GlcA* (Fig. 1 and Table 1). However, we recently showed that a GUS from the gut microbe *Bacteroides uniformis*, termed *BuGUS-1*, processed *p*-

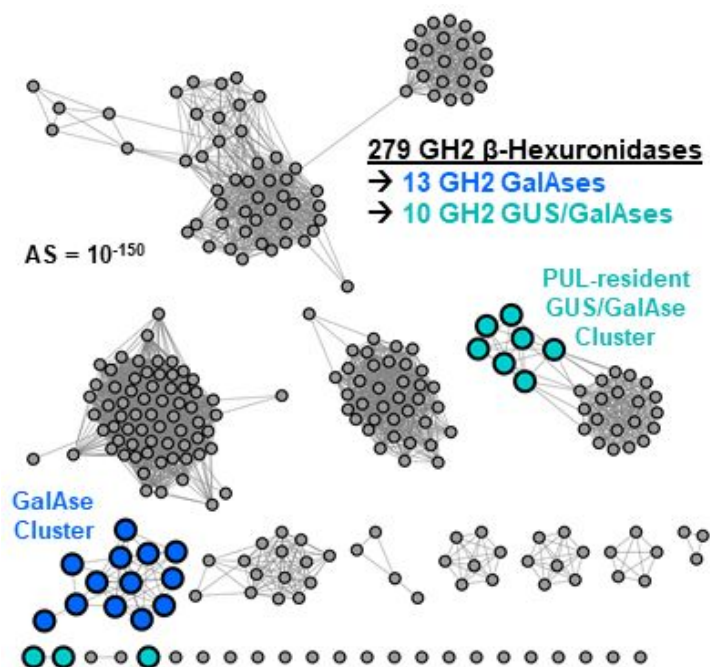
nitrophenyl- $\beta$ -D-galacturonide (*pNP-GalA*) as well as *pNP-GlcA* (Fig. 1)<sup>13</sup>. Thus, we tested if the putative GUS from *E. tayi* could also process *pNP-GalA*. Indeed, the *E. tayi* protein processed *pNP-GalA* efficiently with optimal activity at pH 6 ( $k_{cat} = 98 \pm 3 \text{ s}^{-1}$ ,  $K_M = 160 \pm 6 \text{ }\mu\text{M}$ ,  $k_{cat}/K_M = 6.2 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ ) (Table 1, Fig. S2A, and Fig. S3A). The catalytic efficiency of the *E. tayi* enzyme (*EtGalAse*) with *pNP-GalA* is similar to *pNP-GlcA* cleavage by *bona fide* GUS enzymes<sup>9</sup>. The results presented here demonstrate that the putative GUS from *E. tayi*, annotated as a GUS by sequence analysis, is instead a GH2 GalAse.

### Mutagenesis of a single residue converts *EtGalAse* into a GUS

We determined the 2.10 Å and 2.25 Å resolution crystal structures of glycerol-bound and GalA-bound *EtGalAse* (Table S3). *EtGalAse* is a tetramer (Fig. 2A) similar to previously characterized GUS enzymes<sup>7,9</sup>. The active site also resembles previously characterized GUSs, as it contains the key catalytic glutamates, Glu-378 (acid/base) and Glu-465 (nucleophile), and the NxK motif previously identified as essential for GUS activity (Fig. 2B)<sup>13</sup>. A

1  
2  
3 **1, Fig. S2B, and Fig. S3B).** Together, these structural and functional data pinpoint Arg-337 as the key active site feature that differentiates a GH2 GalAse from a GH2 GUS.

### Structure-guided analysis of putative GUSs from the HMP reveals 12 additional putative GH2 GalAses



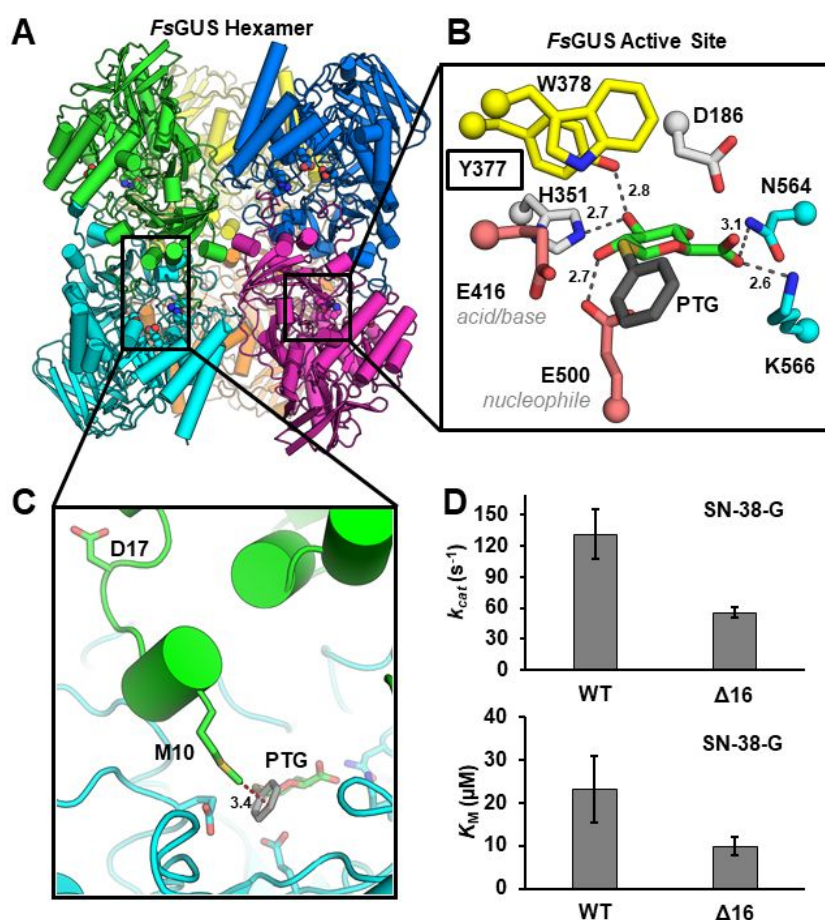
**Figure 3. Sequence similarity network (SSN) generated from putative GUS sequences from the HMP stool sample metagenomic database reveals distinct clusters for the 13 putative GH2 GalAses and 10 putative hybrid GUS/GalAses.** GH2 GalAse cluster proteins colored blue and hybrid GUS/GalAses in cyan. The SSN was generated with the EFI-EST web tool with an alignment score of  $10^{-150}$ .

of a set of sequence similarity networks using the 279 putative GUS sequences from the HMP showed that the putative GalAses cluster together at an alignment score of  $10^{-150}$ , suggesting they share a common function (**Fig. 3** and **Fig. S4**)<sup>16</sup>. This is a lower alignment score (less stringent) than utilized previously to create an SSN of the human gut microbial GUSome ( $10^{-220}$ ; **Fig. S1**)<sup>19</sup>. While this lower alignment ( $10^{-150}$ ) score yielded SSNs that did not differentiate many of the putative GUS enzymes as effectively, it did group the putative GalAses together that were mostly singletons at an alignment score of  $10^{-220}$ , suggesting they are likely isofunctional (**Fig. 3**; **Fig. S1**).

To experimentally validate our structure- and function-guided GalAse discovery, we characterized a putative GalAse from *Faecalibacterium* sp. CAG:74\_58\_120 (*FcGalAse*, UniProtKB: A0A1Q6Q230). *FcGalAse* hydrolyzed *pNP*-GalA efficiently ( $k_{cat} = 42 \pm 3 \text{ s}^{-1}$ ,  $K_M = 58.8 \pm 0.2 \text{ }\mu\text{M}$ ,  $k_{cat}/K_M = 7.1 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ ; **Table 1, Fig. S2C, and Fig. S3D**) and was unable to process *pNP*-GlcA (**Fig. S5A**), validating our bioinformatic identification of GH2 GalAses. The R335A mutant of *FcGalAse* abolished galacturonide processing; however, unlike *EtGalAse*, this single residue mutation did not convert *FcGalAse* into a GUS (**Table 1**). A structural model of *FcGalAse* derived from the *EtGalAse* structure suggests that the absence of GUS activity for the R335A variant of *FcGalAse* may be due to the presence of Leu-409 in *FcGalAse* instead of Cys-411 in *EtGalAse*, which may contribute to a less stable active site in the absence of Arg-335 (**Fig. S6**). Further structural analysis by circular dichroism (CD) showed that the R335A mutant of *FcGalAse* displays reduced secondary structure in comparison to WT protein (**Fig. S7C**). No difference was observed in

secondary structure for WT and R337A *EtGalAse* (Fig. S7A). WT *FcGalAse* and the R335A mutant do not have significantly different melting temperatures, but the *EtGalAse* proteins do not appear to melt at all (Fig. S7B and S7D). While future studies will be needed to pinpoint the molecular details that enable some GalAses to be converted into GUSs, CD analysis suggests that structural differences between *FcGalAse* and *EtGalAse* may mediate these functional differences.

### Structure and function a hexameric hybrid GUS/GalAse from *Fusicatenibacter saccharivorans*



**Figure 4. Structure and function of *FsGUS* reveal Y377 and N-terminus as important structural features for substrate processing.** (A) The *FsGUS* hexamer with each chain in a distinct color and active site residues shown as spheres. (B) Zoom-in of active site of *FsGUS* bound to phenyl-thio- $\beta$ -D-glucuronide (PTG, phenyl in dark grey, GlcA in green) with catalytic glutamates colored deep salmon, NxK motif in cyan, and unique YW motif in yellow. All distances shown are in units of angstroms. (C) Active site of *FsGUS* with PTG bound and N-terminus of adjacent protomer shown in green. (D) Catalytic turnovers and Michaelis constants for WT and  $\Delta 16$  mutant of *FsGUS* with SN-38-G. Values represent an average of three replicates  $\pm$  SD.

Previous analysis of a GUS from *Bacteroides uniformis*, *BuGUS-1*, revealed that it processes both glucuronides and galacturonides, and that it possessed a distinct tyrosine and tryptophan residue in its active site (YW motif) in comparison to previously characterized GUS enzymes<sup>13</sup>. Using the YW motif as a sequence guide, we identified 9 additional hybrid GUS/GalAses from the 279 putative GUS sequences from the HMP stool sample database, 6 of which had been previously annotated as containing the N-terminal loop (NTL) discovered in *BuGUS-1* from *Bacteroides uniformis* (Fig. 3 and Table S5)<sup>13</sup>. We selected the putative GUS/GalAse from *Fusicatenibacter saccharivorans* (*FsGUS*, UniProtKB: A0A174EHD1) for further characterization. Like *BuGUS-1*, *FsGUS* processed both *pNP-GlcA* ( $k_{cat} = 15 \pm 1 s^{-1}$ ,  $K_M = 30 \pm 6 \mu M$ ,  $k_{cat}/K_M = 5.0 \times 10^5 M^{-1} s^{-1}$ ) and *pNP-GalA* ( $k_{cat} = 34 \pm 2 s^{-1}$ ,  $K_M = 360 \pm 20 \mu M$ ,  $k_{cat}/K_M = 9.4 \times 10^4 M^{-1} s^{-1}$ , Table 1, Fig. S5C, and Fig. S5D), validating our bioinformatic search.

We next determined the crystal structure of the apo and phenyl-thio- $\beta$ -D-glucuronide (PTG)-bound structures of *FsGUS* to 2.05 Å and 2.20 Å resolution, respectively (Table S3). The crystal structure of *FsGUS* revealed a unique hexamer among characterized GUS enzymes (Fig. 4A), which we validated in

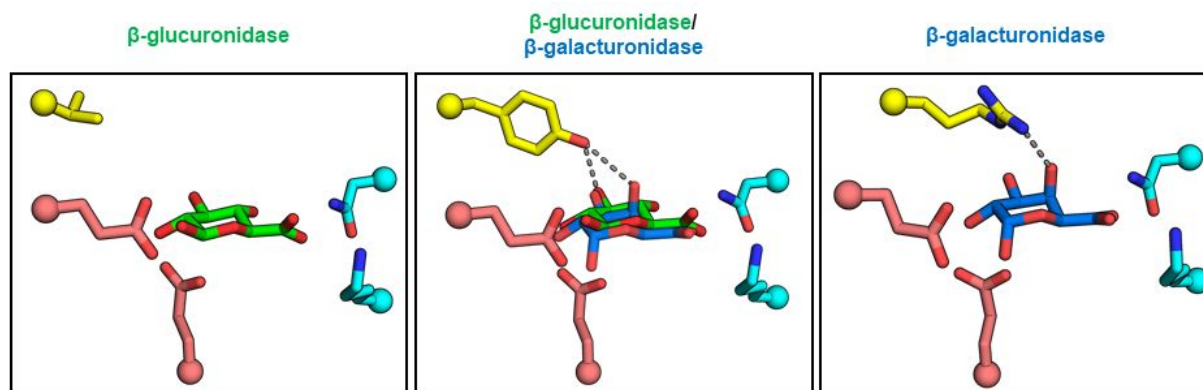
1  
2  
3 solution by size exclusion chromatography-multi-angle light scattering (SEC-MALS) (**Fig. S8**). The PTG-  
4 bound structure of *FsGUS* reveals a shift in the catalytic acid/base Glu-416, presumably to avoid a steric  
5 clash with the thioether linkage in PTG (**Fig. S9**). The *FsGUS* active site also reveals Tyr-377 and Trp-378  
6 previously identified in *BuGUS-1* (**Fig. 4B**). Tyr-377 is key in recognizing both GlcA and GalA because it can  
7 hydrogen bond to the 3-hydroxyl of GlcA (**Fig. 4B**) and can also hydrogen bond to the axial 4-hydroxyl of  
8 GalA when docked in the *FsGUS* active site (**Fig. S10**). Tyr-377 in *FsGUS* occupies nearly the same position  
9 as Arg-337 in *EtGalAse*; thus, we were curious if it aided hydrolysis of glucuronides and galacturonides.  
10 Indeed, mutation of Tyr-377 to phenylalanine in *FsGUS* results in a ten-fold lower catalytic efficiency for  
11 *pNP-GalA* ( $k_{cat}/K_M = 8.9 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$ ) and 3-fold lower catalytic efficiency for *pNP-GlcA* ( $k_{cat}/K_M = 1.6 \times 10^5 \text{ M}^{-1}$   
12  $\text{s}^{-1}$ ) compared to WT *FsGUS* (*pNP-GalA*:  $k_{cat}/K_M = 9.4 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$ , *pNP-GlcA*:  $k_{cat}/K_M = 5.0 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ ) (**Table**  
13 **1**, **Fig. S5C**, **Fig. S5D**, and **Fig. S11**). The greater loss of catalytic efficiency in GalAse function upon mutation  
14 of Tyr-377 suggests that it is particularly important for the processing of GalA-containing substrates.  
15 Furthermore, we performed the analogous Y382F mutation in the previously characterized *BuGUS-1*,  
16 which resulted in a complete loss of activity against both *pNP-GlcA* and *pNP-GalA* (**Fig. S11**). Together, this  
17 structure-function analysis of *FsGUS* reveals the first hexameric GUS and identifies Tyr-377 as an  
18 important residue for efficient processing of both glucuronides and galacturonides.  
19  
20  
21  
22

### 23 **The N-termini of adjacent protomers form the aglycone binding sites of the *FsGUS* hexamer**

24  
25 Further inspection of the *FsGUS* hexamer revealed that the N-termini of adjacent protomers swap  
26 into the active site of each monomer (**Fig. 4C**). For example, in the PTG-bound structure of *FsGUS*, Met-  
27 10 from an adjacent protomer (green) is 3.4 Å from the phenyl ring of PTG, suggesting that it may be  
28 important to substrate recognition of the aglycone moieties of potential glucuronide or galacturonide  
29 substrates (**Fig. 4C**). We performed mutagenesis to remove the first 16 residues ( $\Delta 16$ ) of *FsGUS* to assess  
30 the role the N-terminus plays in substrate processing (**Fig. 4C**). The  $\Delta 16$  *FsGUS* mutant elutes at the same  
31 time as WT *FsGUS* by size-exclusion chromatography, suggesting that this mutant occupies the same  
32 hexamer as WT *FsGUS* (**Fig. S8**). Kinetic analysis of  $\Delta 16$  *FsGUS* revealed a ten-fold and two-fold reduction  
33 in catalytic efficiency for *pNP-GlcA* and *pNP-GalA*, respectively (**Table 1**). Thus, it appears the N-terminus  
34 of *FsGUS* plays a key role in efficient substrate processing.  
35  
36  
37

38 Because *FsGUS* was able to efficiently process the small-molecule glucuronide *pNP-GlcA*, we  
39 tested whether it could also process the therapeutically relevant glucuronide SN-38-G, the inactive  
40 metabolite of the anticancer drug irinotecan, and our original impetus for studying bacterial GUSs from  
41 the gut microbiome<sup>7</sup>. *FsGUS* was able to process SN-38-G efficiently ( $k_{cat} = 130 \pm 20 \text{ s}^{-1}$ ,  $K_M = 23 \pm 8 \text{ }\mu\text{M}$ ,  
42 **Fig. 4D**). We also tested the  $\Delta 16$  mutant of *FsGUS* with SN-38-G, which displayed a reduced  $k_{cat}$  and  $K_M$   
43 ( $k_{cat} = 56 \pm 5 \text{ s}^{-1}$ ,  $K_M = 10 \pm 2 \text{ }\mu\text{M}$ ) in comparison to WT *FsGUS* (**Fig. 4D**). Lastly, we examined if the WT and  
44 R337A variant of *EtGalAse* could process SN-38-G. As expected, WT *EtGalAse* was unable to process SN-  
45 38-G efficiently but the R337A mutant displayed activity ( $k_{cat} = 2.3 \pm 0.2 \text{ s}^{-1}$ ,  $K_M = 39 \pm 6 \text{ }\mu\text{M}$ ,  $k_{cat}/K_M =$   
46  $6.3 \times 10^4 \text{ s}^{-1} \text{ M}^{-1}$ ) (**Fig. S12**).  
47  
48

49 Given the unique structural features of *FsGUS*, we performed a final analysis of the 279 putative  
50 GUS enzymes to determine the potential distribution and frequency of *FsGUS*-like proteins in the gut  
51 microbiome. Unlike our search for putative GalAses or GUS/GalAses, which were predicated on a  
52 conserved arginine and tyrosine-tryptophan motif, the unique *FsGUS* hexamer and N-terminus do not  
53 appear to be conserved sequence features. That is, we were unable to identify common sequence motifs  
54 that are responsible for these unique structural features. Thus, we performed pairwise sequence  
55  
56  
57



**Figure 5. Active site nuances differentiate gut microbial GUSs, hybrid GUS/GalAses, and GalAses.** Representative active site structures of GUS (PDB: 4JKL [*Sa*GUS]), GUS/GalAse (*Fs*GUS), and GalAse (*Et*GalAse) with catalytic glutamates in deep salmon, NxK motif in cyan, key differentiating residues in yellow, GlcA in green, and GalA in blue.

alignments with the complete *Fs*GUS sequence against the remaining 278 putative  $\beta$ -hexuronidases to identify potential *Fs*GUS-like proteins. Alignments revealed only one close relative to *Fs*GUS, a sequence with no clear match to a genome termed GUS/GalAse-3 (**Table S5**) that shares approximately 68% sequence identity with *Fs*GUS. Based on this analysis, there only appear to be two unique *Fs*GUS-like proteins in the human gut microbiome.

## DISCUSSION

The GalAses and GUS/GalAses characterized here reveal modifications on a shared active site structure to differentiate the epimers GalA and GlcA. The enzymes characterized here were originally discovered using two features thought to be specific for GUS activity, the catalytic glutamates and NxK motif<sup>11</sup>. We show here that this structure- and function-guided bioinformatic analysis yielded  $\beta$ -hexuronidases in general, which modify this active site template to yield similar, yet unique activities. For example, the GalAses and GUS/GalAses characterized here also utilize the catalytic glutamates and NxK motif, but the remaining active site residues recognize the other hydroxyls present in GlcA and GalA and are what differentiates GUS, hybrid GUS/GalAses, and GalAses (**Fig. 5**). In bacterial GUS enzymes, the equatorial 4-hydroxyl of GlcA is recognized by an aspartic acid and a tryptophan<sup>13</sup>. The hybrid GUS/GalAses and GalAses characterized here, in which the 4-hydroxyl of their GalA substrate is axial, tyrosine (GUS/GalAse) or arginine (GalAse) residues are positioned above GalA and can hydrogen bond to the 4-hydroxyl (**Fig. 5**). The active site architectures and ‘rules’ of selectivity observed could be used to discover  $\beta$ -hexuronidases in the gut microbiome, to design novel  $\beta$ -hexuronidases, and/or to change the function of existing  $\beta$ -hexuronidases.

The crystal structures presented here reveal active site and quaternary structures unique among characterized GUS enzymes. The tertiary structures of *Et*GalAse and *Fs*GUS exhibit the same GH2 fold observed for *E. coli* GUS (*Ec*GUS) and the other GUS enzymes of known structure, with a core TIM-barrel fold and two  $\beta$ -sandwich domains<sup>7,9,11</sup>. Structural alignments of *Et*GalAse and *Fs*GUS with *Ec*GUS reveal root mean square deviations of 2.6 Å over 520 C $\alpha$  carbons and 2.9 Å over 528 C $\alpha$  carbons, respectively (**Fig. S13**). SEC-MALS shows that *Et*GalAse is a tetramer in solution (**Fig. S14**), and the crystal structure reveals that the tetramer interface is mediated by the C-termini of individual protomers like that observed for *Ec*GUS (**Fig. S15**). While *Fs*GUS maintains a similar tertiary structure to *Ec*GUS, it has a hexameric

1  
2  
3 quaternary structure not observed in previously characterized GUSs or any other GH2 enzymes to our  
4 knowledge (**Fig. 4A** and **Fig. S16**). As observed for other GUSs of distinct structures, the quaternary  
5 structure of this enzyme plays a key role in forming the active site and thus affects activity (**Fig. 4D**). Like  
6 the NTL discovered in *BuGUS-1*, the N-terminus of adjacent protomers in *FsGUS* contribute to the  
7 formation of the aglycone binding site. An overlay of *BuGUS-1* and *FsGUS* reveals that these enzymes  
8 utilize distinct structural elements to form the aglycone binding site (**Fig. S17**). Thus, while *BuGUS-1* and  
9 *FsGUS* both share the YW motif that appears to enable hybrid GUS/GalAse activity, they display distinct  
10 oligomeric states (*BuGUS-1* tetramer, *FsGUS* hexamer) and utilize distinct protomer-donated structural  
11 elements to form the aglycone binding sites of their active sites (**Fig. S17**). Taken together, quaternary  
12 structure is critical to understanding the function of bacterial GUS enzymes.  
13  
14  
15

16 The hybrid GUS/GalAses characterized here likely process glucuronate and galacturonate-  
17 containing polysaccharides in the gut. We performed further bioinformatic analysis of the GUS/GalAses  
18 identified from the HMP stool sample database using the EFI-GNT web tool<sup>16</sup>. This revealed that 7 of the  
19 9 identified GUS/GalAses are embedded in polysaccharide utilization loci (PUL), gene clusters that  
20 coordinate the degradation of complex carbohydrates (**Fig. S18**)<sup>4</sup>. These observations support the  
21 conclusion that the putative GUS/GalAses discovered here process polysaccharides that contain  $\beta$ -linked  
22 glucuronate or galacturonate moieties.  
23  
24

25 *EtGalAse* is functionally related to the previously characterized GH2 GalAse from *B.*  
26 *thetaitotaomicron*, BT\_0992, which was shown to process the  $\beta$ -GalA linkage present in  
27 rhamnogalacturonan-II (RG-II)<sup>20</sup>. While *EtGalAse* and BT\_0992 share 26% sequence identity (**Fig. S19**),  
28 BT\_0992 does not encode the conserved NxK motif and is much longer than the GH2 GalAses  
29 characterized here, suggesting that BT\_0992 likely utilizes a distinct active site architecture to recognize  
30 and cleave  $\beta$ -GalA linkages. Analysis of the genome neighborhoods of the GH2 GalAses identified here  
31 revealed a variable genetic context. The genomic neighborhoods of the GH2 GalAses from *E. taylori* and  
32 *Fusicatenibacter sp. 2789STDY5834925A* contain carbohydrate active enzymes, including  $\beta$ -L-  
33 arabinofuranosidase,  $\alpha$ -L-fucosidase,  $\alpha$ -L-rhamnosidase, and  $\beta$ -galactosidase, but are not located in a  
34 polysaccharide utilization locus (**Fig. S20**). While the linkages that these enzymes cleave are present in RG-  
35 II, it is not clear that RG-II is the native substrate of the GH2 GalAses identified in this study without  
36 extensive experimental validation. The genomic neighborhoods of the remaining 11 GH2 GalAses do not  
37 point to obvious potential substrates (**Fig. S20**). Future studies on the gut microbial GH2 GalAses identified  
38 will be required to determine cognate substrates.  
39  
40  
41  
42

## 43 **CONCLUSION**

44 Here we show that gut bacterial GH2 enzymes utilize subtle active site changes to differentiate  
45 between the epimers glucuronate and galacturonate. We present the first structure of a GH2 GalAse and  
46 show that mutating a single residue in *EtGalAse* transforms it into a GUS. Using the structural and  
47 functional data from *EtGalAse* and the previously characterized *BuGUS-1*, we discovered 12 additional  
48 GalAse genes and 9 additional hybrid GUS/GalAses in the previously defined gut bacterial GUSome.  
49 Through these efforts, we identified the molecular determinants that differentiate bacterial GUSs from  
50 GalAses.  
51  
52

## 53 **ASSOCIATED CONTENT**

Supporting information contains figures describing the characterization and analysis of GalAses and GUS/GalAses described in this study.

### **Accession Codes**

UniProtKB IDs for the proteins characterized in this study are as follows: *EtGalAse* (A0A1E3AEY6), *FcGalAse* (A0A1Q6Q230), and *FsGUS* (A0A174EHD1).

### **REFERENCES**

- (1) Cantarel, B. I., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycomics. *Nucleic Acids Res.*
- (2) Bouvier, J. T., Sernova, N. V., Ghasempur, S., Rodionova, I. A., Vetting, M. W., Al-Obaidi, N. F., Almo, S. C., Gerlt, J. A., and Rodionov, D. A. (2018) Novel Metabolic Pathways and Regulons for Hexuronate Utilization in Proteobacteria. *J. Bacteriol.* *201*, e00431-18.
- (3) Dutton, G. J. (1966) Glucuronic Acid, free and combined, biochemistry, pharmacology, and medicine. *Acad. Press.* Academic Press, New York.
- (4) Koropatkin, N. M., Cameron, E. a., and Martens, E. C. (2012) How glycan metabolism shapes the human gut microbiota. *Nat. Rev. Microbiol.* *10*, 323–335.
- (5) Pellock, S. J., and Redinbo, M. R. (2017) Glucuronides in the gut: Sugar-driven symbioses between microbe and host. *J. Biol. Chem.* *292*, 8569–8576.
- (6) Dutton, G. J. (1980) Glucuronidation of drugs and other compounds. CRC press, Boca Raton.
- (7) Wallace, B. D., Wang, H., Lane, K. T., Scott, J. E., Orans, J., Koo, J. S., Venkatesh, M., Jobin, C., Yeh, L. A., Mani, S., and Redinbo, M. R. (2010) Alleviating cancer drug toxicity by inhibiting a bacterial enzyme. *Science (80-. ).* *330*, 831–835.
- (8) LoGuidice, A., Wallace, B. D., Bendel, L., Redinbo, M. R., and Boelsterli, U. A. (2012) Pharmacologic Targeting of Bacterial  $\beta$ -Glucuronidase Alleviates Nonsteroidal Anti-Inflammatory Drug-Induced Enteropathy in Mice. *J. Pharmacol. Exp. Ther.* *341*, 447–454.
- (9) Wallace, B. D., Roberts, A. B., Pollet, R. M., Ingle, J. D., Biernat, K. A., Pellock, S. J., Venkatesh, M. K., Guthrie, L., O’Neal, S. K., Robinson, S. J., Dollinger, M., Figueroa, E., McShane, S. R., Cohen, R. D., Jin, J., Frye, S. V., Zamboni, W. C., Pepe-Ranne, C., Mani, S., Kelly, L., and Redinbo, M. R. (2015) Structure and Inhibition of Microbiome  $\beta$ -Glucuronidases Essential to the Alleviation of Cancer Drug Toxicity. *Chem. Biol.* *22*, 1238–1249.
- (10) Saitta, K. S., Zhang, C., Lee, K. K., Fujimoto, K., Redinbo, M. R., and Boelsterli, U. A. (2014) Bacterial  $\beta$ -glucuronidase inhibition protects mice against enteropathy induced by indomethacin, ketoprofen or diclofenac: Mode of action and pharmacokinetics. *Xenobiotica* *44*, 28–35.
- (11) Pollet, R. M., D’Agostino, E. H., Walton, W. G., Xu, Y., Little, M. S., Biernat, K. A., Pellock, S. J., Patterson, L. M., Creekmore, B. C., Isenberg, H. N., Bahethi, R. R., Bhatt, A. P., Liu, J., Gharaibeh, R. Z., and Redinbo, M. R. (2017) An Atlas of  $\beta$ -Glucuronidases in the Human Intestinal Microbiome. *Structure* *25*, 967–977.
- (12) Little, M. S., Ervin, S. M., Walton, W. G., Tripathy, A., Xu, Y., Liu, J., and Redinbo, M. R. (2018) Active site flexibility revealed in crystal structures of Parabacteroides merdae  $\beta$ -glucuronidase from the human gut microbiome. *Protein Sci.* *27*, 2010–2022.

- (13) Pellock, S. J., Walton, W. G., Biernat, K. A., Torres-Rivera, D., Creekmore, B. C., Xu, Y., Liu, J., Tripathy, A., Stewart, L. J., and Redinbo, M. R. (2018) Three structurally and functionally distinct  $\beta$ -glucuronidases from the human gut microbe *Bacteroides uniformis*. *J. Biol. Chem.* *293*, 18559–18573.
- (14) Kabsch, W. (2010) XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.* *66*, 125–132.
- (15) Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* *69*, 1204–1214.
- (16) Gerlt, J. A., Bouvier, J. T., Davidson, D. B., Imker, H. J., Sadkhin, B., Slater, D. R., and Whalen, K. L. (2015) Enzyme function initiative-enzyme similarity tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim. Biophys. Acta - Proteins Proteomics* *1854*, 1019–1037.
- (17) Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003) Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Res.* *13*, 2498–2504.
- (18) Pellock, S. J., Creekmore, B. C., Walton, W. G., Mehta, N., Biernat, K. A., Cesmat, A. P., Ariyaratna, Y., Dunn, Z. D., Li, B., Jin, J., James, L. I., and Redinbo, M. R. (2018) Gut Microbial  $\beta$ -Glucuronidase Inhibition via Catalytic Cycle Interception. *ACS Cent. Sci.* *4*, 868–879.
- (19) Biernat, K. A., Pellock, S. J., Bhatt, A. P., Bivins, M. M., Walton, W. G., Tran, B. N. T., Wei, L., Snider, M. C., Cesmat, A. P., Tripathy, A., Erie, D. A., and Redinbo, M. R. (2019) Structure, function, and inhibition of drug reactivating human gut microbial  $\beta$ -glucuronidases. *Sci. Rep.* *9*, 825.
- (20) Ndeh, D., Rogowski, A., Cartmell, A., Luis, A. S., Baslé, A., Gray, J., Venditto, I., Briggs, J., Zhang, X., Labourel, A., Terrapon, N., Buffetto, F., Nepogodiev, S., Xiao, Y., Field, R. A., Zhu, Y., O'Neill, M. A., Urbanowicz, B. R., York, W. S., Davies, G. J., Abbott, D. W., Ralet, M. C., Martens, E. C., Henrissat, B., and Gilbert, H. J. (2017) Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature* *544*, 65–70.

#### **FIGURE LEGENDS**

**Figure 1. Structures of *p*-nitrophenyl- $\beta$ -D-glucuronide (*p*NP-GlcA) and *p*-nitrophenyl- $\beta$ -D-galacturonide (*p*NP-GalA) with stereochemical difference highlighted in red.**

**Figure 2. Crystal structure of the GH2  $\beta$ -galacturonidase *EtGalAse*. (A) Quaternary structure of *EtGalAse* shown in cartoon style with active site residues shown as spheres. (B) Active site of *EtGalAse* bound to GalA (blue) with catalytic glutamates shown in deep salmon, NxK motif in cyan, and the unique R337 in yellow. All distances shown are in units of angstroms.**

**Figure 3. Sequence similarity network (SSN) generated from putative GUS sequences from the HMP stool sample metagenomic database reveals distinct clusters for the 13 putative GH2 GalAses and 10 putative hybrid GUS/GalAses. SSN of the 279 putative GUS enzymes identified in the HMP stool sample database with the GH2 GalAse cluster proteins colored blue and hybrid GUS/GalAses in cyan. The SSN was generated with the EFI-EST web tool with an alignment score of  $10^{-150}$ .**

**Figure 4. Structure and function of *FsGUS* reveal Y377 and N-terminus as important structural features for substrate processing. (A) The *FsGUS* hexamer with each chain in a distinct color and active site residues shown as spheres. (B) Zoom-in of active site of *FsGUS* bound to phenyl-thio- $\beta$ -D-glucuronide (PTG, phenyl in dark grey, GlcA in green) with catalytic glutamates colored deep salmon, NxK motif in cyan, and unique YW motif in yellow. All distances shown are in units of angstroms. (C) Active site of *FsGUS***

1  
2  
3 with PTG bound and N-terminus of adjacent protomer shown in green. (D) Catalytic turnovers and  
4 Michaelis constants for WT and  $\Delta 16$  mutant of *FsGUS* with SN-38-G. Values represent an average of three  
5 replicates  $\pm$  SD.  
6

7 **Figure 5. Active site nuances differentiate gut microbial GUSs, hybrid GUS/GalAses, and GalAses.**  
8 Representative active site structures of GUS (PDB: 4JKL [*SaGUS*]), GUS/GalAse (*FsGUS*), and GalAse  
9 (*EtGalAse*) with catalytic glutamates in deep salmon, NxK motif in cyan, key differentiating residues in  
10 yellow, GlcA in green, and GalA in blue.  
11

12 **Table 1. Kinetic parameters of *pNP-GlcA* and *pNP-GalA* processing.**

13  
14 **Figure S1. Sequence similarity network (SSN) of putative GUS enzymes from the HMP stool sample**  
15 **metagenomic database.** Proteins analyzed in this study are labelled and colors designate loop class.  
16

17 **Figure S2. pH profiles of GUS enzymes characterized in this study.** pH profiles of (A) *EtGalAse* with *pNP-*  
18 *GalA*, (B) *EtGalAse* R337A with *pNP-GlcA*, (C) *FcGalAse* WT with *pNP-GalA*, (D) *FsGUS* WT with *pNP-GalA*,  
19 and (E) *FsGUS* WT with *pNP-GlcA*.  
20

21 **Figure S3. Michaelis-Menten plots and progress curves for *EtGalAse* and *FcGalAse*.** (A) Michaelis-Menten  
22 plot for WT *EtGalAse* with *pNP-GalA*, (B) Michaelis-Menten plot for R337A *EtGalAse* mutant with *pNP-*  
23 *GlcA*, (C) Progress curves of WT and R337A mutant with *pNP-GalA*, and (D) Michaelis-Menten plot of WT  
24 *FcGalAse* with *pNP-GalA*.  
25

26 **Figure S4. Sequence similarity networks of HMP ‘GUSome’ at different alignment scores with putative**  
27 **GalAses highlighted in blue and GUS/GalAses in cyan.**  
28

29 **Figure S5. Progress curves and Michaelis-Menten plots for *EtGalAse*, *FcGalAse*, and *FsGUS*.** (A) Progress  
30 curves for *EtGalAse*, *FcGalAse*, and *FsGUS* with *pNP-GlcA*. (B) Progress curves of WT and R335A mutant of  
31 *FcGalAse* with *pNP-GalA*. (C) Michaelis-Menten plots of WT and Y377F mutant of *FsGUS* with *pNP-GlcA*.  
32 (D) Michaelis-Menten plot of *FsGUS* WT and Y377F mutant with *pNP-GalA*.  
33  
34

35 **Figure S6. Structural overlay of *FcGalAse* model and *EtGalAse* crystal structure reveal active site**  
36 **difference that may impact arginine mutant activity.**  
37

38 **Figure S7. Circular dichroism spectra and melting temperatures of WT and mutant proteins of *EtGalAse***  
39 **and *FcGalAse*.**  
40

41 **Figure S8. SEC-MALS trace of purified WT *FsGUS* (left) and SEC traces of WT and  $\Delta 16$  *FsGUS* (right).**

42 **Figure S9. Active site structure of PTG-bound *FsGUS*.** (A) PTG bound to the *FsGUS* active site with 2Fo-Fc  
43 density shown at 1.5  $\sigma$  for the ligand and acid/base Glu-416. (B) Overlay of the apo (cyan) and PTG-bound  
44 (green) *FsGUS* active site reveals the movement of the acid/base Glu-416. Distance shown is in Å.  
45

46 **Figure S10. GalA docked in the active site of *FsGUS*.** Distance shown is in Å.  
47

48 **Figure S11. Progress curves of *FsGUS*, *BuGUS-1*, and mutants with (A) *pNP-GalA* and (B) *pNP-GlcA*.**  
49

50 **Figure S12. SN-38-G processing by *EtGalAse* R337A mutant.** (A) Progress curves for WT and R337A mutant  
51 of *EtGalAse* with SN-38-G. (B) Michaelis-Menten plot of SN-38-G processing by *EtGalAse* R337A.  
52

53 **Figure S13. Tertiary structures and structural alignments of *EtGalAse*, *FsGUS*, and *EcGUS*.**

54 **Figure S14. SEC-MALS trace of purified *EtGalAse*.**  
55  
56  
57  
58  
59  
60

1  
2  
3 **Figure S15. Comparison of *EtGalAse* and *EcGUS* tetramer with each chain colored rainbow style.**  
4

5 **Figure S16. Quaternary structure of *FsGUS*.**

6 **Figure S17. Structural overlay of *BuGUS-1* and *FsGUS* reveals distinct N-terminal structural elements**  
7 **that contribute to the formation of the aglycone binding site.**  
8

9 **Figure S18. Genome neighborhood diagrams of putative GUS/GalAses identified in the HMP stool**  
10 **sample database.** GUS/GalAses are outlined in a black box, other potential coordinating enzymes are  
11 labeled in the legend.  
12

13 **Figure S19. Portion of sequence alignment for BT\_0992, *FsGUS*, *EtGalAse*, and *FcGalAse* with key**  
14 **residues highlighted.**  
15

16 **Figure S20. Genome neighborhood diagrams of putative GalAses.** GalAses are outlined with a box, other  
17 potential coordinating genes are labeled in the legend, and asterisks denote the genome neighborhoods  
18 from which the GalAses were experimentally characterized.  
19

20 **Table S1. Ligation independent cloning (LIC) primer sequences.**  
21

22 **Table S2. Site-directed mutagenesis primer sequences.**  
23

24 **Table S3. Collection and refinement statistics of crystal structures.**  
25

26 **Table S4. Organism and sequence ID for 13 putative GH2 GalAses identified from the HMP 'GUSome'.**

27 **Table S5. Organism and sequence ID for 10 putative GH2 GUS/GalAses identified from the HMP**  
28 **'GUSome'.**  
29

## For Table of Contents Use Only

Selecting a Single Stereocenter: The Molecular Nuances that Differentiate  $\beta$ -Hexuronidases in the Human Gut Microbiome  
Samuel J. Pellock, William G. Walton, Matthew R. Redinbo

