Undergraduate Honors Thesis

Investigating the structural and functional diversity of gut microbial β-glucuronidases

Lianjie Wei

The University of North Carolina at Chapel Hill

Faculty advisor: Dr. Matthew Redinbo, Ph.D.

Thesis Committee Members:

Dr. Bo Li, Ph. D.

Dr. Matthew Redinbo, Ph. D.

Dr. Dorothy Erie, Ph. D.

Department of Chemistry, the University of North Carolina at Chapel Hill

NOTE TO READERS

Because of the COVID-19 pandemic at the time this thesis was written, I could not perform experiments or retrieve data from the laboratory during this time. As a result, some of the figures presented in this thesis are still preliminary and not of publication quality. I apologize for any of the inconveniences caused by using these figures in my thesis.

ACKNOWLEDGEMENT

I would like to thank my thesis committee members, Dr. Bo Li, Dr. Dorothy Erie, and Dr. Matthew Redinbo, for being an integral part of the last leg of my journey at UNC. Especially, I would like to thank Dr. Redinbo for taking me on as a first-year undergraduate student and molding me to become a better scientist. During my three years in the Redinbo laboratory, Matt was supportive of all of my crazy ideas. He listened to all of them and would always give me his insights on the matters. He also gave me advice on how to become an aspiring scientist, and more importantly, a better person. I could not have learned that much without him.

I would also like to thank my mentors during my time here. They are: Dr. Kristen Biernat, Dr. Sam Pellock, Josh Simpson, Parth Jariwala, Marissa Bivins, and William "Bill" Walton. Kristen, my first mentor in the lab, taught me how to do experiments, starting from pipetting. Often, they pointed out my mistakes and gave me suggestions on how to do science the "right" way. It is from them that I became a better scientist. I am also grateful that Josh, Marissa, and Parth helped me through the graduate school application process and guided me to the next step of my graduate career.

I would like to point out the people who helped making the thesis possible. There people include: Dr. Kristen Biernat, Bill Walton, Dr. Ashutosh Tripathy, Marissa Bivins, and Josh Simpson. They taught me the skills along the way by sacrificing their time to be involved in my work. I would like to thank them for their contribution.

Lastly, I would like to thank my family in China for supporting me for my studies and my high school host family in Georgia for giving me the motivation to pursue science. They are the ones who cheered me on through the highs and lows, and I am forever thankful for that.

ABSTRACT

Gut microbial β-glucuronidases are one of the key players in reactivating drug metabolites in the gut lumen. This reactivation can sometimes lead to dose-limiting side effects, such as severe diarrhea and intestinal bleeding[1]. Thus, understanding the structures and functions of these β-glucuronidases is critical for developing adjuvant therapies to alleviate the gut microbial β-glucuronidase-induced gastrointestinal toxicities[2]. However, the structural and functional diversities of gut microbial β-glucuronidase enzyme class are still poorly characterized. In this thesis, we describe three independent studies to understand these diversities. First, we studied the biochemical aspects of *Streptococcus Agalactiae* β-glucuronidase and *Bacteroides dorei* β-glucuronidase and of structural aspects *Bacteroides dorei* β-glucuronidase inhibition. Next, we used loop swap mutants in an attempt to understand the function diversity caused by the extended active site loops. Last, we performed biochemical assays on a previously poorly characterized class of β-glucuronidase. Our works are a part of a collective effort to explore the gut microbial β-glucuronidase as an enzyme family. In the future, we can use our knowledge gained from these studies to further our knowledge on the interactions between the gut microbiota and drug metabolites and select β-glucuronidases as targets for inhibitor design.

INTRODUCTION

The human body is an intricate machinery with trillions of cells working interdependently to perform various biological functions. However, it is not just the human cells that reside on the human body: the human epithelial surfaces are home to the human microbiota, a collection of commensal microorganisms of which the total genome outnumbering the human genome 3:1[3].

There are multiple microbiota communities in a human body, including skin microbiota, oral microbiota, and gut microbiota. The human gut microbiota, in particular, plays a major role in modulating human health by carrying out numerous biological processes critical for the human health[4] (Figure 1).

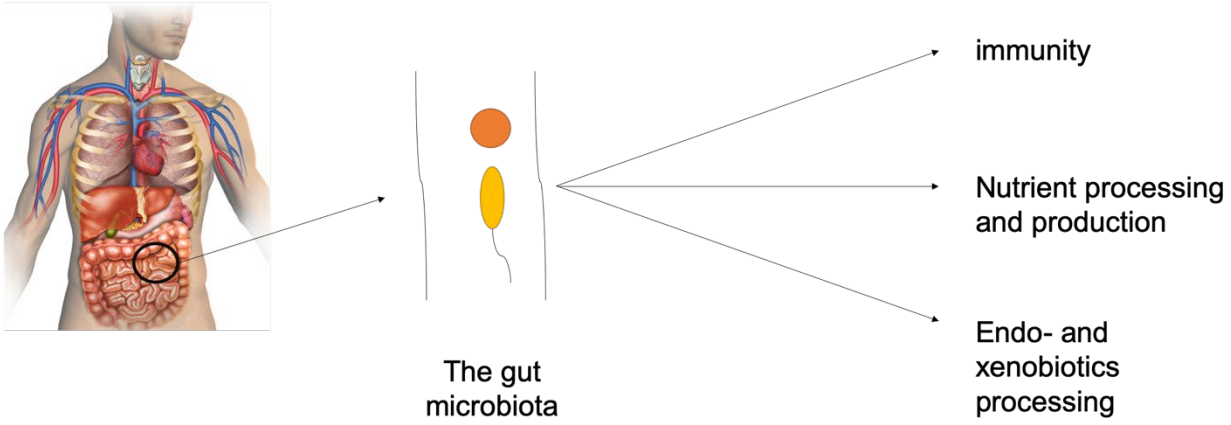


Figure 1: the major biological functions of the human gut microbiota

One of the major ways that the gut microbiota impact human health is through becoming an integral part of host immune system. This is achieved in two folds. First, the normal gut microbiota plays a pivotal role in mucosal immunity by stimulating host production of immunoglobulin A-secreting plasma cells in lamina propria, intraepithelial lymphocytes, and microbicidal protein Angiogenin-4[5]. Second, the gut microbiota can modulate host immunity. This is implicated in type I hypersensitivity, Alzheimer's Disease, type I diabetes, and inflammatory bowel disease[3, 6].

The second major way that the gut microbiota impact human health is through key metabolite processing. The gut microbiota is responsible for the breakdown of complex carbohydrates to provide energy for the host, and this process accounts for at least 10% of the

total available energy to the human body[7]. Other than carbohydrate fermentation, the gut microbiota synthesizes vitamins such as biotin, cobalamin, and niacin[8]. *E. coli* and other gut bacteria also synthesize a siderophore called enterobactin. Enterobactin was previously thought to be detrimental to human health as it is a genotoxin and its siderophore activity to only benefit the bacteria, not the host. However, recent studies show that the host can also use enterobactin for iron(III) intake[9].

The third major way that the gut microbiota influences human health is through endo- and xenobiotic processing. The gut microbiota can process endobiotics and xenobiotics, converting them into compounds that can either positively or negatively impact the human health. For example, the gut microbiota can produce estrogen by cleaving the glyosidic bond off estrogen glucuronides, an estrogen metabolite produced by the liver for secretion[10, 11]. Our group is interested in the gut microbial enzymes responsible for endo- and xenobiotics processing.

We were first interested in irinotecan metabolite processing Irinotecan is the first-line chemotherapy drug for patients with colorectal cancer (Figure 2). After irinotecan is administered, it is metabolized to the active metabolite SN-38 in the liver. Also in the liver, a glucuronic acid moiety is added to SN-38 at β-conformation to form SN-38 glucuronide (SN-38G) by UDP-glucuronyltransferase to be marked for excretion. As a result, SN-38G reaches the gastrointestinal tract. Inside, the gut microbiota recognizes the glucuronic acid moiety and uses gut microbial β-glucuronidases (GUSs) to cleave the glyosidic bond in order to utilize the free glucuronic acid as a carbon source. The free SN-38 is then released to the gut lumen and causes gastrointestinal tract toxicity. This leads to dose-limiting side effects ranging from intestinal bleeding to massive diarrhea. Thus, we became interested in gut microbial GUSs, and we found

that many other drug metabolites undergo similar processes in the gut lumen to cause gastrointestinal toxicity. For us, understanding the functions of gut microbial GUSs is critical to develop methods alleviate drug gastrointestinal toxicity[6].

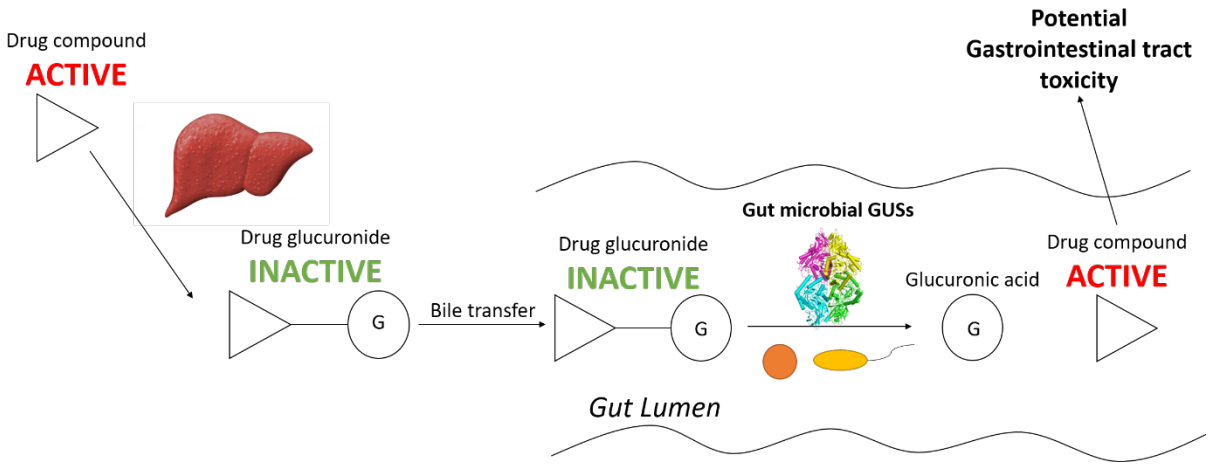


Figure 2: drug glucuronidation in liver and drug glucuronide activation by microbial GUSs

We first studied *E. coli* GUS (*Ec*GUS) to characterize SN-38G processing and developed small molecule inhibitors to inhibit SN-38G by *Ec*GUS. So far, we have developed two small molecule inhibitor series, Inhibitor 1 and UNC10201652. These inhibitors are specific to bacterial GUSs over human GUSs, as deficiencies in human GUSs cause the Sly Syndrome, a lethal lysosomal storage disease. Also, these inhibitors are not lethal to bacterial cells, so the GI tract microbiota composition is not altered by these inhibitors alone. These inhibitors exhibit high nM to low μM IC50 against *Ec*GUS, and they can alleviate the GI side effects caused by irinotecan[2, 12].

We then mapped the primary sequence diversity of the gut microbial GUSs in order to better understand the causes of drug-induced gut toxicity. As the result, we learned that there are at least 279 unique GUS sequences in the GI tract microbiome[13]. We also learned that around 50% of the unique GUS sequences have an extended loop region around the active site. Based

on the primary sequence position of the loop and on the length of the loop, we categorized the GUSs based on 6 different subclasses: Loop 1 (L1), Loop 2 (L2), mini Loop 1 (mL1), mini Loop 2 (mL2), mini Loop 1 and 2 (mL12), and no Loop (nL)[13]. We initially discovered that some L1 GUSs and process SN-38 glucuronide (SN-38G). The one mL1 GUS we then characterized, *Bacteroides fragilis* GUS also processes SN-38G, but with catalytic efficiency ~50% of active L1 GUSs[14]. Our small molecule inhibitors, however, are only effective against selected L1 GUSs. Thus, our long-term goal is to develop inhibitors for non-L1 GUSs that are also important in drug glucuronide reactivation. To do this, we first need to fully understand the structural and functional diversity of the gut microbial GUS enzyme family.

Here, we describe our work on this topic. First, we characterized *Bacteroides dorei* GUS biochemically, and we tested inhibition of *Streptococcus Agalactiae* GUS and *Bacteroides dorei* GUS by UNC10201652. Surprisingly, we found that *Bacteroides dorei* GUS is inhibited by UN10201652, so we attempted to identify the structural basis of this phenomenon. This work has been published in 2019[14]. Then, to understand the functional diversity of GUSs, we developed loop swap mutants but found them to be catalytically inactive. Thus, we deployed biophysical analyses to understand the functions of these active site loops. Lastly, we performed kinetic assays on mL1 GUSs to investigate the functional diversity of this class of GUSs. Through these works, we showed that these GUSs exhibit drastically different kinetics and inhibitory parameters and provided some structural insights on these differences. By doing so, we furthered our understandings on the functional and structural diversity of gut microbial GUSs.

MATERIALS AND METHODS

1. Primary sequence analysis

   The primary sequences of GUSs of interests were obtained from UniProt (EMBL-EBI) or from our internal primary sequence database. Then, primary sequence alignments were performed by ClustalOmega (EMBL-EBI). For these alignments, information such as percent identity matrix were obtained from the built-in functions of ClustalOmega.

2. Protein expression and purification

   The method was adopted from Kiernat et al 2019 and Pellock et al 2019[14, 15]. In short, pLIC-plasmids containing his-tagged GUSs of interests were first transformed into BL21-Gold *E. coli* cells in Miller lysogeny broth (LB) for overnight growth at 37°C. Afterwards, the media containing transformed *E. coli* cells were transferred to 100mL LB for a second round of overnight growth at 37°C. Then, the cells were transferred to 1.5L LB for incubation at 37°C. When the optical density of the cell media reached 0.6, 200μL 1M isopropyl β-D-thiogalactopyranoside was added to the LB to induce GUS protein expression, and the cells were incubated at 18°C overnight. The cells were pelleted by centrifugation at 4500×g for 20 minutes and stored at -80°C before purification.

   For protein purification, the cell pellets containing GUS of interest was first suspended in the lysis buffer (20mM potassium phosphate, 50mM imidazole, 500mM NaCl with one Roche EDTA-free protease tablet, 150μL DNAse1, and pinch lysozyme; pH7.4 for all except for UNC361-9 and *Eubacterium spp*. GUS (pH8.0)). The cells were then lysed via sonication, and the soluble content was subsequently separated from insoluble content via

centrifugation at 17000×g at 4°C for 45 minutes. The supernatant was then filtered by syringe filtration, and the filtrate was then loaded onto nickel column with the nickel column loading buffer (20mM potassium phosphate, 50mM imidazole, 500mM NaCl; pH7.4 for all except for UNC361-9 and *Eubacterium spp*. GUS (pH8.0)) (General Electric ATKA FPLC system). Afterwards, the hexahistidine-tagged sample was then eluted from the nickel column with the wash buffer (20mM potassium phosphate, 500mM imidazole, 500mM NaCl; same pH as loading buffer for each GUS). The largest sample peak from the affinity chromatography was then loaded onto Superdax 200 size exclusion chromatography column (General Electric ATKA FPLC system) and was eluted by super200 column buffer (20mM HEPES, 50mM NaCl; pH same as previous buffers for each GUS). The largest sample peak was taken as the GUS of interest. If needed, the eluted GUS was concentrated in a concentration cassette tube by rounds of centrifugation at 3000×g at 4°C. The GUS was then aliquoted and flashed frozen in liquid nitrogen. The aliquots were then stored in -80°C freezer.

3. *in vitro* Kinetic assays

In this thesis, p-nitrophenyl β-glucuronide (*p*NPG) and 4-methylumbelliferyl β-glucuronide (4MUG) were used as substrates for *in vitro* GUS kinetic assays. The pNPG assay procedure was adapted from Biernat et al. 2019. In short, pNPG was purchased as a solid from Sigma Aldrich and was resuspended in water to form a 100mM solution. The assay was run in a 96-well, clear bottom assay plate (Costar) at 37°C. The final volume for each well in a reaction was 50μL, consisting of 10μL assay buffer (50mM sodium

acetate(pH4-6.0)/50mM HEPES (pH6.5-7.4), 50mM NaCl, at GUS of interest optimal pH), 10μL GUS (various concentrations), and 30μL *p*NPG (various concentrations). Product formation was measured at 410nm with PHERAstar Plus Microplate reader (BMG Labtech). To determine the optimal pH of GUSs of interests, assays similar to aforementioned were performed, with the except of using 800μM *p*NPG at pH4.0-7.4. For assays run at pH6.0 or lower, reaction was quenched by 100μM 0.2M sodium bicarbonate at various time points and final product formation was measured by the plate reader. For pH>6.5, quenching was not needed, and the product formation was monitored by the plate reader over time. The Michaelis-Menten kinetics module in SigmaPlot was used to calculate K_m, k_{cat}, and catalytic efficiency[15].

4MUG assay procedure was adopted from Pellock et al 2019. In short, reactions were performed a 96-well, clear bottom assay plate (Costar) at 37°C. 4MUG was purchased from Sigma Aldrich as a solid and was resuspended in 100% DMSO into 100mM solution. The final volume for each well in a reaction 50μL, consisting of 35μL water, 5μL assay buffer (25mM sodium acetate(pH4-6.0)/50mM HEPES (pH6.5-7.4), 50mM NaCl, at GUS of interest optimal pH), 5μL GUS (50nM, for 5nM final concentration unless otherwise noted), and 5μL 4MUG (varies concentrations). Reactions were initiated by addition of 4MUG and were continuously monitored with excitation at 350nm and emission at 450nm in PHERAstar Plus Microplate reader (BMG Labtech). Initial velocities from resultant data were fit by linear regression with a custom MATLAB program. In instances where K_m, k_{cat}, and catalytic efficiency were all reported, SigmaPlot 13.0 was used for

perimeter calculations. In instances where only catalytic efficiency was reported, apparent k_{cat} was first calculated, and catalytic efficiency was the slope of apparent k_{cat} as a function of substrate concentration[15].

4. *in vitro* inhibition assay

   in vitro inhibition assay was as described in Biernat et. al 2019. In short, reactions consisting of 5μL GUS (15nM final for *Sa*GUS, 150nM final for *Bd*GUS), 5μL inhibitor (various concentrations), 30μL *p*NPG, and 10μL assay buffer (25mM NaCl, 25mM HEPES, pH6.5 or 7.4) were added into a Costar 96-well clear well plate. The reactions were initiated by *p*NPG addition, and the product formation was monitored by BMG Labtech PHERAstar Plus Microplate reader for 60 minutes. Due to the slow-binding nature of UNC10201652, IC_{50} was instead used as a measure for inhibitor efficacy. This was analyzed on SigmaPlot 13.0 by plotting the percent concentration as a function of log of inhibitor concentration and fitting with four-parameter logistic functions. The percent inhibition was calculated as:

$$\%\ \text{inhibition} = \left(1 - \frac{\mathrm{A_{exp}} - \mathrm{A_{bg}}}{\mathrm{A_{max}} - \mathrm{A_{bg}}}\right) \times 100\%$$

   Where $\mathrm{A_{exp}}$ is the end point absorbance at a particular inhibitor concentration, $\mathrm{A_{max}}$ is the absorbance of the uninhibited reaction, and $\mathrm{A_{bg}}$ is the background absorbance of the reaction[15].

5. Circular dichroism and melting temperature experiment

The secondary structures and thermostabilities of WT *Ec*GUS and loop swap mutants were compared by using circular dichroism and melting temperature experiment. 2.5μM GUS in CD buffer (10mM potassium phosphate, 100mM potassium fluoride; pH7.4) was loaded into 1-mm pathlength cuvette. Chirascan-plus instrument (Applied Photophysis Limited) was used to measure the circular dichroism of the GUSs. The melting profile of the samples were monitored at 211nm for all samples.

6. Size-exclusion chromatography-multi-angle light scattering

   Size-exclusion chromatography-multiangle light scattering was used to determine the quaternary structure of WT *Ec*GUS and loop swap mutants. The GUSs were analyzed on a Superdex 200 size exclusion column connected to Agilent FPLC system with Wyatt DAWN HELEOS II multi-angle light scattering instrument and Trax refractometer. 50μL of each GUSs (various concentrations) were injected, and SEC-MALS buffer (50mM HEPES, 150mM NaCl, pH7.4) was used to elute the proteins. Light scattering and refractive index data were collected and analyzed using Wyatt ASTRA v. 6.1 software.

7. Structural analysis

   All structural analyses were done on PyMOL. For docking experimental, molecules were manually docked into the active site of GUS. For hydrophobicity analyses, open-resource code provided the Protein Research Institute at Osaka University (Osaka, Japan) was used.

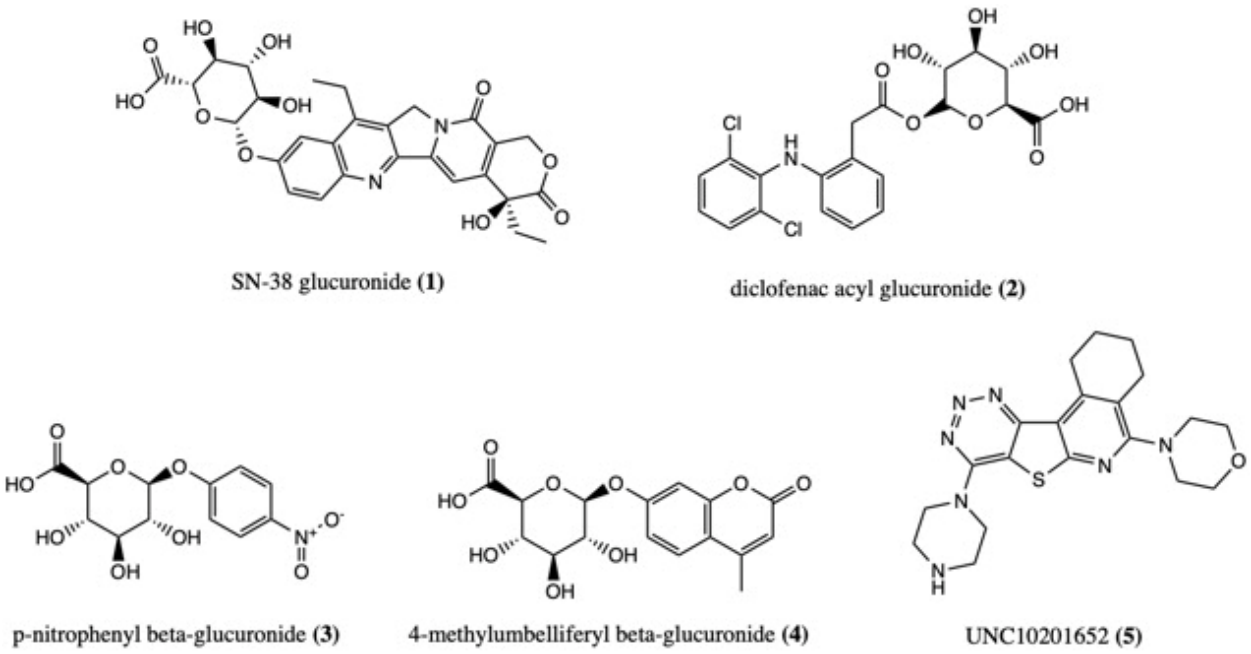


Figure 3: the compounds of interests for this thesis. SN-38G **(1)** and diclofenac acyl glucuronide (DCF-G) **(2)** are the two biologically relevant compounds we are interested in, whereas *p*NPG **(3)** and 4MUG **(4)** were used as model substrates for **(1)** and **(2)**. UNC10201652 (4-(8-(piperazin-1-yl)-1,2,3,4-tetrahydro-[1,2,3]triazino[4',5':4,5]thieno[2,3-c]isoquinolin-5-yl)morpholine; **(5)**) is the gut microbial GUS inhibitor tested in this thesis.

RESULTS

NL GUS from Bacteroides dorei (BdGUS) is a poor hydrophobic small substrate processor

Our landmark bioinformatics article published in 2017 indicated that *Bd*GUS does not have an extended active site loop capping the active site pocket and predicted that *Bd*GUS could not efficiently process SN-38G, but we had not yet obtained the kinetics parameters of *Bd*GUS to test our prediction[13]. The chemical space of DCF-G and SN-38G is drastically different (Figure 3 **(1)** and **(2)**). For instance, in SN-38, five rings are linked and forms two distinct conjugated rings

systems, whereas DCF is essentially a secondary amine group with two aryl rings as substituents. However, these two compounds are both largely hydrophobic with the exception of glucuronic acid moiety. This enabled us to use a small, hydrophobic model substrate to conveniently test *Bd*GUS kinetics when processing a hydrophobic compound.

To do this, we used *p*-nitrophenyl β-glucuronide (*p*NPG, Figure 3 **(3)**) because the product of GUS *p*NPG processing, *p*-ntirophenyl (pNP^-) absorbs light at 410nm if pH≥6.5. Thus, we could use UV/Vis spectroscopy to detect pNP^- formation and calculate the kinetics of GUS of interest. Although pNP^- does not absorb light when pH<6.5, we could relatively easily compensate for this by quenching the GUS processing reaction with 2M sodium bicarbonate at certain time points and obtain the concentrations of pNP^- in quenched reaction mixture using the same spectrochemical method. In the case of *Bd*GUS, we tested the kinetics at its optimal pH (pH6.0).

Our experimentally-obtained kinetic parameters of *Bd*GUS confirmed our predictions. The catalytic efficiency of BdGUS ($5.2\times10^3\ s^{-1}M^{-1}$) is ~176-fold slower than that of *Ec*GUS ($9.2\times10^5\ s^{-1}M^{-1}$) (Figure 4A and B). Thus, *Bd*GUS is not a good processor for *p*NPG and is also likely not a good processor for SN-38G and DCF-G. Nevertheless, we gained more insights into the kinetics of nL GUSs and functional diversity of the gut microbial GUS enzyme family.

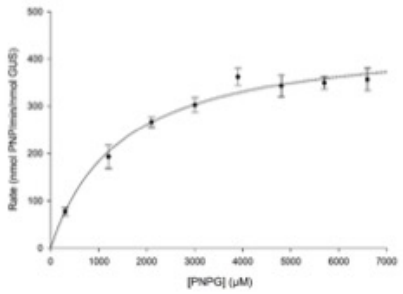
UNC10201652 can inhibit some GUSs, including BdGUS

To counter the GI toxicity caused by DCF, we decided to test if our previously developed inhibitor, UNC10201652 (Figure 3 **(5)**), inhibit the novel GUSs discovered by the 2017 bioinformatics study. *Streptococcus Agalactiae* GUS (*Sa*GUS), was categorized as a L1 GUS. As we found that UNC10201652 inhibited *Ec*GUS at high nanomolar level, we predicted that

UNC10201652 could also inhibit *Sa*GUS at similar inhibitor concentration[13]. In contrast, *Bd*GUS, just as human GUS (*Hs*GUS), does not have an extended active site loop capping the pocket. Then, our proposed UNC10201652 inhibition mechanism indicated that the extended active site loop needed to make contact with UNC10201652 to promote inhibitor binding to GUS active site pocket. Thus, we predicted that UNC10201652 could not inhibit *Bd*GUS[12, 13, 14].

Our data confirmed our prediction on *Sa*GUS inhibition by UNC10201652. The IC_{50} value of UNC10201652 against *Sa*GUS was 450±30nM at pH6.5 and 133±3nM at pH7.5. However, to our surprise, *Bd*GUS was also inhibited by UNC10201652 (IC_{50}>250000nM at pH7.5), whereas *Hs*GUS was not (Figure 4C)[14]. Although this finding is not clinically important, it remained an important scientific question to understand the basis of *Bd*GUS inhibition by UNC10201652.

A



B

GUS	pH tested	k_{cat} (s^{-1})	K_M ($\times10^{-3}$ M)	k_{cat}/K_M ($\times10^3 s^{-1}M^{-1}$)
Ec	7.4	120±10	0.13±0.01	920±10
Bd	6.0	7.5±0.4	1.4±0.3	5.20±0.5

C

GUS	IC50 (UNC10201652, pH6.5; nM)	IC50 (UNC10201652, pH7.5; nM)
Ec	310±20	100±8
Sa	450±30	133±3
Bd	N. I.	>250000

Figure 4: the kinetics and inhibition of selected GUSs. **A**: the Michaelis-Menten kinetics of *Bd*GUS. **B**: the kinetics parameter comparison between *Bd*GUS and *Ec*GUS. The catalytic efficiency of BdGUS ($5.2\times10^3\ s^{-1}M^{-1}$) is ~176-fold slower than that of *Ec*GUS ($9.2\times10^5\ s^{-1}M^{-1}$). **C**: the IC50 values of EcGUS, SaGUS, and BdGUS against UNC10201652 (N. I.= not inhibited)[15].

BdGUS inhibition by UNC10201652 remains inconclusive

In order to understand the structural basis of *Bd*GUS inhibition by UNC10201652, we first used a residue-based approach. By docking UNC10201652 in *Bd*GUS active site pocket *in silico*, we found that F456 and W458, two unique residues to *Bd*GUS over *Hs*GUS, may form π-stacking interactions with pyridine and triazine rings of UNC10201652. Moreover, a carbohydrate binding domain (CBM) loop unique to *Bd*GUS over *Hs*GUS rests above active site pocket opening, resembling the extended active site loop in L1 GUSs. (Figure 6A). We hypothesized that these unique features of *Bd*GUS combined contribute to inhibition of *Bd*GUS by UNC10201652.

To test our hypothesis, we made three mutants: his-*Bd*GUS F456A, his-*Bd*GUS W458A, and his-*Bd*GUS ΔCBM Loop and expressed the mutants in *E. coli* BL21G cell line. We then purified these mutants via affinity and size-exclusion chromatography. To examine the stability of these mutants, we ran SDS-PAGE for both WT and mutant *Bd*GUS. *Bd*GUS W358A and *Bd*GUS ΔCBM Loop are of the same molecular weight as WT *Bd*GUS, whereas *Bd*GUS F456A showed signs of degradation (Figure 5). We then tested the kinetics and inhibitions of intact *Bd*GUS mutants. However, we did not see the diminished inhibition as we expected (data not shown).

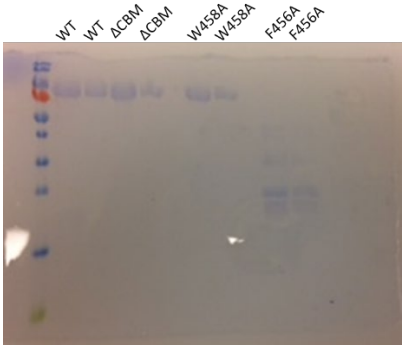


Figure 5: SDS-PAGE gel of purified *Bd*GUS mutants.

We then decided to revisit the *Bd*GUS structure (Figure 6A). *Bd*GUS forms a dimer, with the active site pockets facing outward. This configuration effectively exposes the pocket to solvent, including F456 and W458. As a result, F456 and W458 may be stabilizing each other, and this interaction may also be the stabilizing factor for the α-helix like motif of which F456 is the N-terminal end, because the other regions of this motif are more hydrophilic and subject to denaturation by dissolution without F456-W458 interaction. F456A seems to have a more profound effect of W458A on *Bd*GUS, most likely because F456 also forms an edge-to-face π-stacking interaction with W423. The F456-W423 interaction may act as a staple for the F456-W458 initiated α-helix like motif to the rest of *Bd*GUS. Thus, losing F456-W423 interaction could cause global protein misfolding, making *Bd*GUS F456A prone to degradation, as inferred in the FPLC chromatogram for *Bd*GUS F456A and SDS-PAGE of *Bd*GUS F456A. The W423-F456-W458 interaction in BdGUS may also explain the cause for enhanced IC_{50} observed in *Bd*GUS W558A mutant. Losing F456-W458 interaction is less likely to cause a global misfolding event because F456 and W458 are found on the same chain. However, loss of this π-stacking interaction increases the degree of freedom of α-helix like motif random walk. Thus, F456 and potentially W423 can be better positioned to interact with UNC10201652, thereby increasing its binding efficiency and in turn lowering IC_{50}. Thus, we proposed that the π-stacking interactions among W423-F456-W458 may be instrumental for UNC10201652 binding to *Bd*GUS, but F456-W458 interaction found in WT *Bd*GUS may hinder this process. However, our experimental results were inconclusive to test our new proposition.

Because UNC10201652 is largely hydrophobic, we sought to explore the presence of a hydrophobic binding site on *Bd*GUS for UNC10201652. To test this, we used an open-source

python code to visualize the hydrophobicity of *Bd*GUS residues. However, we found no such pockets. An encrypted pocket is unlikely to be found on *Bd*GUS because GUSs do not go through conformational changes and are known to be structurally stiff and thermostable (Figure 6B). Thus, the mechanism of *Bd*GUS inhibition by UNC10201652 remains inconclusive.

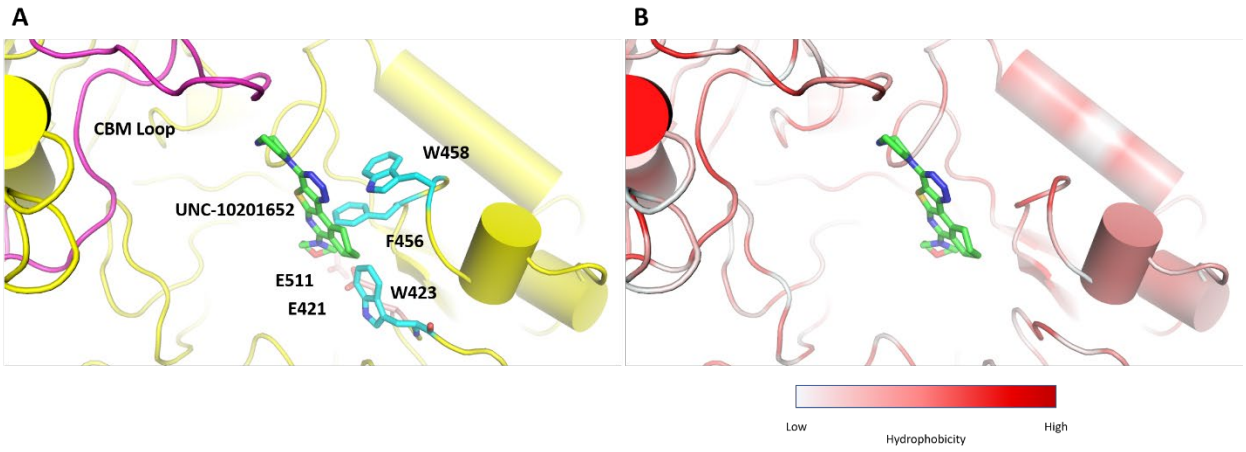


Figure 6: the structural basis of *Bd*GUS inhibition by UNC10201652 is inconclusive. **A**: The *Bd*GUS (yellow) active site. The CBM loop (teal) W423, F456, and W458 (purple) are all unique to the *Bd*GUS and may contribute to the inhibition of the enzyme by UNC10201652 (green). W423-F456-W458 form hydrophobic interactions to stabilize the α-helix like motif. **B**: hydrophobicity map of *Bd*GUS, with UNC10201652 docked. There is no observed hydrophobic caves inside *Bd*GUS active site.

Loop swap GUS mutants are catalytically inactive

GUS is a diverse structurally and functionally diverse enzymes. Although the GUSs are all capable at cleaving glucuronic acid moiety, they do so at different rates, and they have different substrate preferences. Also, it became evident that UNC10201652 could only inhibit only a portion of L1 GUSs. We categorized the GUSs into different loop classes, but this categorization

does not reflect the substrate preferences and differential kinetics of these GUSs[13, 14, 15]. Even within the same group of GUSs, the kinetics of these GUSs are drastically different when the same substrate is tested. One of the major differences among GUSs is their extended active site loops (Figure 7A-B). Because the catalytic residues of GUSs are conserved, we hypothesized that the identity of the extended active site loops affects GUS catalytic activities and inhibition by UNC10201652.

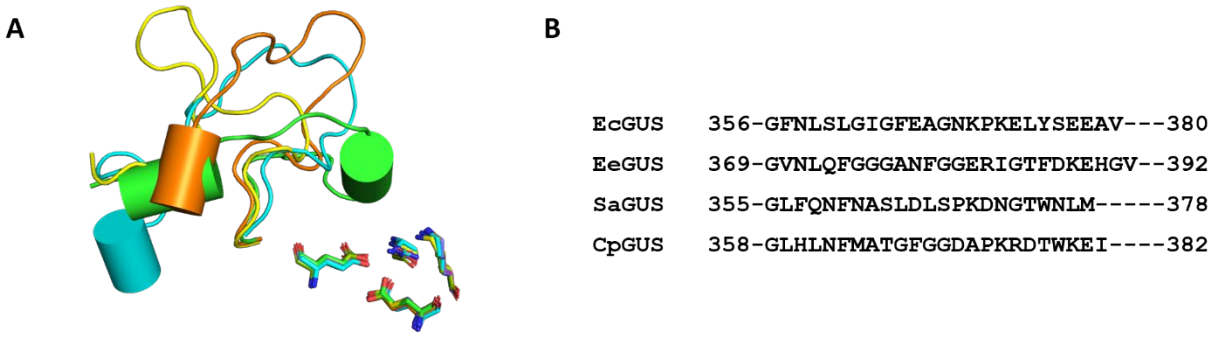


Figure 7: L1 GUSs adopt highly variable active site loop structures. **A**: the overlay of the extended active site loops on EcGUS. **B**: the active site loop sequences of L1 GUSs used to make loop-swap mutants.

To test this hypothesis, we designed three "loop swap" mutants by swapping the loops of *Sa*GUS, *Eubacterium eligens* GUS (*Ee*GUS), and *Clostridium perfringens* GUS (*Cp*GUS) each with *Ec*GUS loop. We reasoned that if the extended active site loop is responsible for modulating GUS kinetics, these loop swap GUSs should adopt the kinetics of loop donor GUSs. We used pNPG as our model substrate. To our surprise, these loop swap mutants have no significant catalytic activities (Figure 7A-C). We then used SDS-PAGE to check if the loop swap mutants were prone to degradation. Although we found that the mutants do not degrade in denaturing environments,

we noticed that polymeric populations were prevalent in WT GUSs even with SDS denaturation, whereas all of the populations for loop swap mutants are monomers (Figure 8A). We reasoned that these loop swap mutants may have the same fold as WT *Ec*GUS but are less thermostable. To test this hypothesis, we used circular dichroism to perform wavelength scans and melting temperature experiments for WT *Ec*GUS and all loop swap mutants (Figure 8B-C). We found that while WT *Ec*GUS and loop swap mutants have the same secondary structures, loop swap mutants melt at ~60˚C, whereas the melting point for EcGUS exceeds 94˚C (Figure 8C). We also used size exclusion chromatography-multiangle light scattering to determine the oligomerization state of WT *Ec*GUS and all loop swap mutants and found that these enzymes are all tetramers (Figure 8C-D). Our biophysical experiences confirmed our hypothesis that the loop swap mutants are less thermostable than WT *Ec*GUS. However, the reason behind diminished activities of these loop swap GUSs remained to be unknown.

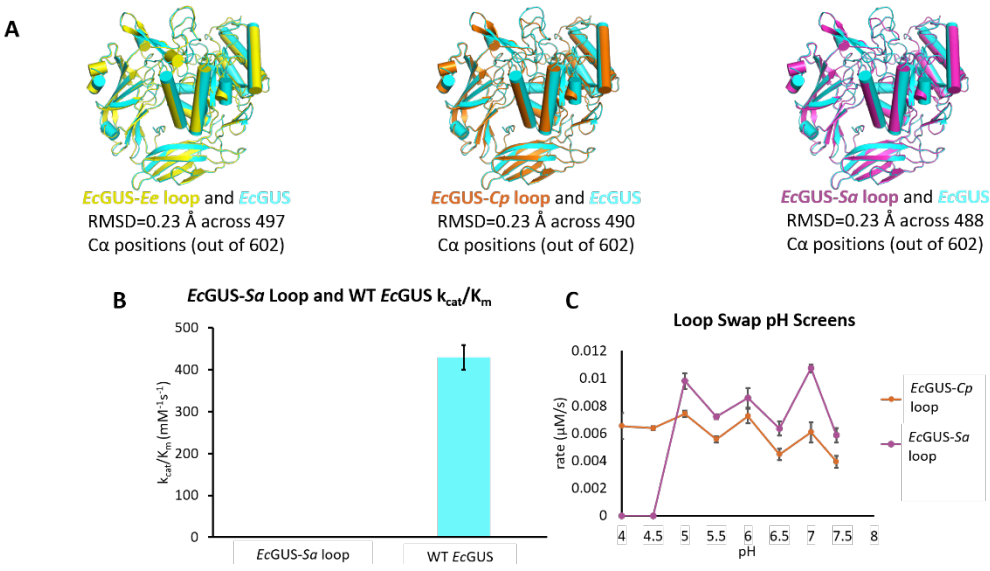
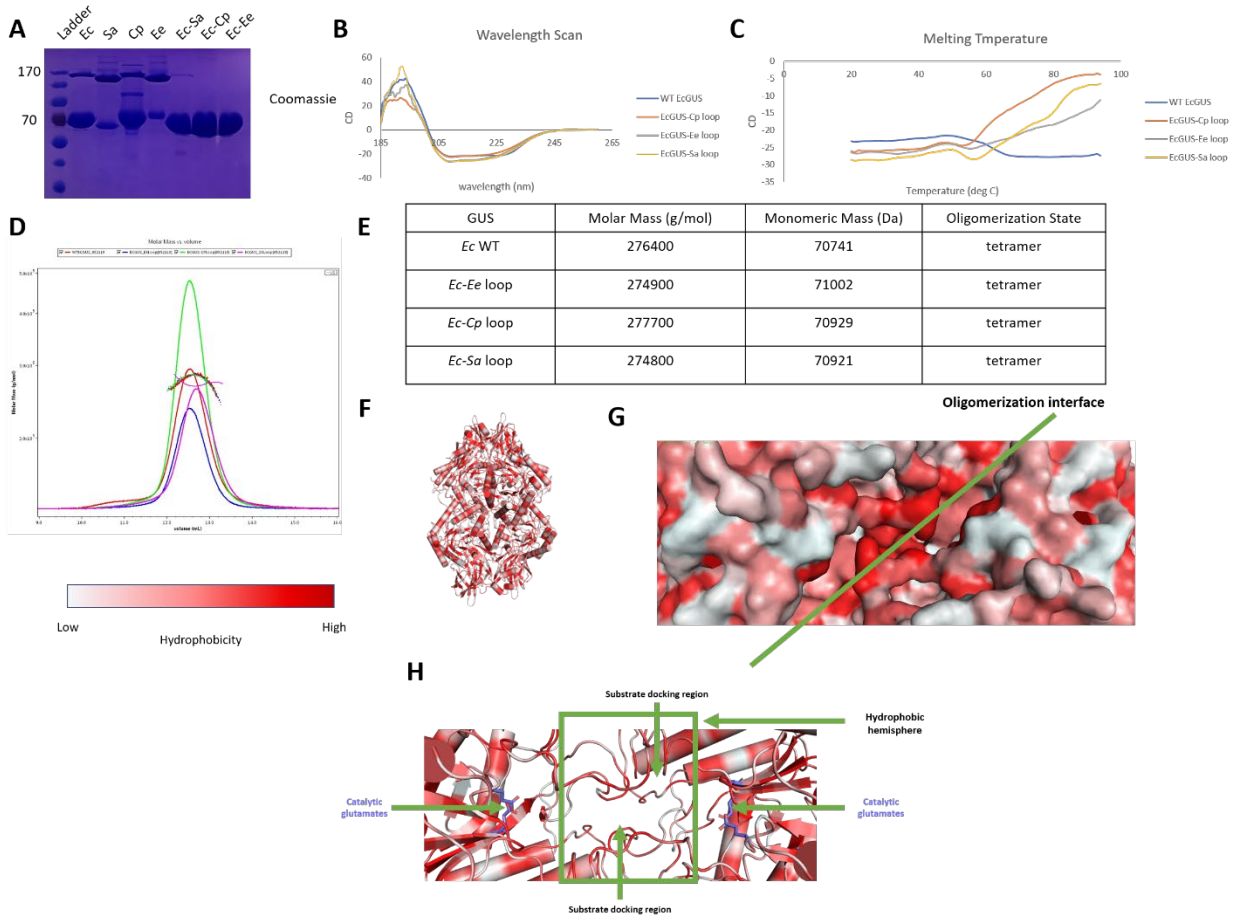


Figure 8: loop-swap mutants are predicted to have the conserved fold but are shown to be inactive. **A** alignments of Phyre2 models of *Ec*GUS-*Ee*GUS loop hybrid (*Ec*GUS-*Ee* loop; left,

yellow), *Ec*GUS-*Cp*GUS loop hybrid (*Ec*GUS-*Cp* loop; middle, orange), and *Ec*GUS-*Sa*GUS loop hybrid (*Ec*GUS-*Sa* loop; right, purple) with WT *Ec*GUS (cyan) show that the predicted loop swap mutants share the same tertiary structures with WT *Ec*GUS; **B** pH screen with 100nM loop swap mutants (*Ec*GUS-*Ee* loop is inactive) and 200μM 4-methylubelliferone β-glucuronide showed nM/s-scale rate; **C** k_{cat}/K_m of *Ec*GUS-*Sa* loop is ~1000-fold slower than WT *Ec*GUS.

To answer this question, we used the open-access hydrophobicity visualization code to analysis *Ec*GUS structure (Figure 8F-H). *Ec*GUS forms a tetramer of which the center is the symmetry point of the structure (Figure 8F). The active site of EcGUS are outward-facing, with the active site within a protomer facing toward each other. This causes the active sites in a protomer to be in close proximity to each other (Figure 8G). The active site pockets of *Ec*GUS is buried beneath the oligomerization interface. These features combined confine the active site of *Ec*GUS (Figure 8G). Two largely hydrophobic loops from each monomer, one of which the extended active site loop used for loop class categorization, support the pocket by interacting with the base of *Ec*GUS and provide the space needed for substrate entry and product exit (Figure 8G-H). Because of these loops, the top hemisphere of *Ec*GUS active site is largely hydrophobic fairly stable (Figure 8H). Thus, interrupting the hydrophobic loop-base interactions by loop swapping may destroy the stability of the hydrophobic hemisphere of *Ec*GUS. The decreased active site pocket stability may lead to two possible outcomes. First, the instability can decrease the volume of the active site pocket and make substrate entry increasingly difficult. If this is true, the loop swap mutants cannot process any substrates. The second possible outcome is that the decreased stability causes the active site pocket to be more solvent-exposed. This means that

while the loop swap mutants are poor processor for *p*NPG and other small hydrophobic substrates, these mutants may be better at processing hydrophilic substrates. While we have a deeper insight on the structural basis of the diminished kinetics of *p*NPG processing by loop swap mutants, more substrates need to be tested to understand the effect of loop swapping on *Ec*GUS active site pocket.



GUS	Molar Mass (g/mol)	Monomeric Mass (Da)	Oligomerization State
Ec WT	276400	70741	tetramer
Ec-Ee loop	274900	71002	tetramer
Ec-Cp loop	277700	70929	tetramer
Ec-Sa loop	274800	70921	tetramer

Figure 8: loop-swap mutants have similar secondary and quaternary structures as WT *Ec*GUS but are less thermostable than WT *Ec*GUS. **A**: Nonreducing SDS-PAGE gel of loop swap mutants. Lanes 1-4: WT EcGUS, SaGUS, CpGUS, and EeGUS. Lanes 5-7: EcGUS-Sa loop, EcGUS-Cp loop, and EcGUS-Ee loop. Protein concentration is normalized at 36uM with pH 6.5 GUS buffer (25mM HEPES and 25mM NaOAC) before the addition of 3X Laemmli sample buffer and loading. The

monomeric molecular weights of all GUSs are 70kDa. A 170kDa band exists for all WT bands, but not for loop swap GUSs. **B**: circular dichroism of WT EcGUS and loop-swap mutants show no differences in secondary structure. **C**: melting experiments of WT *Ec*GUS and loop-swap mutants see a decrease of thermal stability of loop-swap mutants comparing to WT *Ec*GUS. **D:** chromatogram and molar mass of WT *Ec*GUS and loop-swap mutants. **E:** summary of **D** with monomeric mass of WT *Ec*GUS and of loop-swap mutants and their respective oligomerization states. **F**-**H**: EcGUS hydrophobicity contour map. The Hydrophobicity gradient is shown at the bottom of the figure. Green=methionine, purple=catalytic glutamic acid.

UNC361-9 forms aggregates during protein purification

Previously, we have characterized *Bacteroides fragilis* mL1 GUS (*Bf*GUS1) and *Roseberia hominis* mL1 GUS (*Rh*GUS1)[11, 13]. Using primary sequence alignment tools, we found that *Bf*GUS1 and *Rh*GUS1 lack similarity. Thus, we decided to characterize more mL1 GUSs to further our understandings on the functional diversity of these GUSs.

We chose five novel mL1 GUSs based on the sequence similarity network: *Facecalibacterium prausnitzii* mL1 GUS (*Fp*GUS4), *Bacteroides massiliensis* GUS (*Bm*GUS), *Eubacterium spp.* GUS, and two GUSs from unknown species (UNC361-9 and UNC361-11, respectively). In order to predict the functions of these mL1 GUSs, we first aligned the sequences of these novel Ml1 GUSs with Ml1 GUSs *Bf*GUS1 and *Rh*GUS as well as known β-galactosideases (GalAse): *Eisenbergiella tayi* GalAse (*Es*GUS) and *Fusicatenibacter saccharivorans* GUS/GalAse hybrid (*Fs*GUS/GalAse)[16]. We found that *Bm*GUS and *Bf*GUS1 share 70% primary sequence identity. We reasoned that *Bm*GUS and *Bf*GUS1 should possess similar biochemical functions.

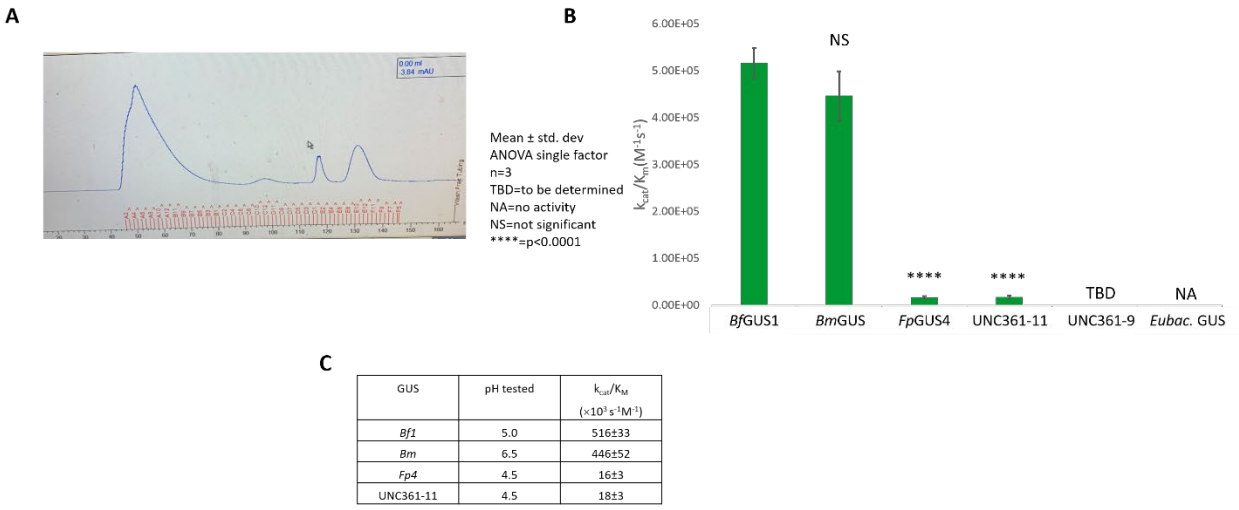
Otherwise, these GUSs share 18%-42% identity with each other, suggesting that they are unlikely to share similar substrate processing ability or substrate preferences.

To study these GUSs, we expressed his-*Fp*GUS4, his-*Bm*GUS, his- *Eubacterium spp*. GUS, his-UNC361-9, and his-UNC361-11 in *E. coli* BL21-Gold cell line for protein purification by fast protein liquid chromatography. All cell cultures were expressed nominally, and His-*Fp*GUS4, his-*Bm*GUS, his- *Eubacterium spp*. GUS and his-UNC361-11 were successfully purified by using Ni^{2+} column and Super200 size exclusion chromatography column. UNC361-9 was also purified by using the aforementioned columns. However, chromatogram of the size exclusion step shows that the majority of UNC361-9 population forms aggregates, as indicated by the wide band ranging from 45mL-76mL column volume (Figure 9A). There are three other bands in the chromatogram, most likely indicating three possible non-aggregate oligomerization state UNC361-9 exhibits in a periplasm-free environment. Because the absorbance of pNP$^-$ at 410nm diminishes at pH<6.5, we used an analog 4-methylumbelliferyl glucuronide (4MUG), which is fluorescence-active across pH4.0-7.4, to examine the catalytic activities of the eluents corresponding the major peaks from pH4.0-7.0. However, we found that none of the eluents possess 4MUG processing abilities at any of the pH we tested (data not shown). It is unknown to us why UNC361-9 has a strong tendency to form aggregates in a cell-free environment. First, we reasoned that UNC361-9 is soluble in aqueous buffers used in the protein purification, as the protein was eluted during both FPLC steps. Then, we used SignalP-5.0 to predict if there are any signal peptides in UNC361-9 primary sequence that might cause aggregation to occur, but we did not find any. Previously, we showed that some GUSs bind to flavin mononucleotide (FMN) to stabilize the protein structure. We did not anticipate UNC361-9 to bind to FMN, as the enzyme in buffer solution is transparent, not

yellow as is the color of FMN-binding GUSs. Nevertheless, we still checked if UNC361-9 contains motifs important for FMN binding[16]. As we expected, UNC361-9 lacks the tyrosine needed to form π-stacking interactions with the nucleotide. Therefore, UNC361-9 is unlikely to be an FMN-binding GUS. One probable explanation is that UNC361-9 directly interacts with other macromolecules, such as molecular chaperones or functional proteins, to form stable heterooligomer. Without interactions with other subunits of this heterooligomer, UNC361-9 will aggregate. To test this theory, we can either use experimental methods to identify interactors of UNC361-9 in the native species, or we can use published protein-protein interaction databases to predict the possible interactions. The greatest obstacle we face to test this hypothesis is that UNC361-9 is of unknown species. For either of our approaches, we need species information to proceed. Thus, the reason for UNC361-9 aggregation remains inconclusive for now.

mL1 GUSs exhibit diverse catalytic activities

After successfully purifying his-*Fp*GUS4, his-*Bm*GUS, his-*Eubacterium spp*. GUS, and his-UNC361-11, we tested their catalytic activities by using 4MUG and compared them with that of *Bf*GUS1. We did not compare them to *Rh*GUS1 because these novel GUSs share low primary sequence similarity with *Rh*GUS1. We found that *Bm*GUS and *Bf*GUS1 have similar kcat/Km values and that *Fp*GUS4 and UNC361-11 possess catalytic efficiency ~30-fold less than *Bm*GUS and *Bf*GUS1 (Figure 9B). Surprisingly, *Eubacterium spp*. GUS does not possess glyosidic bond cleavage activity of glucuronic acid from pH4.0 to pH7.4, even when [GUS] was increased to 100nM from 5nM (not shown). Through this experiment, we gained a deeper understanding on the functional diversity of mL1 GUSs.



C

GUS	pH tested	k_{cat}/K_M ($\times 10^3 s^{-1}M^{-1}$)
Bf1	5.0	516±33
Bm	6.5	446±52
Fp4	4.5	16±3
UNC361-11	4.5	18±3

Figure 9: characterization of novel mL1 GUSs. A: the FPLC size exclusion chromatogram of UNC361-9. B and C: catalytic efficiencies of active novel mL1 GUSs.

DISCUSSIONS

The gut microbiota possesses diverse chemical processing capabilities. The gut microbial GUSs are great examples of this functional diversity. A family of at least 279 unique sequences, gut microbial GUSs has been shown to process small, largely hydrophobic substrates such as SN-38G to largely charged oligosaccharides such as haparan sulfate[13]. In order to understand the importance of GUS in drug metabolite processing and to identify targets for novel GUS inhibitors, our works have focused on understanding the structural basis of this functional diversity. First, we sought to understand the molecular basis of GUS inhibition by UNC10201652, especially for *Bd*GUS. Although our study was inconclusive, we still gained information about *Bd*GUS structure and residues important for *Bd*GUS stability[14]. Then, we characterized the biochemical and biophysical features of loop-swap GUS mutants and gained insights on the importance of the extended active site loop on GUS activity. Lastly, we expressed, purified, and characterized a

panel of new mL1 GUSs. The knowledge we gained from these studies can be used to design small molecules as GUS inhibitors.

So far, we are only starting to uncover the complexity of the GUS enzyme family. In the gut microbiota only, there are at least 279 unique GUS sequences. Right now, we have characterized or are characterizing some of these GUSs, including *Bd*GUS, *Sa*GUS, *Fp*GUS4, *Bm*GUS, *Eubacterium spp*. GUS, UNC361-9, and UNC361-11[2, 11, 12, 13, 14, 15, 16,]. Using biochemical and structural methods to characterize all of the members in the gut microbial GUS family would be extremely time consuming. Thus, we need novel methods to characterize or predict the biochemical properties of GUSs. So far, we have mainly used bioinformatics to select GUS sequences using a specific rubric and to construct sequence similarity networks on the gene products of these selected sequences. From assessing the clustering and comparing the edge lengths between GUS nodes on the sequence similarity network, we can make educated guesses on the biochemical and structural properties of novel GUSs[13]. However, we still need to conduct experiments to test our predictions, and these prediction tools do not tell us information about the preferred substrates of these GUSs. Thus, using more complex computational models to predict the kinetics and substrate preferences of these GUSs are necessary to reduce the workload of novel GUS characterization. To do this, we can compile information about GUS substrate processing for all substrates and chemical properties of these substrates. By correlating these data, we can find trends to GUS substrate preference and can use this information to better predict the substrate preferences of uncharacterized GUSs. Further, computational modeling can guide us to a better understand on the underlying mechanisms of difference in GUS substrate preferences. Our *Bd*GUS inhibition study and loop swap study were attempts to experimentally

study how a region of GUS affects its kinetics and inhibition, but our results were inconclusive in both occasions. By using computational modeling methods, such as molecular dynamics and molecular kinetics, can help us visualize and give direct explanations to the effects of altering the key residues identified in structural studies.

In our loop swap study, we noticed that the WT *Ec*GUS has a high melting temperature (>94°C). GUSs consist of two immunoglobulin regions at the N-terminus, followed by a triose phosphate isomerase (TIM)-barrel domain. At the C-terminus, the GUSs consist 0-2 carbohydrate binding modules (CBMs) or domains of unknown functions (DUFs). Because the extended active site loop is within the TIM-barrel domain, we decided to investigate if some TIMs would also have similar melting temperatures. We found that Katebi and Jernigan have described that TIMs from thermophiles tend to form tetramers for the extra stability and that TIMs from mesophiles are likely to form dimers. *Ec*GUS and other known L1 GUSs, also form tetramers[17]. This observation raises questions about the evolutionary origins of gut microbial GUSs. Indeed, our group have found that gut microbial GUSs have different oligomerization states, and our studies outlined in this thesis have indicated that GUSs have different substrate preferences. It is interesting to see if there is a difference in evolutionary origins between each class of GUSs and if this difference contributes to the differences in substrate preferences. Also, we have only tested GUSs catalytic activities at 37°C. As *Ec*GUS has a high melting temperature, we can investigate *Ec*GUS catalytic activity as a function of temperature. The absorbance of *Ec*GUS changes at ~60°C. Although this does not signify a global denaturation event, it is possible that the immunoglobulin region of *Ec*GUS becomes unfolded. We can use this strategy to investigate the function of the

immunoglobulin region of *Ec*GUS. By understanding these problems, we will be able to unveil the origins of GUSs and of the differences of substrate preferences.

Our work on novel mL1 GUSs is still ongoing as this thesis is written. So far, we have established that all novel mL1 GUSs except *Bm*GUS are poor small hydrophobic substrate processors. It is possible that these mL1 GUSs are better processors for larger, more hydrophilic substrates. It is also possible that these mL1 GUSs are not glucuronide processors but are non-GUS β-hexuronidases. We need to test this theory by using such substrates (such as haparan sulfate and *p*-ntirophenyl β-galactoside). We also need to use structural techniques to understand the molecular details of the active site pockets of these GUSs. Because *Bm*GUS can process 4MUG with similar efficiency as *Bf*GUS1, we need to test *Bm*GUS against biologically important glucuronides, such as SN-38G, DCF-G, and serotonin glucuronide and find the importance of *Bm*GUS in processing these substrates. By continuing to characterize these mL1 GUSs, we will gain insights in the functional diversities and clinical significance of this class of GUSs.

CONCLUSION

In this thesis, we described our works on biochemical and biophysical characterization of *Bd*GUS, *Sa*GUS, loop swap mutants, and novel mL1 GUSs. First, we characterized *Bd*GUS and *Sa*GUS biochemically, and surprisingly, we found that *Bd*GUS is inhibited by UN10201652, so we attempted to identify the structural basis of this phenomenon. Then, we developed loop swap mutants to understand the biochemical importance of the extended active site loop region but found them to be catalytically inactive. As a result, we used biophysical methods to understand

the functions of these active site loops. Lastly, we characterized novel mL1 GUSs biochemically to investigate the functional diversity of this class of GUSs. Our work is instrumental in describing the functional and structural diversity of gut microbial GUS enzyme class. In the future, we can use our knowledge to further our understandings on the interactions between the gut microbiota and drug metabolites at and beyond GUSs. Also, we can refer to our insights gained in these studies to select non-L1 GUSs as targets for inhibitor development.

REFERENCES

1. Stein, A, Voigt, W, Jordan, K. Chemotherapy-Induced Diarrhea: Pathophysiology, Frequency and Guideline-Based Management. *Ther Adv Med Oncol* **2**(1): 51-63 (2010): DOI: 10.1177/1758834009355164
2. Wallace BD, Wang H, Lane KT, Scott JE, Orans J, Koo JS, Venkatesh MM, Jobin C, Yeh LA, Mani S, and Rebindo MR. Alleviating Cancer Drug Toxicity by Inhibiting a Bacterial Enzyme. *Science* **330**, 831 (2010): DOI: 10.1126/science.1191175.
3. Sender R, Fuchs S, and Milo R. Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol*. **14** (8): e1002533 (2016): DOI: 10.1371/journal.pbio.1002533.
4. Goodrich JK, Davenport ER, Clark AG, and Ley RE. The relationship betweenthe human genome and microbiome comes into view. *Ann. Rev. Genet.* **51**: 413-433 (2017): DOI: 10.1146/annurev-genet-1107110155532.
5. Shi, N, Li N, Duan X, and Niu H. Interaction between the gut microbiome and mucosal immune system. Mil. Med. Res. **4**: 14 (2017): DOI: 10.1186/s4779-017-0122-9.

6. Davis-Richardson AG, Ardissone AN, Raquel D, Simell V, Leonard MT, Kemppainen KM, Drew JC, Schatz D, Atkinson MA, Kolaczkowski B et al. *Bacteroides dorei* Dominates Gut Microbiome Prior to Autoimmunity in Finnish Children at High Risk for Type 1 Diabetes. *Front. Microbiol.* **5**: 678 (2014). DOI: 10.3389/fmicb.2014.00678.
7. Rowland I, Gibson G, Heinken A, Scot K, Swann J. Thiele I, and Tuohy K. Gut microbiota functions: metabolism of nutrients and other food components. *Eur. J. Nutr*. 57 (1): 1-24 (2018): DOI: 10.1007/s00394-017-1445-8.
8. LeBlanc JG, Milani C, de Giori GS, Sesma F, van Sinderen D, Ventura M. Bacteria as vitamin suppliers to their host: a gut microbiota perspective. *Curr. Opin. Biiotechnol*. **24** (2): 160-168 (2013): Doi: 10.1016/j.copnio.2012.08.005.
9. Qi B, Han M. Microbial Siderophore enterobactin promotes mitochondrial iron uptake and development of the host via interaction with ATP Synthase. *Cell*. **175**: 571-582 (2018): DOI: 10.1016/j.cell.2018.07.032.
10. Koppel N, Rekdal VM, and Balskus EP. Chemical transformation of xenobiotics by the human gut microbiota. *Science.* **356** (6344): eaag2770 (2019): DOI: 10.1126/science.aag2770.
11. Ervin SM, Li H, Lim L, Roberts LR, Liang X, Mani S, and Redinbo MR. Gut microbial β-glucuronidases reactivate estrogens as components of the estrobolome that reactivate estrogens. *J. Bio. Chem.* **294**: 18585-18599 (2019). DOI: 10.1074/jbc.RA119.010950.
12. Wallace, BD, Roberts, AB, Pollet, RM, Ingle JD, Biernat KA, Pellock SJ, O'Neal SK, Robinson SJ, Dollinger M, Figueroa E, McShane SR, Cohen RD, Jin, J, Frye SV, Zamboni WC, Pepe-Ranney C, Mani S, Kelly L, Redinbo MR. Structure and Inhibition of Microbiome β-

Glucuronidases Essential to the Alleviation of Cancer Drug Toxicity. *Chem. Biol.* **22 (9)**: 1238-1249 (2015): DOI: 10.1016/j.chembiol.2015.08.005

13. Pollet RM, D' Agostino EH, Walton WG, Xu Y, Little MS, Biernat KA, Pellock SJ, Patterson LM, Creekmore BC, Isenberg HN, Bahethi RR, Bhatt AP, Liu J, Gharaibeh RZ, and Redinbo MR. An Atlas of β-Glucuronidases in the Human Intestinal Microbiome. *Structure*. **25**: 7 (2017): DOI: 10.1016/j.str.2017.05.003.
14. Biernat BA, Pellock SJ, Bhatt AP, Bivins MM, Walton WG, Tran BNT, Wei L, Snider MC, Cesmat AP, Tripathy A, Erie DA, and Redinbo MR. Structure, Function, and Inhibition of Drug Reactivating Gut Microbial β-Glucuronidases. *Sci. Report*. **9**: 825 (2019): DOI: 10.1038/s41598-018-36069-w.
15. Pellock SJ, Walton WG, and Redinbo MR. Selecting a single stereocenter: the molecular nuances that differentiate β-hexuronidases in the human gut microbiome. *Biochemistry*. **58** (9): 1311-1317 (2019): DOI: 10.1021/acs.biochem.8b01285.
16. Pellock SJ, Walton WG, Ervin SM, Torres-Rivera D, Creekmore BC, Bergan G, Dunn ZD, Li B, Tripathy A, and Redinbo MR. Discovery and characterization of FMN-Binding β-glucuronidases in the human gut microbiome. *J. Mol. Biol*. **431** (5): 970-980 (2019): DOI: 10.1016/jmb.2019.01.013.
17. Katebi AR and Jernigan RL. The critical role of the loops of triosephosphate isomerase for its oligomerization, dynamics, and functionality. *Prot. Soc.* 23: 213-228 (2013): DOI: 10.1002/pro.2407.