Article:
Jariwala, Parth B., Pellock, Samuel J., Cloer, Erica W. et al. (8 more authors) (2020) Discovering the Microbial Enzymes Driving Drug Toxicity with Activity-Based Protein Profiling. ACS Chemical Biology. pp. 217-225. ISSN 1554-8929

https://doi.org/10.1021/acschembio.9b00788

Article

Discovering the Microbial Enzymes Driving Drug Toxicity with Activity-Based Protein Profiling

Parth B Jariwala, Samuel J Pellock, Dennis Goldfarb, Erica W Cloer, Marta Artola, Joshua B Simpson, Aadra P Bhatt, William G. Walton, Lee R Roberts, Michael B Major, Gideon J. Davies, Herman S. Overkleeft, and Matthew R. Redinbo

ACS Chem. Biol., **Just Accepted Manuscript** • DOI: 10.1021/acschembio.9b00788 • Publication Date (Web): 27 Nov 2019

Discovering the Microbial Enzymes Driving Drug Toxicity with Activity-Based Protein Profiling

Parth B. Jariwala[1], Samuel J. Pellock[1], Dennis Goldfarb[6,7], Erica W. Cloer[2], Marta Artola[11], Joshua B. Simpson[1], Aadra P. Bhatt[5], William G. Walton[1], Lee R. Roberts[9], Michael B. Major[7,8], Gideon J. Davies[10], Herman S. Overkleeft[11], and Matthew R. Redinbo*[,1,3,4]

[1]Department of Chemistry, [2]Lineberger Comprehensive Cancer Center, [3]Integrated Program for Biological and Genome Sciences, and [4]Departments of Biochemistry and Microbiology, [5]Center for Gastrointestinal Biology and Disease, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States; [6]Institute for Informatics, [7]Department of Cell Biology and Physiology, [8]Department of Otolaryngology, Washington University, St. Louis, Missouri, United States; [9]Exploratory Science Center, Merck & Company Inc, Cambridge, Massachusetts, United States; [10]York Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York, UK; [11]Department of Bioorganic Synthesis, Leiden Institute of Chemistry, Leiden University, Leiden, the Netherlands

ABSTRACT

It is increasingly clear that interindividual variability in human gut microbial composition contributes to differential drug responses. For example, gastrointestinal (GI) toxicity is not observed in all patients treated with the anticancer drug irinotecan, and it has been suggested that this variability is a result of differences in the types and levels of gut bacterial β-glucuronidases (GUS). GUS enzymes promote drug toxicity by hydrolyzing the inactive drug-glucuronide conjugate back to the active drug, which damages the GI epithelium. Proteomics-based identification of the exact GUS enzymes responsible for drug reactivation from the complexity of the human microbiota has not been accomplished, however. Here, we discover the specific bacterial GUS enzymes that generate SN-38, the active and toxic metabolite of irinotecan, from human fecal samples using a unique activity-based protein profiling (ABPP) platform. We identify and quantify gut bacterial GUS enzymes from human feces with an ABPP-enabled proteomics pipeline and then integrate this information with *ex vivo* kinetics to pinpoint the specific GUS enzymes responsible for SN-38 reactivation. Furthermore, the same approach also reveals the molecular basis for differential gut bacterial GUS inhibition observed between human fecal samples. Taken together, this work provides an unprecedented technical and bioinformatics pipeline to discover the microbial enzymes responsible for specific reactions from the complexity of human feces. Identifying such microbial enzymes may lead to precision biomarkers and novel drug targets to advance the promise of personalized medicine.

INTRODUCTION

The gut microbiota are capable of metabolizing a myriad of drugs,[1] and the biotransformation of these compounds by commensal intestinal bacteria can impact therapeutic outcomes by altering drug efficacy, and in some instances, inducing disease onset.[2] Since each person harbors a unique set of gut microbes, drug response varies considerably between individuals.[3] Although key recent reports have profoundly advanced our understanding of the central microbes and genes implicated in the metabolism of drugs[2,3], only a handful of studies have focused on gut bacterial proteins implicated in the biotransformation of drug metabolites.[4–6] Pinpointing the exact microbial enzymes that process drugs in the gut could lead to the development of precision biomarkers for the determination of therapeutic efficacy and may serve as drug targets for the modulation of the gut microbiota to optimize drug responses.

The gut bacterial β-glucuronidase (GUS) enzyme mediates drug-induced gastrointestinal (GI) toxicity by reversing glucuronidation, a Phase II transformation that inactivates and detoxifies drugs by conjugating them to glucuronic acid (GlcA) (**Figure S1a**).[7] Inactive drug glucuronides created in the liver traverse the biliary duct to reach the intestines where they are excreted from the body.[8] However, once in the gut, drug glucuronides have the potential to be reactivated via the hydrolytic removal of the GlcA tag by gut bacterial GUS enzymes. Intestinal reactivation of drug metabolites has been reported to cause acute, dose-limiting GI toxicities.[9,10] The severity of irinotecan-induced GI toxicity varies considerably between patients and may be due to the interindividual variability of the human gut microbiota.[11,12] Previous analysis of the Human Microbiome Project (HMP) stool sample database revealed that the gut microbiota contains hundreds of putative GUS enzymes with seven unique structural classes that display varying catalytic efficiencies against the reporter substrates *p*-nitrophenyl-β-D-glucuronide (*p*NP-GlcA) and 4-methylumbelliferone-β-D-glucuronide (4-MUG).[13,14] Since gut bacterial GUS enzymes process glucuronide conjugates with varying efficiencies, we hypothesized that interindividual differences in gut bacterial GUS abundance and composition might influence the differential drug response to irinotecan.

Efficient and facile strategies to identify the exact gut bacterial GUS enzymes that process drug glucuronides of interest from fecal material are lacking. Significant advancements in mass spectrometry (MS) and related bioinformatics software have made the identification and quantification of proteins from complex fecal supernatant possible.[15–17] However, recent work has shown that shotgun-based metaproteomics cannot accurately identify and quantify low abundance proteins from fecal lysates.[17] Activity-based probes (ABPs) serve as powerful tools to access low abundance targets and enrich for functionally active proteins from fecal lysate.[17,18] ABPs target the catalytic machinery of specific enzymes and can be outfitted with a chemical handle for target enrichment, enabling identification and quantitation using MS. Activity-based protein profiling (ABPP)-enabled GUS abundance data obtained from fecal

metaproteomes can then be correlated with *ex vivo* drug glucuronide processing data to identify the exact GUS enzymes that process drug glucuronides of interest (**Figure S1b**).

Using a unique pipeline that integrates ABPP-enabled GUS abundance data with *ex vivo* SN-38-G processing data, we pinpoint, from human feces, the exact bacterial GUS enzymes that reactivate SN-38, the active metabolite of the anti-cancer drug irinotecan. For the first time, we show that cyclophellitol-based ABPs can be used to identify and quantify gut bacterial GUS enzymes from human fecal lysate. We identify Loop 1 (L1) GUS enzymes as key modulators of SN-38 reactivation and verify this finding with *in vitro* kinetic data and structural modeling. Finally, we use the ABPP-enabled pipeline outlined in this study to provide a rationale for differential GUS inhibition between human fecal samples by previously designed piperazine-containing GUS inhibitors.

RESULTS

Cyclophellitol-based inhibitors and ABPs target structurally diverse gut bacterial GUS enzymes. Cyclophellitol-based epoxide and aziridine inhibitors **1** and **2** and activity-based probes (ABPs) **3** and **4** were previously developed to profile GUS in human cells (**Figure 1a**).[19] Since human and bacterial GUS utilize the same retaining mechanism to catalyze glucuronide hydrolysis, we hypothesized that **1–4** could also be used to target gut bacterial GUS enzymes from the human gut.[20] To confirm that **1–4** covalently label the catalytic glutamate in bacterial GUS enzymes, we determined the 2.4 Å resolution crystal structure of a bacterial GUS from the human gut commensal strain *B. uniformis* (*Bu*GUS-2) in complex with the unsubstituted cyclophellitol-based aziridine inhibitor (**2**) (**Table S1**). Examination of the active site revealed inhibitor **2** covalently linked to the catalytic nucleophile (E526) of *Bu*GUS-2 indicating that it is also an inhibitor of bacterial GUS (**Figure 1b**). Key contacts were also observed between the carboxylic acid of inhibitor **2** and N591 and K593, the conserved NxK motif that is essential for recognition of glucuronides by bacterial GUS (**Figure 1b**).[9] In-gel labelling of wild type and mutant enzymes using Cy5-ABP (**4**) further indicated that a functionally active GUS is necessary for labelling and that the NxK motif is essential for recognition of ABP **4** by bacterial GUS enzymes (**Figure 1c**).

The gut microbiota contains a structurally diverse assortment of bacterial GUS enzymes.[13] Using in-gel labelling studies, we found that ABP **4** labels most exogenously purified GUS enzymes from this structurally and functionally diverse group of enzymes (**Figure 1d**). Labelling was not observed for a GUS from *B. uniformis* (*Bu*GUS-3), which corroborates a recent study reporting that *Bu*GUS-3 does not process small molecule glucuronides and poorly processes GlcA-containing polysaccharides.[14] *In vitro* apparent IC_{50} values showed that **1–4** inhibit *E. coli* GUS (*Ec*GUS), *B. uniformis* GUS-1 (*Bu*GUS-1), and *Bu*GUS-2 with values ranging from 20 nM to 4 μM (**Figure S2**). Further kinetic analysis of GUS inactivation by **1–4** displayed k_i/K_I values that mirrored the IC_{50} values (**Table S2 and Figure S3**). Taken

together, these data establish that cyclophellitol-based inhibitors and ABPs target structurally diverse and functionally active gut bacterial GUS enzymes.

Cyclophellitol-based ABPs label GUS enzymes in mouse fecal mixtures. As a controlled proof-of-concept for labelling of GUS enzymes by the cyclophellitol-based ABPs, we collected fecal samples from wild-type germ-free mice and mice mono-associated with *gus*$^{+}$ *E. coli* (*Ec*GUS$^{M.A.}$; M.A., mono-associated). Labelling of *Ec*GUS$^{M.A.}$ fecal extracts with ABP **4** revealed a single, prominent band with a molecular weight indicative of recombinant *Ec*GUS (**Figure S5**). Heat denaturation of the fecal extracts from the mono-associated mice (*Ec*GUS$^{M.A. + H.K.}$; H.K., heat-killed) resulted in complete loss of labelling, which further establishes that these ABPs only label functionally active GUS enzymes. No significant labelling was observed in the fecal mixtures collected from germ-free mice which indicates that labelling of non-microbial protein is minimal. Finally, we show that labelling of *Ec*GUS by ABP **4** can be blocked in a complex fecal setting in a dose-dependent manner using the pan-GUS inhibitor, D-glucaro-1,4-lactone. These results demonstrate successful labelling of bacterial GUS enzymes in a controlled fecal matrix.

Gut bacterial GUS enzymes can be identified and quantified using cyclophellitol-based aziridine ABPs. After confirming that the cyclophellitol-based inhibitors and ABPs **1–4** label bacterial GUS enzymes *in vitro* and in a controlled mouse model, we performed ABPP to identify and quantify bacterial GUS enzymes present in human fecal samples collected from two females (F1 and F2) and two males (M1 and M2). We extracted total protein from human fecal lysates and enriched for GUS using the biotin-ABP (**3**) (**Figure 2a**). Resultant samples were analyzed by liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) and a bioinformatics pipeline that queried the integrated gene catalog (IGC) using MetaLab to assemble and quantify enriched protein groups (**Table S3 and Table S5**).[16,21] Protein groups were defined as GUS enzymes if sequences shared similarity to either *Ec*GUS, *C. perfringens* GUS (*Cp*GUS), *S. agalactiae* GUS (*Sa*GUS), or *B. fragilis* GUS (*Bf*GUS) and contained the catalytic glutamates as well as the NxK motif (**Figure 2a and Figure S6**). Analysis of the identified GUS protein groups revealed significant variations in taxa, structure, and abundance of the GUS enzymes present in the four fecal samples (**Figure 2b and Figure S7**). Individuals contained between 15–29 bacterial GUS protein groups, similar to a recent metagenomic study which showed that individuals harbor between 4–38 bacterial *gus* genes (**Figure S7a**).[13] Phylum-level analysis revealed that all four individuals predominantly contained GUS enzymes from Firmicutes but displayed substantial variation in GUS composition at lower taxa levels (**Figure 2b and Figure S7b**). Further examination using a previously defined GUS structure rubric allowed us to analyze the identified GUS protein groups based

on three-dimensional structure, which revealed significant structural diversity (**Figure S7c**).[13] We have developed an ABPP-enabled proteomics pipeline to identify and quantify functionally active GUS enzymes present in human fecal material.

Cyclophellitol-based aziridine ABPs also target GH3 β-glucosidases. Because the human gut microbiota contains a diverse assortment of glycoside hydrolases (GHs), performing ABPP from fecal material is a veritable test of the selectivity of the GUS ABPs.[22] Sequence analysis of the protein groups identified from human fecal extracts revealed a major off-target hit, GH3 β-glucosidases (**Figure 3a and Table S5**). The GH3 β-glucosidases enriched are structurally similar but occupy two topologically distinct categories that we have termed "Type I" and "Type II" (**Figure S8**). Manual docking analysis of the untagged ABP in structurally characterized GH3 β-glucosidases revealed favorable positioning of the catalytic nucleophile for attack of the aziridine ring (**Figure 3b**). Additionally, an arginine residue was also present (R538 and R50 in Type I and Type II, respectively) that may contact the carboxylic acid moiety of the probe, enabling recognition and subsequent processing of ABPs by GH3 β-glucosidases. We expressed and purified both a Type I and Type II β-glucosidase identified in the fecal samples and confirmed *in vitro* that they are labelled by high concentrations of ABP **3** (**Figure 3c and Figure S8**). Despite labelling of the GH3 β-glucosidase by GlcA-like aziridine probes, neither type of β-glucosidase processed *p*NP-GlcA, suggesting that off-target labelling of β-glucosidases is probably due to the reactive aziridine moiety of the GUS ABPs (**Figure 3d, e**).

Gut bacterial Loop 1 GUS enzymes are key mediators of SN-38 reactivation. After successfully identifying and quantifying bacterial GUS enzymes from human feces, we investigated whether we could identify the exact bacterial GUS enzymes responsible for SN-38 reactivation in the gut by integrating ABPP-enabled GUS abundance information with *ex vivo* SN-38-G processing data. We measured *ex vivo* SN-38-G hydrolysis by human fecal extracts, which revealed faster processing for F2 and M1 than F1 and M2 (**Figure 4a and 4b, Figure S9**). We found a strong correlation between Loop 1 (L1) GUS abundance and rate of SN-38-G hydrolysis when compared to total bacterial GUS abundance (**Figure 4c**). No correlation was found between either human GUS or other GUS structural classes and the rate of SN-38-G hydrolysis (**Figure S10**). We validated the correlation by assessing the catalytic efficiency of SN-38-G processing by a panel of purified GUS enzymes from various GUS structural classes and found that bacterial L1 GUS enzymes process SN-38-G most efficiently (**Figure 4d**). We also found that F2, M1, and M2 were abundant in L1 GUS enzymes that had sequence identities ≥ 90% to *E. eligens* GUS (*Ee*GUS, PDB: 6BJQ) (**Table S3**). We expressed and purified *Ee*GUS and found that it processed SN-38-G faster than all other examined GUS enzymes *in vitro*. A close examination of the crystal structure of

*Ee*GUS reveals a hydrophobic active site pocket formed at the interface of two monomers in the L1 tetramer (**Figure 4e**).[23] The hydrophobic pocket formed by the oligomeric interface appears to optimally recognize hydrophobic small molecule glucuronides like SN-38-G. Taken together, correlation analysis between *ex vivo* processing data and ABPP-enabled GUS abundance data, further informed by *in vitro* enzyme kinetics and structural modeling, provides a molecular rationale for interindividual variation in SN-38 reactivation in human fecal samples, and identifies L1 GUS enzymes, particularly *Ee*GUS, as key molecular regulators of efficient SN-38-G reactivation.

Piperazine containing small molecules inhibitors target gut bacterial Loop 1 GUS enzymes. Finally, we sought to extend these investigations to explain differential gut bacterial GUS inhibition. We have developed selective, potent, and non-lethal gut bacterial GUS inhibitors that block the reactivation of drug metabolites like SN-38-G (**Figure 5a**).[14,24] The piperazine moiety in both UNC4917 and UNC10201652 acts as a warhead that targets the catalytic machinery of bacterial GUS enzymes by intercepting the catalytic cycle.[23] We find that SN-38-G processing was differentially inhibited in all four human fecal extracts using these GUS inhibitors (**Figure 5b**). Subsequent analyses reveal a strong correlation between inhibition and L1 GUS abundance while no correlation was observed for the other GUS structural classes, confirming previous work that UNC4917 and UNC10201652 act as L1-specific GUS inhibitors (**Figure 5c** and **Figure S11**).[9,24] Furthermore, we verified that UNC4510, a negative control analog of UNC10201652 that contains a methylated piperazine moiety, poorly inhibited SN-38-G processing for all GUS enzymes.[23] These data show that L1-specific GUS inhibitors can block SN-38-G processing only in individuals whose fecal gut microbiota is highly abundant in L1 GUS enzymes.

DISCUSSION

Here we show that cyclophellitol-based epoxide and aziridine inhibitors and ABPs can target gut bacterial GUS enzymes. Using a combination of *in vitro* and in-gel assays, we find that **1**–**4** target structurally diverse GUS enzymes with varying potencies. The variation in GUS inhibition is likely due to differences both in oligomeric states and active site features of the bacterial GUS enzymes examined (**Figure S4**). For example, we observe more potent inhibition of *E. coli* GUS by the biotin-ABP (**3**) when compared to the unsubstituted aziridine inhibitor (**2**). Like *E. eligens* GUS, previous structural work has shown that *E. coli* GUS is a tetramer with a hydrophobic active site formed at the interface of its monomers.[4,9] Thus, the increase in inhibition by ABP **3** compared to inhibitor **2** is likely due to hydrophobic interactions between the *E. coli* GUS active site and the nonpolar alkyl chain present in **3**. Furthermore, ABP **3** and **4** displayed notable differences in inhibition for all GUS enzymes. The Cy5-

ABP (**4**) is weaker at inhibiting GUS enzymes than the biotin-ABP (**3**), and this is likely due to steric clashes between the bulky fluorophore group and the GUS enzymes examined.

Most importantly, we show that gut microbial GUS enzymes can be identified and quantified from human feces using ABPP. We were interested in examining GUS sequence information obtained through our ABPP-enabled pipeline to better understand the structural diversity of GUS enzymes present in the gut microbiome and to correlate GUS structure to SN-38-G processing. By using powerful metaproteomic software tools like MetaLab[16] and Unipept[25], we also show that peptide MS data can be employed to obtain taxon information for GUS-producing bacterial species found in human feces. However, many protein groups could not be assigned to lower taxonomic ranks due to a lack of taxon-specific distinctive peptides. Thus, in the future, strategies that both increase peptide count and yield longer peptides for MS analysis should be explored to improve taxonomy assignment using ABPP. Metagenomic sequencing could be pursued to develop a sample-specific sequence database to query peptides, but this approach may be economically prohibitive.[26] While other methods have coupled deep-sequencing with ABPs to uncover GUS-producing species[27], we provide evidence here that ABPP alone can be used to obtain a strong level of taxa information for GUS-producing bacterial species from human fecal samples.

An unexpected yet exciting finding from our investigation was the identification of GH3 β-glucosidases as an off-target hits. We identified two topologically distinct GH3 β-glucosidases as off-target hits of the GUS ABPs. Since ABPs sample enzyme function, we initially hypothesized that the identified GH3 β-glucosidases may process GlcA-containing substrates, but *in vitro* assays using *p*NP-GlcA revealed that these enzymes do not process glucuronides, and are in fact, off-target hits (**Figure 3d, e**). Further assessment of previously published GH3 β-glucosidase structures reveal a solvent exposed active site and an arginine residue that interacts with the carboxylic acid moiety of GlcA. These features combined with the highly reactive nature of the aziridine moiety in the cyclophellitol-based ABP likely cause labelling of the GH3 β-glucosidases. The identification of only one class of off-target hits is remarkable given that the human gut microbiome is one of the most glycoside hydrolase rich environments found in nature[22], and further demonstrates that cyclophellitol-based GUS ABPs are incredibly precise and effective probes.

Integration of ABPP-enabled GUS abundance with *ex vivo* SN-38-G processing data enabled the identification of L1 GUS enzymes as the key molecular regulators of SN-38-G turnover. Importantly, this predictive correlation was validated by both *in vitro* enzyme kinetics and structural modeling. Although we have strongly correlated L1 GUS enzymes to SN-38-G processing, they are lead biomarkers that will need to be further characterized for clinical use. For example, the ABPP methodology outlined here does not examine the bacterial cell uptake of these glucuronide substrates. Further studies analyzing relevant

gut bacterial isolates will be needed to assess cellular uptake of SN-38-G. Additionally, the gut microbiota contains hundreds of unique GUS enzymes, all of which are not encompassed by the four fecal samples used in this study. The strategy outlined here provides a foundation on which future proteomics and drug processing can be added to extant datasets to re-run correlation analyses and identify new biomarkers.

We also show that SN-38-G processing can be inhibited in complex metaproteomes using previously designed L1-specific GUS inhibitors and that GUS inhibition can be accurately predicted with probe-derived proteomics data. Interestingly, our data indicates that UNC10201652 is more potent than UNC4917 at inhibiting L1 GUS enzymes in fecal samples, a similar result found in a previous study.[23] Structure-activity relationships can be conducted against a large assortment of GUS enzymes found in fecal samples using this strategy to identify the inhibitor chemotypes that block GUS enzymes from reactivating drug glucuronides like SN-38-G. Coupling ABPP-enabled GUS abundance with *ex vivo* inhibition data can serve as a powerful strategy to conduct structure-activity relationships in a high-throughput manner. Since we have a limited understanding of enzyme-substrate pairs in the microbiome, we believe it is imperative that high precision gut bacterial inhibitors be developed in lieu of broad-spectrum drugs like antibiotics or inhibitors that target enzymes classes.

Recent work was published on a distinct GUS ABP composed of a GlcA warhead linked to a quinone methide leaving group at the anomeric position.[27] The main difference between the quinone methide ABP and the cyclophellitol-based aziridine ABP employed here is target specificity. As noted by Wright and co-workers, the quinone methide ABP, once activated, has the potential to leave the enzyme active site and label off-target macromolecules.[27] In contrast, the cyclophellitol-based aziridine ABP employed here reacts directly with the GUS active site in a mechanism-based fashion to form a covalent bond with the glutamate nucleophile, likely reducing the number of off-targets. Although labeling live bacteria with a quinone-methide ABP coupled with FACS sorting and 16S rRNA sequencing can give general taxa information on bacterial populations found in feces[27], it seems less suitable for sequence-level identification and quantification of active GUS enzymes from fecal supernatant due to the promiscuity of the activated quinone-methide leaving group.

In summary, we determined the composition and relative abundance of bacterial GUS enzymes from human fecal samples using ABPP. We utilized these data to identify the key modulators of SN-38 reactivation and to rationalize differential GUS inhibition across fecal samples. While we focused on SN-38-G metabolism in the present study, the combination of proteomics data and functional assays can be employed to pinpoint specific GUS enzymes implicated in the reactivation of other drug glucuronides. Furthermore, proteomics-activity correlations provide a universal tool to identify a specific molecular target for any enzyme activity in the microbiome, an approach that is only limited and facilitated by the

current set and continued development of ABPs that target gut bacterial enzymes.[17–19,27–30] Together, the data gained from this ABPP approach enables the identification of potential gut bacterial drug targets for the molecular modulation of the gut microbiota and can be employed to reveal highly precise biomarkers for possible diagnostic development in the era of personalized medicine.

METHODS

Full details for all materials and methods are provided in the Supporting Information.

ASSOCIATED CONTENT

Supporting Information

This material is available free of charge via the internet at http://pubs.acs.org.

Methods and Figures S1-S12 (**PDF**)

Table S1, Crystallographic statistics for *B. uniformis* GUS-2 bound to the unsubstituted cyclophellitol-based aziridine ABP (**2**). (**XLSX**)

Table S2, Kinetic parameters for inhibition of select gut bacterial GUS enzymes by cyclophellitol-based inhibitors and ABPs. (**XLSX**)

Table S3, Protein groups identified as GUS and compiled loop abundance data. (**XLSX**)

Table S4, Taxa annotation for each GUS identified peptide and compiled taxa abundance data. (**XLSX**)

Table S5, All protein groups identified. (**XLSX**)

AUTHOR INFORMATION

Corresponding Author

*(M.R.R.) E-mail: redinbo@unc.edu

Author Contributions

P.B.J, S.J.P, and M.R.R. developed the project. P.B.J. designed and performed all experiments. S.J.P. and L.R.R. aided in designing experiments. D.G. and E.W.C. designed and performed all MS-based experiments and provided helpful discussions about the project. D.G. designed and performed all bioinformatic analyses. M.A. synthesized the cyclophellitol-based compounds used in this study. J.B.S. assisted with collecting data and provided helpful discussions about the project. A.P.B conducted all animal studies and provided helpful discussions about the project. W.G.W expressed and purified all recombinant enzymes used in this study. P.B.J, S.J.P, and M.R.R. wrote the manuscript and all authors provided helpful edits and commentary.

Notes

The authors declare the following competing interest(s): M.R.R. is a Founder and Board Member of Symberix, Inc., which is developing microbiome-targeted therapeutics.

Acknowledgements

The authors thank members of their respective laboratories for experimental assistance, helpful discussions, and manuscript revisions. Funded by National Institutes of Health grants CA098468 and CA207416, and the Merck Exploratory Science Center (M.R.R.), T32-DK007737 and Pilot and Feasibility funds from P30 DK034987 (A.P.B.), American Cancer Society grant RSG-14-1657 068-01-TBE and North Carolina University Cancer Research Funds (M.B.M.), the Netherlands Organization for Scientific Research NWO; TOP grant (H.S.O.), the European Research Council grants ERC-2011-AdG-290836 (H.S.O.) and ERC-2012-AdG-32294 (G.J.D.), and Biotechnology and Biological Sciences Research Council grants BB/R001162/1 and BB/M011151/1 (G.J.D.). G.J.D. also thanks the Royal Society for the Ken Murray Research Professorship.

REFERENCES

(1) Zimmermann, M., Zimmermann-Kogadeeva, M., Wegmann, R., and Goodman, A. L. (2019) Mapping human microbiome drug metabolism by gut bacteria and their genes. *Nature 570*, 462–467.

(2) Koppel, N., Rekdal, V. M., and Balskus, E. P. (2017) Chemical transformation of xenobiotics by the human gut microbiota. *Science 356*, 1246–1257.

(3) Lam, K. N., Alexander, M., and Turnbaugh, P. J. (2019) Precision Medicine Goes Microscopic: Engineering the Microbiome to Improve Drug Outcomes. *Cell Host Microbe 26*, 22–34.

(4) Wallace, B. D., Wang, H., Lane, K. T., Scott, J. E., Orans, J., Koo, J. S., Venkatesh, M., Jobin, C., Yeh, L., Mani, S., and Redinbo, M. R. (2010) Alleviating Cancer Drug Toxicity by Inhibiting a Bacterial Enzyme. *Science 330*, 831–835.

(5) van Kessel, S. P., Frye, A. K., El-Gendy, A. O., Castejon, M., Keshavarzian, A., van Dijk, G., and El Aidy, S. (2019) Gut bacterial tyrosine decarboxylases restrict levels of levodopa in the treatment of Parkinson's disease. *Nat. Commun. 10*, 1–11.

(6) Koppel, N., Bisanz, J. E., Pandelia, M. E., Turnbaugh, P. J., and Balskus, E. P. (2018) Discovery and characterization of a prevalent human gut bacterial enzyme sufficient for the inactivation of a family of plant toxins. *Elife 7*, 1–32.

(7) Dutton, G. J. (1966) Glucuronic Acid, Free and Combined, Biochemistry, Pharmacology, and Medicine. Academic Press, New York.

(8) Dutton, G. J. (1980) Glucuronidation of Drugs and Other Compounds. CRC press, Boca Raton.
(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-Ranney, C., Mani, S., Kelly, L., and Redinbo, M. R. (2015) Structure and Inhibition of Microbiome β-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 β-glucuronidase inhibition protects mice against enteropathy induced by indomethacin, ketoprofen or diclofenac: mode of action and pharmacokinetics. *Xenobiotica 44*, 28–35.
(11) Kweekel, D., Guchelaar, H. J., and Gelderblom, H. (2008) Clinical and pharmacogenetic factors associated with irinotecan toxicity. *Cancer Treat. Rev. 34*, 656–669.
(12) Guthrie, L., Gupta, S., Daily, J., and Kelly, L. (2017) Human microbiome signatures of differential colorectal cancer drug metabolism. *NPJ Biofilms Microbiomes 3*, 1–8.
(13) 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 β-Glucuronidases in the Human Intestinal Microbiome. *Structure 25*, 967–977.
(14) 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 -glucuronidases from the human gut microbe Bacteroides uniformis. *J. Biol. Chem. 293*, 18559–18573.
(15) Zhang, X., and Figeys, D. (2019) Perspective and Guidelines for Metaproteomics in Microbiome Studies. *J. Proteome Res. 18*, 2370–2380.
(16) Cheng, K., Ning, Z., Zhang, X., Li, L., Liao, B., Mayne, J., Stintzi, A., and Figeys, D. (2017) MetaLab: an automated pipeline for metaproteomic data analysis. *Microbiome 5*, 1–10.
(17) Mayers, M. D., Moon, C., Stupp, G. S., Su, A. I., and Wolan, D. W. (2017) Quantitative Metaproteomics and Activity-Based Probe Enrichment Reveals Significant Alterations in Protein Expression from a Mouse Model of Inflammatory Bowel Disease. *J. Proteome Res. 16*, 1014–1026.
(18) Parasar, B., Zhou, H., Xiao, X., Shi, Q., Brito, I. L., and Chang, P. V. (2019) Chemoproteomic Profiling of Gut Microbiota-Associated Bile Salt Hydrolase Activity. *ACS Cent. Sci. 5*, 867–873.
(19) Wu, L., Jiang, J., Jin, Y., Kallemeijn, W. W., Kuo, C. L., Artola, M., Dai, W., Van Elk, C., Van Eijk, M., Van Der Marel, G. A., Codée, J. D. C., Florea, B. I., Aerts, J. M. F. G., Overkleeft, H. S., and Davies, G. J. (2017) Activity-based probes for functional interrogation of retaining β-glucuronidases. *Nat. Chem. Biol. 13*, 867–873.
(20) Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014) The

carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res. 42*, 490–495.

(21) Li, J., Wang, J., Jia, H., Cai, X., Zhong, H., Feng, Q., Sunagawa, S., Arumugam, M., Kultima, J. R., Prifti, E., Nielsen, T., Juncker, A. S., Manichanh, C., Chen, B., Zhang, W., Levenez, F., Wang, J., Xu, X., Xiao, L., Liang, S., Zhang, D., Zhang, Z., Chen, W., Zhao, H., Al-Aama, J. Y., Edris, S., Yang, H., Wang, J., Hansen, T., Nielsen, H. B., Brunak, S., Kristiansen, K., Guarner, F., Pedersen, O., Doré, J., Ehrlich, S. D., and Bork, P. (2014) An integrated catalog of reference genes in the human gut microbiome. *Nat. Biotechnol. 32*, 834–841.

(22) Berlemont, R., and Martiny, A. C. (2016) Glycoside Hydrolases across Environmental Microbial Communities. *PLoS Comput. Biol. 12*, 1–16.

(23) Pellock, S. J., Creekmore, B. C., Walton, W. G., Mehta, N., Biernat, K. A., Cesmat, A. P., Ariyarathna, Y., Dunn, Z. D., Li, B., Jin, J., James, L. I., and Redinbo, M. R. (2018) Gut Microbial β-Glucuronidase Inhibition via Catalytic Cycle Interception. *ACS Cent. Sci. 4*, 868–879.

(24) 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 β-glucuronidases. *Sci. Rep. 9*, 1–15.

(25) Mesuere, B., Willems, T., Van Der Jeugt, F., Devreese, B., Vandamme, P., and Dawyndt, P. (2016) Unipept web services for metaproteomics analysis. *Bioinformatics 32*, 1746–1748.

(26) Proctor, L. M., Creasy, H. H., Fettweis, J. M., Lloyd-Price, J., Mahurkar, A., Zhou, W., Buck, G. A., Snyder, M. P., Strauss, J. F., Weinstock, G. M., White, O., and Huttenhower, C. (2019) The Integrative Human Microbiome Project. *Nature 569*, 641–648.

(27) Whidbey, C., Sadler, N. C., Nair, R. N., Volk, R. F., Deleon, A. J., Bramer, L. M., Fansler, S. J., Hansen, J. R., Shukla, A. K., Jansson, J. K., Thrall, B. D., and Wrig. (2019) A Probe-Enabled Approach for the Selective Isolation and Characterization of Functionally Active Subpopulations in the Gut Microbiome. *J. Am. Chem. Soc. 141*, 42–47.

(28) Schröder, S. P., De Boer, C., McGregor, N. G. S., Rowland, R. J., Moroz, O., Blagova, E., Reijngoud, J., Arentshorst, M., Osborn, D., Morant, M. D., Abbate, E., Stringer, M. A., Krogh, K. B. R. M., Raich, L., Rovira, C., Berrin, J. G., Van Wezel, G. P., Ram, A. F. J., Florea, B. I., Van Der Marel, G. A., Codée, J. D. C., Wilson, K. S., Wu, L., Davies, G. J., and Overkleeft, H. S. (2019) Dynamic and Functional Profiling of Xylan-Degrading Enzymes in Aspergillus Secretomes Using Activity-Based Probes. *ACS Cent. Sci. 5*, 1067–1078.

(29) Jiang, J., Kuo, C. L., Wu, L., Franke, C., Kallemeijn, W. W., Florea, B. I., Van Meel, E., Van Der Marel, G. A., Codée, J. D. C., Boot, R. G., Davies, G. J., Overkleeft, H. S., and Aerts, J. M. F. G. (2016) Detection of active mammalian GH31 α-glucosidases in health and disease using in-class, broad-spectrum activity-based probes. *ACS Cent. Sci. 2*, 351–358.

(30) Li, K. Y., Jiang, J., Witte, M. D., Kallemeijn, W. W., Donker-Koopman, W. E., Boot, R. G., Aerts, J. M. F. G., Codée, J. D. C., Van Der Marel, G. A., and Overkleeft, H. S. (2014) Exploring functional cyclophellitol analogues as human retaining beta-glucosidase inhibitors. *Org. Biomol. Chem. 12*, 7786–7791.

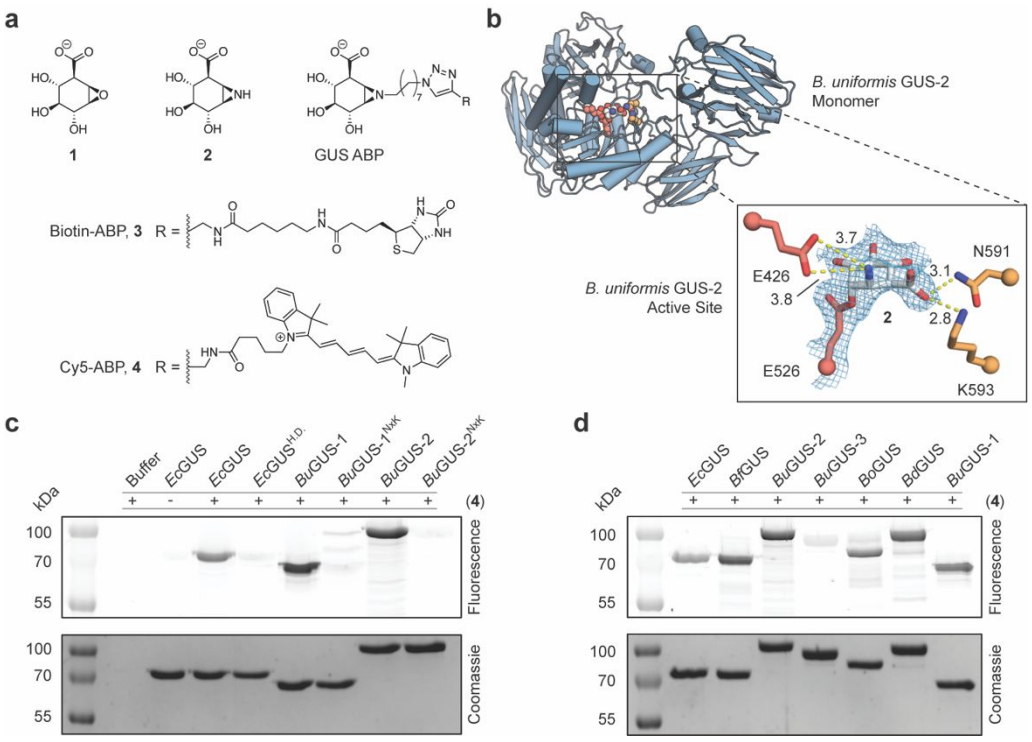


Figure 1. Cyclophellitol-based inhibitors and ABPs label structurally diverse gut bacterial GUS enzymes. (**a**) Cyclophellitol-based epoxide and aziridine inhibitors **1** and **2** and ABPs **3** and **4**. (**b**) A 2.4 Å resolution crystal structure (PDB: 6NZG) of inhibitor **2** bound to *Bu*GUS-2. Inset shows $2F_o - F_c$ map (after refinement) at 1 σ and distances are shown in Å. (**c**) In-gel fluorescence labelling of wild type and inactive GUS controls by ABP **4**. *E. coli* GUS (*Ec*GUS), heat-denatured *E. coli* GUS ($EcGUS^{H.D.}$), *B. uniformis* GUS-1 (*Bu*GUS-1), *B. uniformis* GUS-2 (*Bu*GUS-2), and *Bu*GUS-1 and *Bu*GUS-2 mutants ($BuGUS\text{-}1^{NxK}$ and $BuGUS\text{-}2^{NxK}$) where the asparagine and lysine residues of the NxK motif have been mutated to alanines. (**d**) In-gel fluorescence labelling of structurally diverse gut bacterial GUS by ABP **4**. *B. fragilis* GUS (*Bf*GUS), *B. uniformis* GUS-3 (*Bu*GUS-3), *B. ovatus* GUS (*Bo*GUS), and *B. dorei* GUS (*Bd*GUS). All wild type and mutant proteins were exogenously purified.

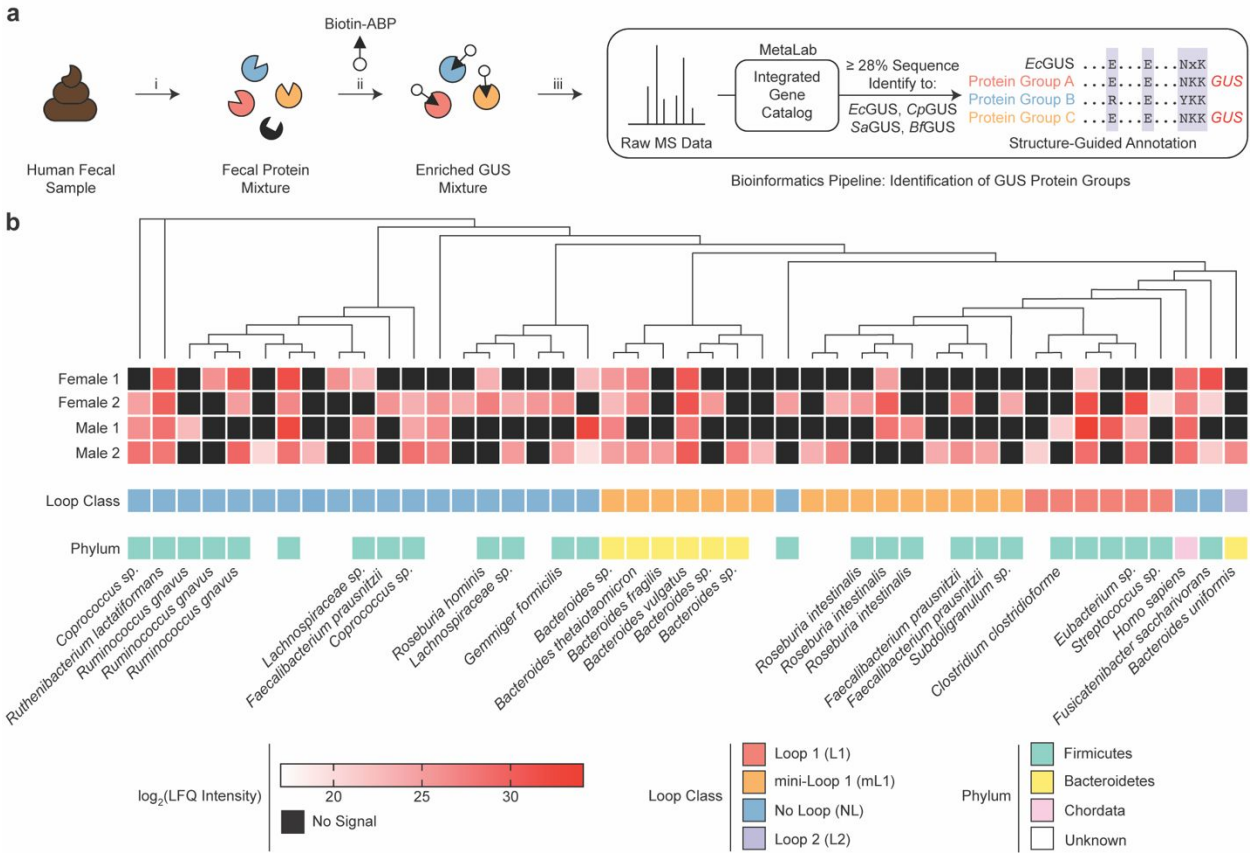


Figure 2. Probe-enabled proteomics and structure-guided bioinformatics enable identification and relative quantitation of bacterial GUS enzymes from human fecal samples. (**a**) General schematic of the probe-enabled proteomics pipeline used to identify and quantify GUS from fecal material. In brief, (**i**) proteins are extracted from feces using ultrasonication, (**ii**) GUS enzymes are enriched using the pre-clicked biotin-ABP (**3**) and streptavidin beads, and (**iii**) MetaLab is used to query the integrated gene catalog using raw MS data to assemble and quantify protein groups. Only proteins with the GUS fold and active site features, including the catalytic glutamates (E) and NxK motif, are defined as GUS enzymes. (**b**) Heatmap of identified GUS protein groups organized by sequence similarity and color coded by abundance. GUS abundance is represented by LFQ intensities, which are normalized and combined peptide signal intensities as determined by the MaxLFQ algorithm in MaxQuant. Further taxonomic classifications are shown below the abundance heatmap. Unknown refers to protein groups where the phylum assignment was ambiguous due to mapping of GUS peptides to multiple phyla. Sequence-level information for each protein group can be found in **Table S3.**

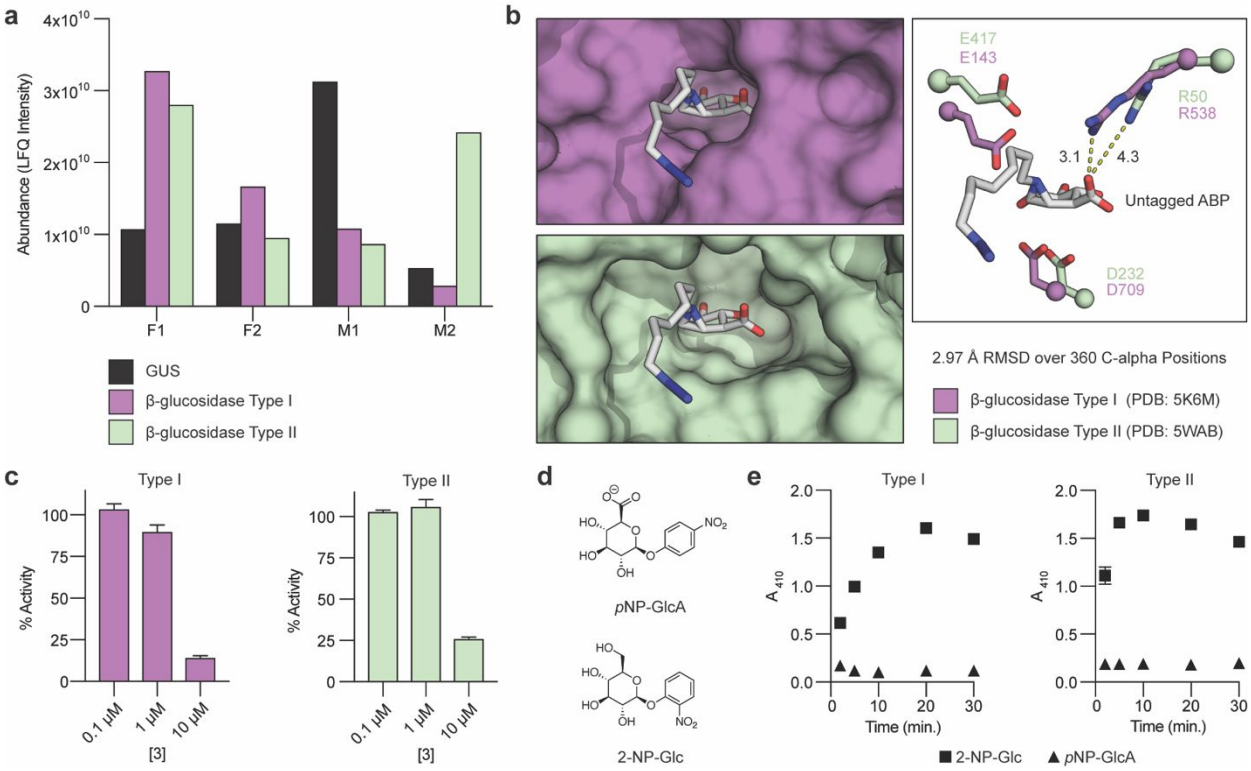


Figure 3. β-glucosidase is a specific off-target of GUS ABPs. (**a**) Protein abundance of GUS, Type I β-glucosidase, and Type II β-glucosidase identified from human fecal samples. (**b**) Conserved active sites of topologically distinct Type I (PDB: 5K6M) and Type II (PDB: 5WAB) β-glucosidases with the untagged ABP manually docked in PyMol. Distances are shown in Å. (**c**) Type I and Type II β-glucosidase inhibition by the biotin-ABP (**3**). (**d**) Chemical structures of 2-nitrophenyl β-D-glucopyranoside (2-NP-Glc) and *p*-nitrophenyl-β-D-glucuronide (*p*NP-GlcA). (**e**) *In-vitro* processing of 2-NP-Glc and *p*NP-GlcA by Type I and Type II β-glucosidases. All percent activity and rate values shown are mean values ± standard deviation using *N*=3 biological replicates.

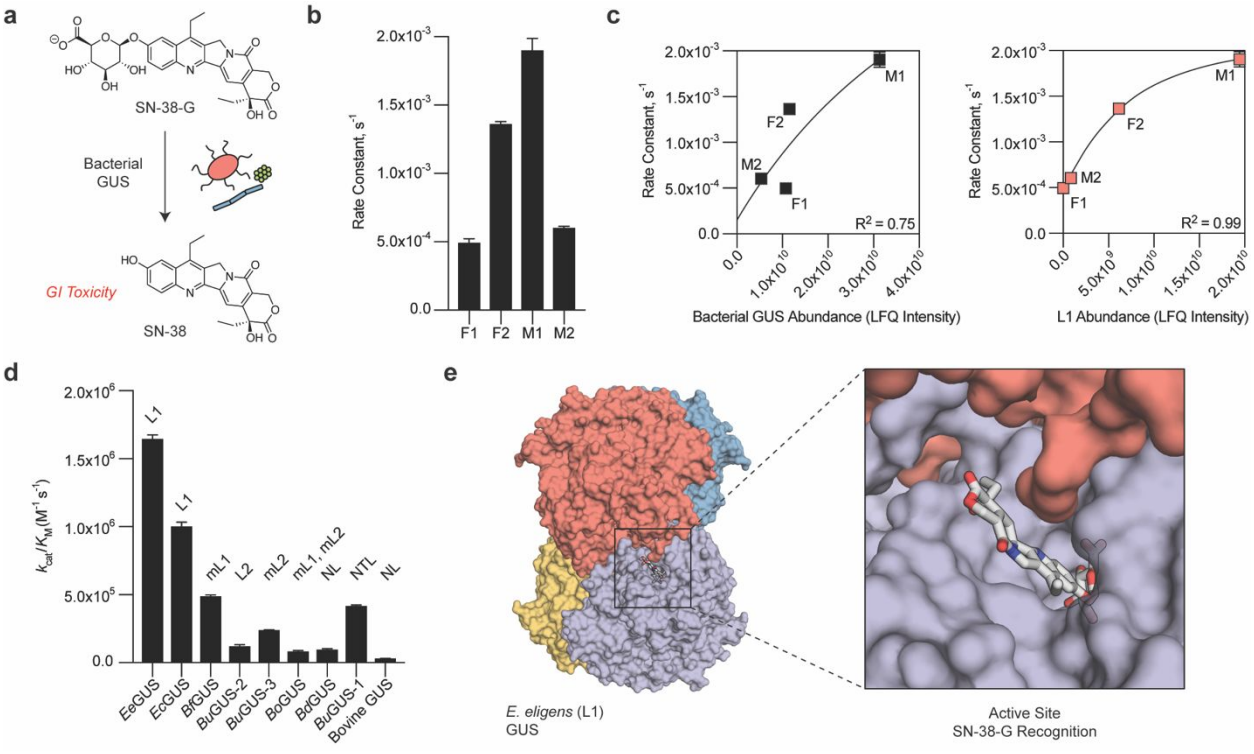


Figure 4. ABPP coupled with *ex vivo* processing data provides a molecular rationale for GUS-mediated SN-38 reactivation. (**a**) SN-38 glucuronide (SN-38-G) is the inactive metabolite of the topoisomerase I inhibitor irinotecan and is reactivated to SN-38 in the gut by bacterial GUS enzymes, resulting in acute, dose-limiting GI toxicity. (**b**) *Ex vivo* processing of SN-38-G by human fecal protein extracts. (**c**) Correlation analysis between total bacterial GUS abundance and Loop 1 (L1) GUS abundance against SN-38-G processing. (**d**) *In vitro* catalytic efficiencies of SN-38-G processing for a representative panel of GUS enzymes of different loop types. mini-Loop 1 (mL1); Loop 2 (L2); mini-Loop 2 (mL2); mini-Loop 1, mini-Loop 2 (mL1, mL2); No Loop (NL); N-Terminal Loop (NTL) (**e**) Quaternary structure of *E. eligens* GUS (*Ee*GUS, PDB: 6BJQ) with SN-38-G manually docked in PyMol.

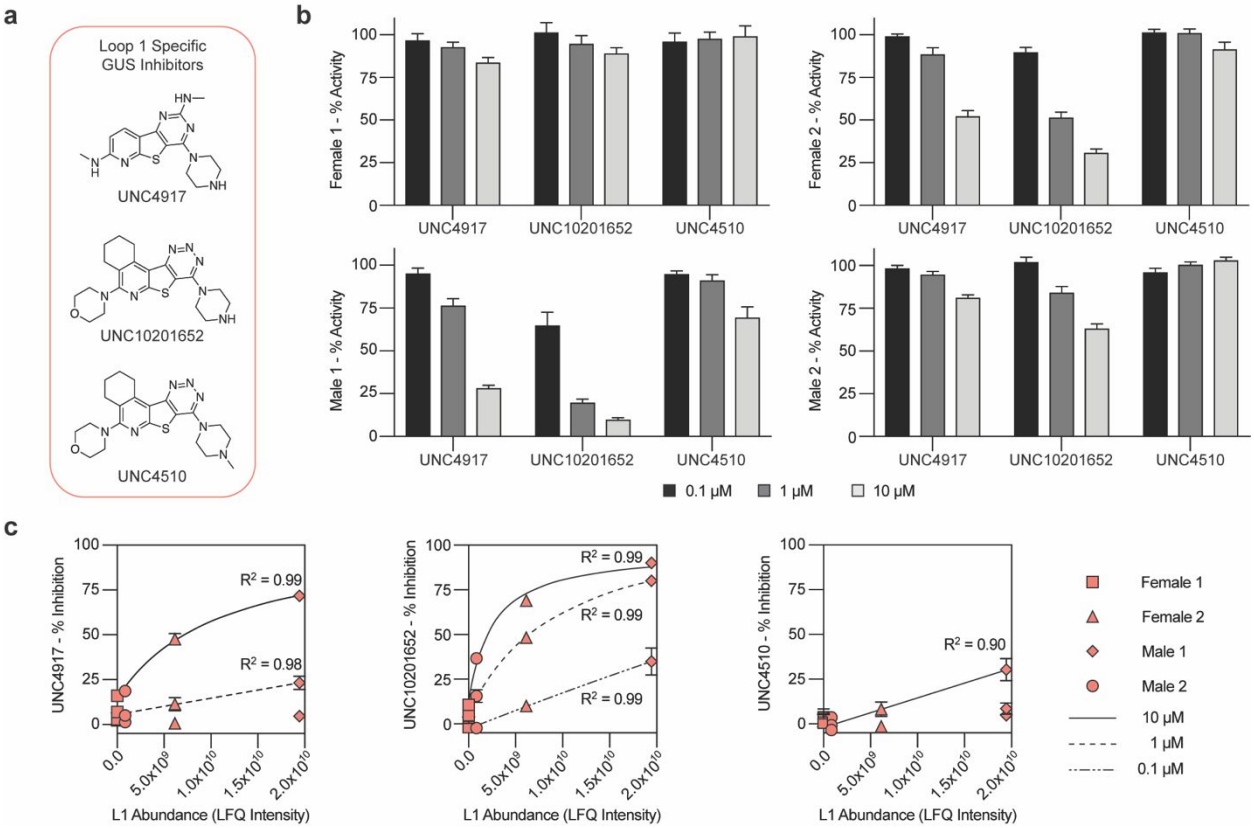
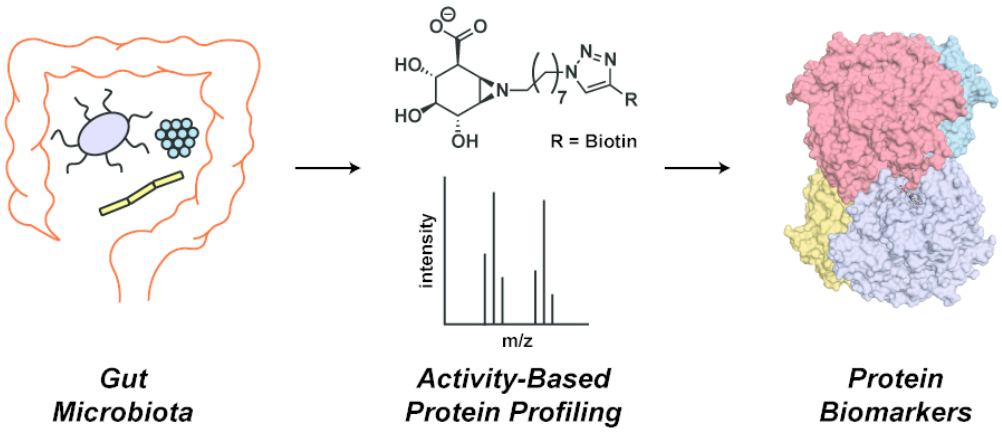


Figure 5. ABPP coupled with *ex vivo* processing data explains differential propensities for GUS inhibition. (**a**) Structures of L1 GUS inhibitors, UNC4917, UNC10201652, and the poor inhibitor, UNC4510 (negative control). (**b**) Inhibition of SN-38 reactivation in human fecal samples by selective bacterial GUS inhibitors. All percent activity values shown are mean values ± standard deviation using *N*=3 biological replicates. (**c**) Correlation analysis between L1 GUS abundance and inhibition data for each GUS inhibitor.


HO
HO
OH
N
N
N
7
R
R = Biotin
intensity
m/z
Gut Microbiota
Activity-Based Protein Profiling
Protein Biomarkers