SUPPLEMENTARY INFORMATION

Three structurally and functionally distinct β -glucuronidases from the human gut microbe *Bacteroides uniformis*

Samuel J. Pellock,^{1,6} William G. Walton,^{1,6} Kristen A. Biernat,¹ Dariana Torres-Rivera,¹ Benjamin C. Creekmore,¹ Yongmei Xu,² Jian Liu,² Ashutosh Tripathy,³ Lance J. Stewart,⁴ and Matthew R. Redinbo^{1,3,5,*}

From the ¹Department of Chemistry, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA; ²Department of Chemical Biology and Medicinal Chemistry, University of North Carolina at Chapel Hill, NC 27599, USA; ³Departments of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA; ⁴Institute for Protein Design, Department of Biochemistry, University of Washington, Seattle, WA 98195, USA; ⁵Departments of Microbiology and Immunology, and Integrative Program for Biological and Genome Sciences, University of North Carolina at Chapel Hill; Chapel Hill, NC, 27599, USA; ⁶These authors contributed equally

TABLE OF CONTENTS

Figure S1	S-3
Figure S2	S-4
Figure S3	S-5
Figure S4	S-6
Figure S5	S-7
Figure S6	S-8
Figure S7	S-9
Figure S8	S-10
Figure S9	S-11
Figure S10	S-12
Figure S11	S-13
Figure S12	S-14
Figure S13	
Table S1	S-16
Table S2	S-17
Table S3	S-18
Table S4	S-19
Table S5	S-20

	N-term loop		
BuGUS1 BuGUS2 BuGUS3	MKTLLKNSLTFLLMIMPVLAFAQQAPQIMNVSARQTTSLDGQWKTIVDPFENG <u>YYD</u> 56 MEREKNTLPQKACHWMAAVI-ISLEVLPEVHAQRQTQTINDGWKFLKEECTA 51 MKARFILLSLL-FSSLLQAHALIPTDRRQRLTEHMSFIRQDMSSIMEV 48 :: ::::::::::::::::::::::::::::::::::	BuGUS1 BuGUS2 BuGUS3	RLIFLSNLANKARSLDSVRLIGAAMEKEEVQPGVLIVNDPLGELLDIISF 477 RKYKTEAELKPVLERTLALANRLERVIKEEDSTRISTMAFHGSNSVNETGLSKITDIVGW 492 FQTTEKDSTRAFMGELHDLAHRLDPTRSTAIRRCEFCKDIVDVYSP 480 L2 mL2 * *.*: .:::: *:.
BuGUS1 BuGUS2 BuGUS3	YRLKPYDGGYAQDKTYSDKTKLQEYDFETDKLLFVPGDWNTQRPQLYYYEGTVWYRK 113 AADSAFDDSKWTSIHLFHTWNTDAYTEKDYYRGTGWYRR 90 MRPITGAGKPETVPLWQKVTLFHCFNAEDAVDPDWNYYQGFGWYRT 94 :::::::::::::::::::::::::::::::::	BuGUS1 BuGUS2 BuGUS3	NEYVGWYDGDSEKCDRVNWTFDTQKPVFISELGGGALYGHH518 NLYQGWYGGDLTGFEKFLAQQHQNHPTHPMIVSEYGAGSDKRLHS-537 TTWAGWYGRFRNYREMETAGINATTRFLHAEMGGDSHAGRHMEVNKETGRKGIDNFD 538 . : *** : : * *:
BuGUS1 BuGUS2 BuGUS3	HFEYSLQ-PGKRLFLNFGAVNYEAIVWINGKRLGRHIGGFTPFNFEITNLLK 164 QLTLPQGWKEKQIILRLDAAGKSATIYINGKNVGEHAGGYTACSFNITPFLSF 143 MLNIENPYTNGRTCLEFEGAGQKTEVYVYTTRIASHVGGYDEWKADITEAVEAFRRTPLC 154 : **: *.: *** :.	BuGUS1 BuGUS2 BuGUS3	GS EXERFTEEYQEDLYI RHVNMLKRI PGLAGTTFWILKDFRSPRRHVPEIQDDENRK 575 LHPRAFDFS IEYQQKYLEHYLFVLEDT PYIOGGTHWNFI DFS SALRD-ESM-PRINNK 593 IEAADRNGDWSESY IIRLFDWHLKEQETHPRLTGSLFWTFKDFSTPLRPDNPI-PYVNQK 597 ::* :: : * : * : * : * : * : * : * : . * : * :
BuGUS1 BuGUS2 BuGUS3	EGINSLVVKVINKRLPEAVPIVNADWWN FGGIIRFVILIEMPATYIRDYYVQLAK-219 DIPNILAVCVINARQDIAPISGDFIFFGGIYRDWLIAVPNQHFNLINHGSDGL197 RERFGKIPIAIRCINSRDIPHIPSDMSDIPNLYGGLYRYVNLVYTPAVHFEQIRIEATT-213 ** * : *: :**: * * * :	BuGUS1 BuGUS2 BuGUS3	GLVSDK32K#KAFFVLQKWYKELTEAYK603 603 GLVYADRT PKDVYHYQAAWRKD IPVIHIASRDWTDRAGVQQAAPVYLPVKIYTNLSEV 653 GVVQRDGT PKE SYYVPQSYWSS-KEMIHIYGHSWPVRWGKPGEPKE ILVYSNCPEV 652 *:* * :
BuGUS1 BuGUS2 BuGUS3	DKNMEGGVQLEGSDKEQKITLDIPELKVKKEVTT 255 FISTPQVSEEQATLSIRGEV-KNDAPEKATLELIHTIYRPDGTLLQTLKKNIQLKAGETY 256 DEKGKQGSISIDIAFGGLSGKEKKAKEFSLRVFSPEGKEVNSISSELT 261 .:* : *::	BuGUS1 BuGUS2 BuGUS3	603 ELFIDGISLGKQKTENYTATFEVPFSNRNPFLFAQGNYQGKT-VQDGLRINFTPI 707 ELFVNGVSQGRKKRNSQDFPAAGLRNNVPLREGRNKITATG-YDGKLRIDDEI 704
BuGUS1 BuGUS2 BuGUS3	DANGYASFLIKSKPIINTPENFKLYAVNLASETDKVSDEIGFRTIRTEGI-KI 307 AFSNEATFVIKPEINTPETPRLYRVETTIRNRKTKTLLDQSNHYTAFRWFRFDGDEGF 314 EISSYQISLKRPQLWSPHSPALYTCVAELIDNGDTLRTTQHFGFRHFRFEEKGPF 316 :* **:*.* :: :	BuGUS1 BuGUS2 BuGUS3	603 PACLDANNIKGLELAVINVESQCFFTSDESQLTWLPDQPYAAGSWGYIGEKEGTAQTEIQN 767 QQEYQTRTWGEPSRILLTQTAQ 726
BuGUS1 BuGUS2 BuGUS3	LLNEKEIFCRGISIHEETPYYSGRAYSKDHAHTLLSNAKELGCNFVRLAHYPHNEEMVRE 367 FLNGKPYKLRGICRHQDQKD-IGPALTDEMHRRDFLIMKEMGANFIRISHYPQDDALLEM 373 YLNEERLLIRGTHRHEDHAG-VGAALTEENMARTEMQQIKEMGANFIRIGHYQQSGIILRL 375 **.: ** *:: ** *:: * *:::::::::::::::::	BuGUS1 BuGUS2 BuGUS3	603 TADGPLFQTLRNEIEGYRFDAPQGVYEIELLFTDIFRRNAGIAYQLDRNGQ@ENRESTFG 827 DAETILVQAELIDENGIRCLDACQFIEFGCTDSEALLRNQGTAQGSRRIQAANGR-A-S 783
BuGUS1 BuGUS2 BuGUS3	AERYGFLUWSE IPVYMT IHWENKDTYQNAEQQLCDMIARDKNRCNI I IWS IANETPHSKT 427 CDKLGMIAWEE IPI IDI VP-NTPGYGENCERNIREMIRGHYNHPSI I WGYANEI LLWTQ 432 CDELGLIWWEE IEWCRGGL-GGE SYKNQARRMLINNIEGHNNPSVI IWGLGNEDWPGD 434 *:*.*.*** L2 mL2	BuGUS1 BuGUS2 BuGUS3	603 ISINGEVVEESLSPCKESGYFRALRKKYYIINDKEYIDIRFHSTSGTCFLNGIKLRNIY 886 IRVNKQHAFVVVSAS 814

Figure S1. Multiple sequence alignment of *Bu***GUS-1**, *Bu***GUS-2**, and *Bu***GUS-3**. Regions of interest discussed in the text are bolded and highlighted in various colors. Unique active site residues highlighted in yellow, N-terminal loop and catalytic glutamates in red, NxK motif in green, Loop 2 (L2) in magenta, mini-loop 2 (mL2) in blue.



Figure S2. pH screen of PNPG hydrolysis for *Bu***GUS-1**, *Bu***GUS-2**, and *Bu***GUS-3**. Error bars represent SD of n = 3 biological replicates.



Figure S3. Structural alignments of *Bu*GUS-1 (cyan), *Bu*GUS-2 (magenta), and *Bu*GUS-3 (blue) with previously elucidated structure of *E. coli* GUS (green, *Ec*GUS, PDB: 3LPF).



Figure S4. Size-exclusion chromatography multi-angle light scattering analysis (SEC-MALS) of *Bu*GUS-1, *Bu*GUS-2, and *Bu*GUS-3 confirms oligomeric states predicted from crystal structures.



Figure S5. Comparison of CBMs of *Bu***GUS-2**, *Bu***GUS-3**, and *Bf***GUS.** Conserved core fold of GUS shown in grey with DUF-1 and CBM 57 or DUF2 (for *Bu***GUS-2** and *Bu***GUS-3**) highlighted in cyan and magenta, respectively.



Figure S6. Circular dichroism wavelength scan and melting temperature for WT and D341A D367A calcium binding mutant of *Bu*GUS-2.

SRS013476.24854-T1-C	EGGFAAFRFDITDLVKPESENVIAVKVNNAPTDYIAPITNQGDFTKMGGIYRD	167
SRS018575.34501-T1-C	IGGYTFFCFDITPYIREGENQLVVCVDNSYDSEIPPLSADFTFFGGIYRD	165
SRS016203.125956-T1-C	KGGYSAFTFDVTDYVHTGRNLVAVSVDNSYNPDIAPLSADFTFFGGLYRD	173
SRS012902.13258-T1-C	KGGYSAFTFDVTDYVHAGRNLIAVSVDNSHNPDIAPLSADFTFFGGIYRD	173
SRS015217.11766-T1-C	KGGYTRFCFDITSHLRYGQENLFAIYVNNVYNPNIPPLSADFTFFGGIYRD	203
SRS050422.73127-T1-C	KGGYTRFCFDITSKLRYGQENLFAICVNNTYNPNIPPLSADFTFFGGIYRD	168
SRS048870.28850-T1-C	KGGYSAFIVDITDKVRNDVSNLLMVHVNNO-NNOIPPLSGDFTIFGGIYRD	166
SRS020869.18805-T1-C	TGGYTACTLDITPFCSLNTPNSLAVCVDNS-RODIPPISGDFTFFGGIYRD	176
SRS045713.76924-T1-C	NGGYTACTEDITPECSEDAPNSLAIHVDNA-RODIPPISADETEEGIYRD	173
SBS015190,42469-T1-C	NGGYTACTEDITSLCSENSPNSLAIRVDNA-RODIPPISADETEEGGIYRD	156
SBS056259.118375-T1-C	NGGYTACTEDITSLCSENSPNSLAIRVDNA-RODIPPISADETEEGGIYRD	156
SRS011405_14712=T1=C	AGGYSAFILDVTGLIRFNNETETTVDNA-RRDITPLW-ADETFWGGIYRD	163
SRS019787_27227-T1-C	AGGYSAFILDVTGLIRKENETETIVDNA-RRDITPLW-ADFTFWGGIYRD	163
SRS049959 38447-T1-C	ACCVSSFTEDITORINTNDAROCEKSENTIEITVONS-RODVTDIM-ADETEWCCIVED	164
SPS011520 64867-T1-C	ACCYSAFTED TOPYI OKNNTTETTVDNS_DED TTOTSADETEWCCTYDD	159
SRS017433 24596-T1-C	LCCVSAFTIDITIL RKENVIFITVDNS-REDITIDS ADDITIO	150
SPS010601 88766-T1-C	VCCVTACIUDITEVID	1/2
SK5019601.00/00-11-C	VGGIIACIVDIILIIKKLNDILIIVDNG-KKDIIFIS-ADFIFWGGIIKD	142
	¥.	
SRS013476.24854-T1-C	EVTQHFGVRSYSIDYDSGFYLNGRSYDLHGVNYHQDSYENGWAMSDSQRERDYGMMMDMG	345
SRS018575.34501-T1-C	EVSNTFGIRTCSFSAEKGFELNGKAVKLLGTNRHQCHSGMGNALKDEMHVRDIELLHEMG	342
SRS016203.125956-T1-C	CVVNPLGIREYHFDAEKGFFLNGKYRKLIGTSRHQDYKGMGNALRDEMHIRDIQLSKDMG	351
SRS012902.13258-T1-C	CVVNPLGIREYRFDAEKGFFLNGKYRKLIGTSRHQDYKGMGNALRDEMHIRDVQLSKDMG	351
SRS015217.11766-T1-C	EVVNPLGLRWFKFDSEKGFFLNGKGRKLIGTARHODYFOKGNALRDELHIODVLLLKEMG	382
SRS050422.73127-T1-C	EVVNPLGLRWFEFDSEKGFFLNGKWRKLIGTARHODYFOKGNALRDELHVODVLMLKEMG	347
SRS048870.28850-T1-C	EMRVPLGVRWFSMDAQEGFKLNGKPMKLIGACRHQDQMPMGIALSDEMHRRDMQLLKDMG	344
SRS020869.18805-T1-C	QSNHYTGFRWFSFDGERGFSLNGKPYKLRGICRHODOKPIGVALTDEMHRRDIKLMKEMG	355
SRS045713.76924-T1-C	HSNHHTGFRWFSFDGKTGFSLNGKPYKLRGICRHQDQKPIGVALTDEMHRRDMKLMKEMG	352
SRS015190.42469-T1-C	RSNHHTGFRWFGFDGKTGFSLNGKPYKLRGICRHQDQKPIGVALTDEMHRRDMMLMKEMG	335
SRS056259.118375-T1-C	RSNHHTGFRWFGFDGKTGFSLNGKPYKLRGICRHODOKPIGVALTDEMHRRDMMLMKEMG	335
SRS011405.14712-T1-C	EISNKIGFRWFSFDADKGFSLNGKPYKLRGVNRHQDQAPVGVAIDDEVNRRDIRQIKEIG	342
SRS019787.27227-T1-C	EVSNKVGFRWFSFDSGKGFSLNGKPYKLRGVNRHQDQAPVGVAIDDEVNRRDIRQMKEIG	342
SRS049959.38447-T1-C	EKNHKVGFRWFSFDGEKGFCLNGKTYKLRGFNRHODOAPVGVALPDEAHRRDIKLMKELG	343
SRS011529.64867-T1-C	RSTHKTGFRWFSFDGEKGFSLNGKPYKLRGVNRHODOAPVGVALDDEAHRRDIRLIKEMG	336
SRS017433.24596-T1-C	EOIHKTGFRWFSFDGNKGFCLNGKPYKLRGVNRHODOAPVGVALDDEAHRRDIRLMKEFG	338
SRS019601.88766-T1-C	EKNHKVGFRWFTFDGSKGFFLNGKSYKLRGLNRHODOAPAGVALDDEAHRRDIFLMKELG	321
	** ** ** ** * * * * * * * * * * * *	
	L.	
000013476 34054 mt c	OBSTRESS UNATE OUR PERFORMENT OF OFFICE PROPERTY IN INTERSTICK TO BE STADOOD OF THE OUT DOOT BET	405
SR5013476.24854-11-C	CTAVKLAHIQHDQIEIDLCDRLGLCVWTEVGLVNTISADNDTLIIADGFAGNIKQQLTEL	405
SRS018575.34501-11-C	GNFLRIAHYPQ D EMVLAACNRLGIVTSVEIPVINAITMNQNFSDNCVEMMKEM	395
SRS016203.125956-T1-C	SNFLRVAHYPQ D PVVMQMCDKLGLLTSVEIPIVNAITQSRAFMDNCVEQATEM	404
SRS012902.13258-T1-C	SNFLRVAHYPQDPVVMQMCDKLGLLTSVEIPVVNAITQSKAFMDNCVEQVTEM	404
SRS015217.11766-T1-C	GNFLRVSHYPQDPVIMEMCDKLGIVTSVEIPVVNAVTETEEFLQNSVEMAKEM	435
SRS050422.73127-T1-C	GNFLRVSHYPQDPVIMEMCDKLGIVTSVEIPVVNAVTETEEFLTNSVEMAKEM	400
SRS048870.28850-T1-C	ANFVRLAHYPQDAVLRACDELGMLVWEEIPVVDLIALGDEFRTNATSALREM	397
SRS020869.18805-T1-C	ANFIRISHYPQDDALLEMCDKLGMLAWEEIPIIDIVPDTPGYTENCERNLREM	408
SRS045713.76924-T1-C	ANFIRISHYPQDDALLEMCDKLGMLAWEEIPIIDIVPDTPGYAENCENNLREM	405
SRS015190.42469-T1-C	ANFIRISHYPQDDALLEMCDKLGMLAWEEIPIIDIVPDTPGYAENCENNLREM	388
SRS056259.118375-T1-C	ANFIRISHYPQDDALLEMCDKLGMLAWEEIPIIDIVPDTPGYAENCESNLREM	388
SRS011405.14712-T1-C	CNFIRISHYPQDDALLDACDELGLLAWEEIPIIDIVPDTPGYDDHCEMNLVEM	395
SRS019787.27227-T1-C	CNFIRISHYPQDALLDACDELGLLAWEEIPIVNMVPDTPGYEDNCETNLVEM	395
SRS049959.38447-T1-C	SNYIRISHYPQDALLDACDELGLLAWEEIPIIDLVPDTPHYADNCERNLREM	396
SRS011529.64867-T1-C	CNFIRIAHYPQDDALVEACDEMGLLAWEEIPIINIVPDTPGYDDNCETNLTEM	389
SRS017433.24596-T1-C	CNFIRISHFPQDDAILEMCDELGLLVWEEIPVINTVPDTPGYDDNCEYNLREM	391
SRS019601.88766-T1-C	CNFIRISHFPQDDAILEMCDELGLLAWEEIPIINIVPNTPGYDDNCEYNLREM	374
	* * * *	

Figure S7. Sequence alignment of putative calcium binding GUS enzymes from the Human Microbiome Project stool sample database. Related to Figure 5.



Figure S8. Modeling of sugar monosaccharides of glucuronate (GlcA), galacturonate (GalA), mannuronate (ManA), and iduronate (IdoA) in the active sites of *Bu*GUS-1, *Bu*GUS-2, and *Bu*GUS-3. Related to Figure 6.



Figure S9. *Bu*GUS-1, but not *Bu*GUS-2 or *Bu*GUS-3, can process both glucuronides and galacturonides. Related to Table 1. *a*, chemical structures of *p*NP-GlcA and *p*NP-GalA. *b*, progress curves for *Bu*GUS-1, *Bu*GUS-2, and *Bu*GUS-3 with *p*NP-GalA. *c*, kinetic parameters of *p*NP-GlcA and *p*NP-GlA hydrolysis by *Bu*GUS-1. *d*, active site of BuGUS-1 with GalA manually docked and potentially important aspartic acid residue highlighted in green. This aspartic acid likely undergoes a rotamer shift, shown in cyan, to accommodate the GalA sugar.



Figure S10. Structural overlay of *Bu*GUS-1 WT (cyan) and BuGUS-1 Δloop (grey) reveals absence of N-terminal loop necessary for efficient processing of SN-38-G. Related to Figure 8.



Supplementary Figure 11.

Figure S11. Structural analysis of BMSP GUS and docking studies of SN-38-G in the active sites of *Bu*GUS-2 and *Bu*GUS-3. Related to Figure 1, Figure 6. *a*, tetrameric biological assembly of BMSP GUS. *b*, active site of BMSP GUS with NxK motif in green and catalytic glutamates in deep salmon. Manually docked SN-38-G in active site of *c*, *Bu*GUS-2 and *d*, *Bu*GUS-3 based on the PTG-bound structure of *Bu*GUS-1.



Figure S12. Chemical structures of Inhibitor 1 and UNC10201652. Related to Table 3.



Figure S13. Predicted structures by Rosetta-based modeling of unique tertiary and quaternary structures discovered by crystallography. (A) Overlay of observed and Rosetta-predicted tertiary structure for *Bu*GUS-1, *Bu*GUS-3, and BMSP GUS. (B) Observed quaternary structures for *Bu*GUS-1, *Bu*GUS-3, and BMSP. (C) Rosetta prediction of quaternary structure for *Bu*GUS-1, *Bu*GUS-3, and BMSP.

Protein	BuGUS-1	BuGUS-3	BMSP	BuGUS-1 + G-1,5-L	BuGUS-2 + G-1,5-L	BuGUS-1 + GlcA	BuGUS-1 ∆loop	BuGUS-2 ∆Ca ²⁺	BuGUS-1 + TPG
Resolution (Å)	29.6 - 2.20	29.48 – 2.35	29.43 - 2.65	29.93-1.90	29.48-2.50	29.98-1.80	29.34-2.00	29.55-3.00	29.30-2.40
Resolution in highest shell (Å)	2.25 – 2.20	2.39 – 2.35	2.71 – 2.65	1.93 -1.90	2.56-2.50	1.83-1.80	2.03-2.00	3.12-3.00	2.49-2.40
Space group	C2	C2	I4 ₁	C2	P212121	C2	P12 ₁ 1	P212121	P3221
Unit cell dimensions a,b,c (Å);	125.8,130.2, 94.2;	209.7, 68.2, 167.4;	213.5, 213.5, 112.6;	126.1, 130.2, 94.5;	74.5, 141.9, 180.8;	134.2, 133.9, 163.4;	76.7,199.2, 84.6;	75.0,142.0, 181.2;	252.2, 252.2, 107.3
α,β,γ (°)	90, 100.3, 90	90, 122.5, 90	90, 90, 90	90, 100.2, 90	90, 90, 90	90,100.3, 90	90,102.8, 90	90, 90, 90	90, 90, 120
Total reflections (F>0)	517245	568681	305377	797212	446802	1761568	995295	525999	3085366
Unique reflections	75592	83443	73245	114979	67106	261960	163378	39506	152030
Multiplicity	6.8 (7.0)	6.8 (6.9)	4.2 (4.2)	6.9 (6.9)	6.7 (6.8)	6.7 (6.9)	6.1 (5.8)	13.3 (14.0)	20.3 (21.0)
Completeness, %(highest shell)	99.9 (100)	99.9 (99.9)	99.7 (100)	97.7 (96.3)	99.9 (100.0)	100 (100)	98.4 (97.7)	99.8 (100)	99.8 (99.7)
Mean l/sigma (I) (highest shell)	18.6 (4.7)	21.0 (4.9)	16.1 (2.6)	16.6 (4.4)	15.9 (4.1)	11.8 (4.3)	12.1 (3.7)	20.9 (3.9)	23.4 (7.0)
Wilson B-factor (Å ²)	21.4	31.9	61.8	14.9	36.5	8.3	25.6	67.7	33.1
R _{merge} (highest shell)	0.078 (0.387)	0.073 (0.445)	0.060 (0.50)	0.081 (0.351)	0.084 (0.438)	0.126 (0.471)	0.118 (0.496)	0.120 (0.69)	0.111 (0.511)
Rwork	0.1494	0.1755	0.2069	0.1377	0.1703	0.18	0.176	0.1670	0.1739
Rfree	0.1841	0.2224	0.2827	0.1694	0.2277	0.229	0.23	0.2409	0.2467
Number of molecules in asymmetric unit	2	2	4	2	2	4	4	2	6
Number of water molecules in asymmetric unit	1075	815	110	1549	535	4087	1808	0	1561
rms bonds (Å)	0.003	0.009	0.008	0.012	0.004	0.006	0.007	0.009	0.007
rms angles (°)	0.631	1.24	1.000	1.072	0.669	0.874	0.849	1.030	0.94
Ramachandran favored (%)	96.53	95.34	92.85	96.79	96.23	96.71	96.30	92.91	92.23
Ramachandran outliers (%)	0.00	0.39	0.19	0.00	0.24	0.00	0.09	0.18	0.20
Clash score	3.04	6.87	13	1.96	2.12	7.05	5	8.33	16.84
Average B-factor (Å ²)	28.3	35.1	67.0	21.2	41.0	14.6	30.0	57.9	37.9
RCSB ID	6D1N	6D1P	6D8K	6D41	6D5O	6D6W	6D89	6D8G	6D7F

Table S1. X-ray data collection and refinement statistics.

HMGC ID	Length	Kingdom	Phylum	Class	Order	Family	Genus	Species	NCBI ID
SRS011529.64867	855	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	Paraprevotella		
SRS049959.38447	872	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	Prevotella	copri	CUO96334.1
SRS018575.34501	885	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides	clarus	WP_009122035.1
SRS050422.73127	891	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides	finegoldii	OKZ23312.1
SRS020869.18805	891	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides	intestinalis	WP_007661425.1
SRS011405.14712	861	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides	intestinalis	CCY84748.1
SRS015217.11766	926	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides	ovatus	WP_004304963.1
SRS019601.88766	840	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides	vulgatus	WP_011965330.1
SRS012902.13258	891	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides		
SRS016203.125956	891	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides		
SRS045713.76924	888	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides		
SRS048870.28850	883	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides		
SRS015190.42469	871	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides		
SRS056259.118375	871	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides		
SRS019787.27227	861	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides		
SRS017433.24596	858	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides		
SRS013476.24854	693	Bacteria	Firmicutes	Clostridia	Clostridiales	Eubacteriaceae	Eubacterium		

 Table S2. Bioinformatic analysis of the HMP GUS enzymes identified 18 GUS enzymes with calcium binding sites.

HMGC ID	Length	Kingdom	Phylum	Class	Order	Family	Genus	Species	NCBI Annotation	NCBI ID
SRS052697.50102	1427	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides	eggerthii	GH2 protein	WP_017140867.1
SRS022137.24481	689	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides	MSP	GH2 protein	WP_007840523.1
SRS055982.32560	875	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides	salyersiae	GUS/b-gal	CUM88360.1
SRS024388.29889	796	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides	sartorii	GH2 protein	WP_016277419.1
SRS011405.4923	1427	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides	thetaiotaomicron	b-gal	WP_081030062.1
SRS014613.5859	1118	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides	thetaiotaomicron	GUS/b-gal	WP_080973891.1
SRS018133.31535	814	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides	uniformis	GH2 protein	WP_044467710.1
SRS017433.149978	1429	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides		b-gal	
SRS011586.60073	826	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Porphyromonadaceae	Parabacteroides	merdae	GH2 protein	WP_005639106.1
SRS014923.34218	642	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiaceae	Faecalibacterium	CAG:74	b-gal	OLA21514.1
SRS023914.87176	647	Bacteria	Firmicutes	Clostridia	Clostridiales					

Table S3. Bioinformatic analysis of the HMP identified 11 GUS enzymes containing the variable active site tryptophan observed in BuGUS-3.

HMGC ID	Length	Kingdom	Phylum	Class	Order	Family	Genus	Species	NCBI Annotation	NCBI ID
SRS020869.249001	613	Bacteria	Bacteroidetes							
SRS015190.3291	603	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Bacteroidaceae	Bacteroides	uniformis	GUS/b-gal	CU058783
SRS017307.68457	600	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Rikenellaceae	Alistipes	senagalensis	GUS	WP_019151945
SRS015264.145530	600	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Barnesiellaceae	Coprobacter	secundus	GUS	WP_021929483
SRS049995.49420	609	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidiales	Rikenellaceae	Alistipes	senagalensis	GUS	WP_019149660
SRS017701.130646	318	Bacteria	Bacteroidetes							

Table S4. Bioinformatics analysis of the HMP identified 6 GUS enzymes containing an N-terminal loop like that observed in BuGUS-1.

PrimerName	Sequence
BuGUS-1 Fwd	TACTTCCAATCCAATGCGTTCGCACAGCAGGCACCGCAGATTATGAAC
BuGUS-1 Rev	TTATCCACTTCCAATGCGCTATTTATAGGCTTCAGTCAGT
BuGUS-3 Fwd	TACTTCCAATCCAATGCGCTGATCCCGACCGATCGTAAACAGCGTC
BuGUS-3 Rev	TTATCCACTTCCAATGCGCTACTGGCCGCTCACAGAGATCAGCGC
BMSP Fwd	TACTTCCAATCCAATGCGCAGAACCTGATTACCAATGTTTACGGTCGTGACATCC
BMSP Rev	TTATCCACTTCCAATGCGCTATTCACCGAATTTGGTTTTGTAGTACTCACGCATCAGGTA
BuGUS-1 ∆N-termloop Fwd	GTTGATCCGTTCGAAAACGGTCGTGGCATGAAAGTTTACCAGGATAAAACCTATAGTGAT
<i>Bu</i> GUS-1 ∆N-termloop Rev	ATCACTATAGGTTTTATCCTGGTAAACTTTCATGCCACGACCGTTTTCGAACGGATCAAC
BuGUS-2 D341A Fwd	GTCCGGCTCTGACGGCTGAAATGCATCGTCG
BuGUS-2 D341A Rev	CGACGATGCATTTCAGCCGTCAGAGCCGGAC
BuGUS-2 D367A Fwd	CACTACCCGCAGGCTGACGCTCTGCTG
BuGUS-2 D367A Rev	CAGCAGAGCGTCAGCCTGCGGGTAGTG
BuGUS-2 N591A/K593A Fwd	CTGCGTGATGAAAGCATGCCGCGTATTGCCAACGCAGGCCTGGTTTAC
BuGUS-2 N591A/K593A Rev	CGTAAACCAGGCCTGCGTTGGCAATACGCGGCATGCTTTCATCACGCAG

Table S5. Primer sequences used in this study.