Supporting Information

Little et al. 10.1073/pnas.1716241115

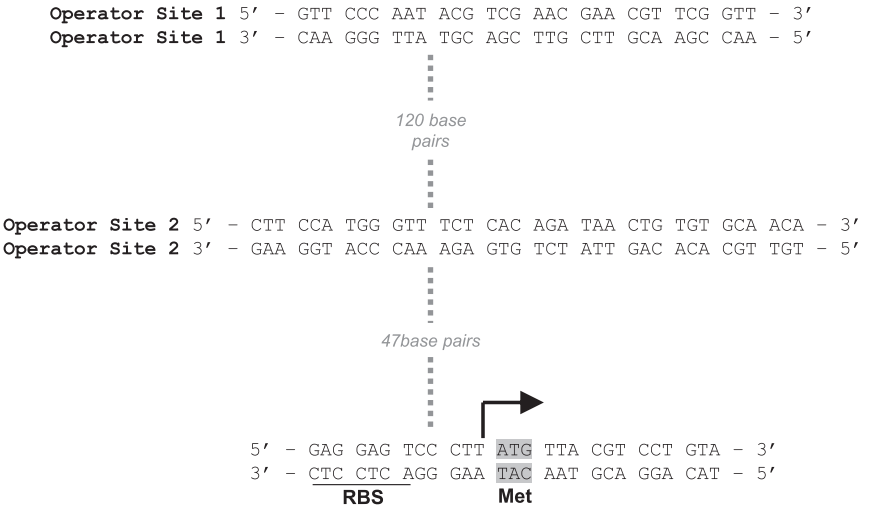

Fig. S1. Oligonucleotide sequences of predicted operator sites of the *E. coli* GUS operon. The ribosome binding site is underlined (RBS), and the *gus*A gene start site is shaded (Met).

	PNP-glucuronide (PNPG)	PNP-glucopyranoside (PNP-gluco)	Glucuronic Acid (GlcA)
	PNPG	PNP-gluco	GlcA
EcGusR	0.21 *(0.07)*	*NB*	*NB*
SeGusR	2.7 *(0.4)*	*NB*	*NB*
EcUxuR	*NB*	*NB*	*NB*
CpFadR	*NB*	*NB*	*NB*
SaGntR	*NB*	*NB*	*NB*

Fig. S2. Binding of the indicated ligands by EcGusR, SeGusR, and related repressor proteins as measured by isothermal titration calorimetry (ITC). EcGusR and SeGusR both bound to PNP-glucuronide (PNPG) but failed to bind PNP-glucopyranoside (PNP-Gluco) or glucuronic acid (GlcA). Error values are reported in parentheses and italicized, and represent SDs. NB, no binding.

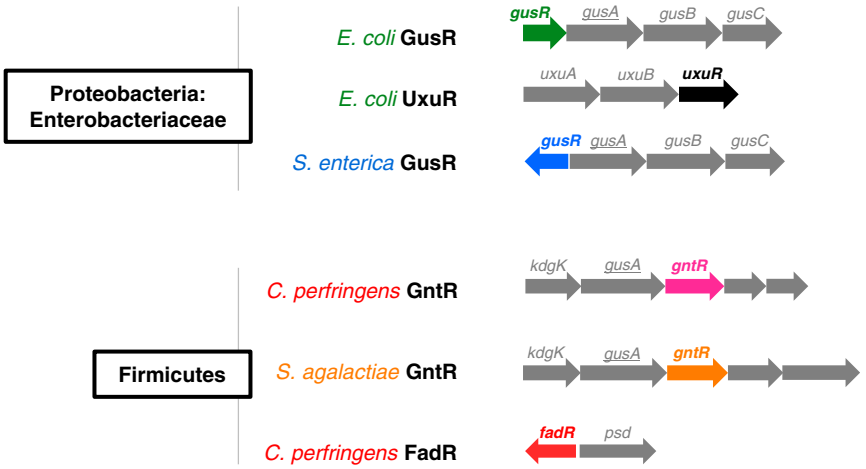

Fig. S3. Schematics showing the organization of genes near the GUS operons in the Enterobacteriaceae and putative GusR homologs in Firmicutes. The GUS operon and Uxu operon are separated by more than 2.5 million bp on the *E. coli* chromosome.

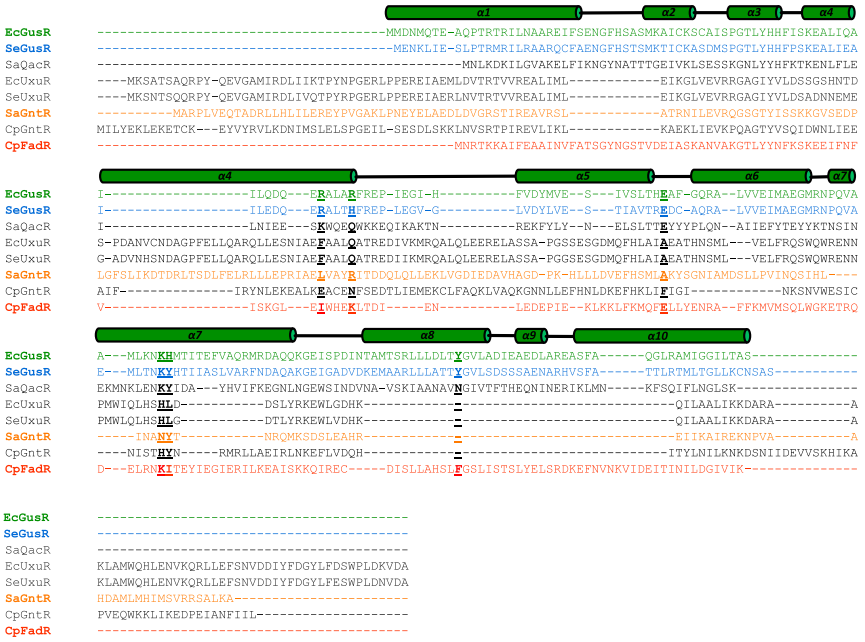

Fig. S4. Sequence alignment of EcGusR, SeGusR, and related repressor proteins, with the secondary structure of EcGusR shown and residues important for glucuronide recognition highlighted in bold and underlined. The names of the proteins whose structures are resolved here are indicated in bold on the left.

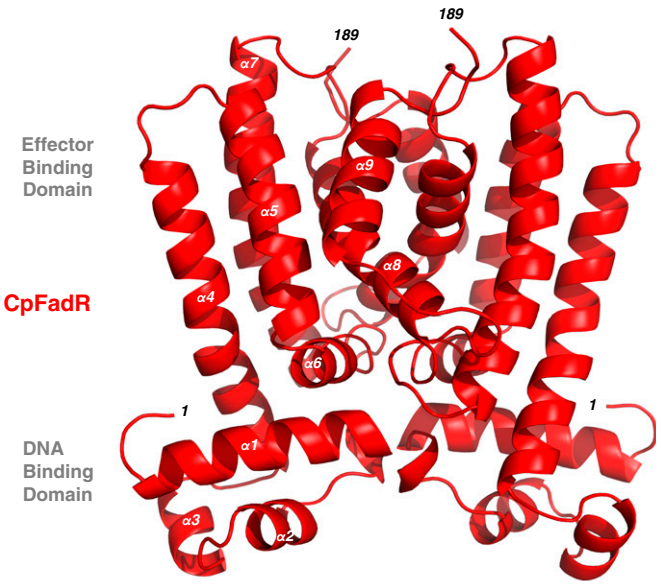

Fig. S5. Cartoon representation of the 1.8-Å crystal structure of the CpFadR homodimer (PDB ID code 6AZH).

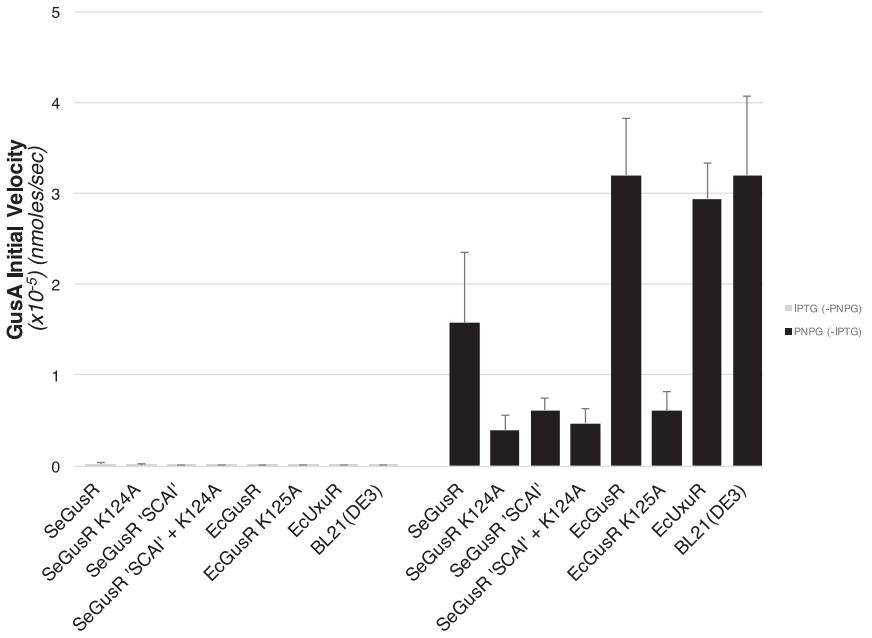

Fig. S6. Controls for *E. coli* cell-based GUS activity assay. A bar graph of initial velocities of GUS in cultured *E. coli* BL21 cells harboring either an empty expression plasmid or expression plasmids containing genes for the indicated proteins, both wild-type and variants. (*Left*) Only the IPTG inducer was added with no PNPG, and nearly negligible GUS activity was observed. (*Right*) Only the GusR effector PNPG was added but no IPTG, and more modest levels of GUS activity were observed. All values are the mean of three biological triplicates, run in triplicate for each experiment, and error bars are SEMs.

Table S1. X-ray crystallographic statistics

Protein	EcGusR	SeGusR	SaGntR	CpFadR
Space group	$P22_12_1$	$P2_12_12$	C2	P 2_1
Unit cell				
a, b, c, Å	51.0, 106.7, 135.3	52.7, 55.2, 60.5	109.6, 105.2, 45.9	39.9, 93.9, 55.1
α, β, γ, °	90, 90, 90	90, 90, 90	90, 112.44, 90	90.00, 96.07, 90.00
Resolution range, Å	29.2–2.09 (2.17–2.09)	27.6–2.06 (2.13–2.06)	27.8–1.91 (1.98–1.91)	29.1–1.78 (1.78–1.75)
Wavelength, Å	1.0332	0.97939	0.97935	0.97935
Unique reflections	43,892 (3,995)	11,394 (1,087)	36,686 (3,545)	39,564 (2,181)
Multiplicity	4.3 (3.0)	2.0 (1.8)	4.3 (2.4)	6.8 (7.1)
Completeness, %	0.95 (0.92)	0.99 (0.96)	0.98 (0.96)	0.98 (0.98)
I/σ	13.5 (2.0)	10.9 (2.2)	16.3 (1.6)	23.7(3.5)
Wilson B factor	33.2	21.9	27.0	19.1
R_{work}	0.189 (0.271)	0.183 (0.275)	0.215 (0.291)	0.196 (0.232)
R_{free}	0.233 (0.322)	0.222 (0.322)	0.250 (0.325)	0.225 (0.285)
Molecules, a.u.	4	1	2	2
Waters, a.u.	408	114	283	438
Residues, a.u.	648	192	404	373
Average B factor	40.1	23.7	30.3	27.0
Rmsd, Å	0.010	0.010	0.010	0.010
Rmsd, °	0.830	0.510	1.06	0.740
Ramachandran favored, %	100	98	99	99
Ramachandran outliers, %	0	0	0	0
PDB ID code	6AYI	6AYH	6AZ6	6AZH

Values for the highest resolution shell are indicated in parentheses.

Table S2. Impact of GusR effector-binding pocket mutations on PNPG binding as measured by isothermal titration calorimetry

Ec/Se mutations	EcGusR, μM	SeGusR, μM
WT	0.21 (*0.07*)	2.7 (*0.5*)
H126Y/Y125H	0.94 (*0.04*)	2.6 (*0.04*)
R73H/H72R	0.52 (*0.01*)	2.1 (*0.06*)
M87L/L86M	1.7 (*0.06*)	2.8 (*0.1*)
H126Y+R73H /Y125H+H72R	N/A	1.4 (0.1)
H126Y+R73H+M87A/Y125H +H72R+L86M	N/A	1.6 (0.01)

Error values are reported in parentheses and italicized, and represent SDs. N/A, not applicable, because the corresponding mutants were not created.

Table S3. Summary of GUS operon-containing organisms

Organism	Strain	Complete GUS Operon	Note
Escherichia coli	K12 MG1655 (*NC_000913.3*)	Yes	
Escherichia coli	O157:H7 Sakai (*NC_002695.1*)	No	No *gus*A and *gus*C
Salmonella enterica	BCW_1559 (*NZ_MXOA01000002.1*)	Yes	
Shigella boydii	600690 600690_30 (*NZ_LPTQ01000029*)	No	Truncated *gus*A and *gus*B
Shigella flexneri	301 (*NC_004337.2*)	No	Truncated *gus*B/no *gus*C
Shigella sonnei	FC1772 (*NZ_NGWB010000048.1*)	Yes	
Shigella dysenteriae	225-75 gss22575.contig.38 (*AKNG01000039.1*)	Yes	
Buttiauxella agrestis	MCE Contig_3 (*NZ_JPRU01000002.1*)	Yes	
Yersinia enterocolitica	ERL053484 (*NZ_CWGL01000004.1*)	Yes	
Yersinia kristensenii	FCF326 (*NZ_CPYI01000006.1*)	Yes	
Erwinia totetana	DAPP-PG 735 scf_2378_174.1 (*NZ_KB372804.1*)	Yes	
Raoultella planticola	CHB Contig_16 (*NZ_JPRG01000013.1*)	Yes	
Klebsiella pneumoniae	KP38731 (*NZ_CP014297*)	Yes	
Klebsiella oxytoca	2880STGY5682666 (*FKYV01000004.1*)	No	Truncated *gus*A and *gus*B/no *gus*C

Organisms are labeled by genera and species, and the NCBI accession codes for each strain are listed (in parentheses and italicized).