

UCSF

UC San Francisco Electronic Theses and Dissertations

Title

Reducing the Rainbow: The intestinal environment shapes gut bacterial azo dye depletion with impacts on drug absorption

Permalink

<https://escholarship.org/uc/item/1v61d95v>

Author

Pieper, Lindsey M

Publication Date

2023

Peer reviewed|Thesis/dissertation

Reducing the Rainbow: The intestinal environment shapes gut bacterial azo dye depletion with impacts on drug absorption

by
Lindsey Michelle Pieper

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY

in

Biomedical Sciences

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

DocuSigned by:

Dr. Sue Lynch

1A38FCA893D34C5...

Dr. Sue Lynch

Chair

DocuSigned by:

Carol Gross

DocuSigned by:

Peter J. Turnbaugh

8C4E22B44AE0441...

Carol Gross

Peter J. Turnbaugh

Committee Members

Copyright 2023

by

Lindsey Michelle Pieper

Dedication

To my parents, Cindy and Toby Pieper. Thank you for everything.

To Jessica Blanton. Without your mentorship, I would never have gotten here.

Acknowledgements

First, to my wonderful advisor, Peter Turnbaugh. Peter, thank you for supporting me, believing in me, encouraging me, and guiding me through this process. Not only have you created an environment that fosters fantastic science and learning, you foster a lab that is welcoming, accepting and believes in the importance of work-life balance making us better scientists. It's quite an impressive feat. Thank you.

To my thesis committee, Carol Gross and Sue Lynch. Thank you for being excellent role models and advisors. Not only did you provide excellent scientific guidance and feedback, you picked up on my fears and anxieties as I moved through my grad school journey and pushed me forward with just the right amount of reassurance and firm nudging into the deep end.

To the Turnbaugh lab. I don't know if I can adequately articulate my appreciation. From the first day of my rotation, I felt warmly welcomed and accepted in this lab. I appreciate your enthusiasm, dedication, and compassion in both research and friendship. To the post-docs, thank you all for answering my questions. Especially thank you to Peter S for being my main postdoc mentor teaching me about anaerobic culturing and *E. coli* genetics. Thank you to Maggie, Kathy, Cecilia, Ben, Jordan, and Christine for not only answering my lab questions, but also my life questions. To the grad students. I wouldn't have joined this lab if it weren't for Qi Yan, Paola, Veronica, and Than. I'm so glad we all became friends, not just colleagues.

To my grad school friend, Camillia, Denise, Tara, and Casey, and pre-grad school friends, Courtney, and Manoshi. I wouldn't have gotten through a pandemic, personal life ups and downs, and a PhD without your friendships.

To my previous mentors. Thank you for taking a chance on me and letting me learn in your labs and work on your projects. Thank you, Eric Allen, for answering my email when I was an undergrad, thank you for confirming my love of microbiology and continuing enthusiasm and support of my research and career. Thank you to Jessica Blanton. Words cannot describe how you opened the world of research and microbiology to me. Not only did you encourage me scientifically, you made me believe in myself and my ability to do whatever I set my mind to. I still hear your voice when I do certain lab tasks.

To my therapist. This PhD has occurred at the same time a lot of other personal life has occurred. Thank you to her for her support and keeping me grounded so I could still work on these projects.

To the community that raised me, especially the women. Thank you for paving the way for me to do this PhD. I know not that long ago even getting to the doorsteps of a lab, let alone being in a lab and grad program cohort made of mostly women would have been an entirely different endeavor. I appreciate how far we've come even if there is still room to grow.

To my girls, Myla, the dog who joined me in college, and Duchess, the horse that joined me in grad school. Thank you for unconditional love and reminding me what's truly important in life. I've appreciated that you both get me outside and enjoy the bay area. Mostly, thank you for the hugs. I don't think I'd have enjoyed this time as much without you.

Lastly, to my family. Mom and Dad, you have literally been there from day one. You've supported my every interest from ballet to horses to science. You've encouraged me and Lance to pursue educational endeavors that interest and inspired us to study and work in fields we're genuinely interested in. I don't think I can adequately express how grateful I am for helping me get here and continuing to support me. To Lance, I couldn't ask for a better brother. Thanks for always listening. To Victor, I know you didn't experience this entire roller coaster, but you've been a great addition to the ride. Thank you for supporting me, loving me, and accepting me.

Contributions

Chapter 2 contains work previously published in:

- Zou L*, Spanogiannopoulos P*, Pieper LM, Chien H-C, Cai W, Khuri N, Pottel J, Vora B, Ni Z, Tsakalozou E, Zhang W, Shoichet BK, Giacomini KM, Turnbaugh PJ. 2020. Bacterial Metabolism rescues the inhibition of intestinal drug absorption by food and drug additives. *Proc Natl Acad Sci U S A* 117:16009 – 16018. *Equal contribution.

Chapter 3 contains work currently in press:

- Pieper LM, Spanogiannopoulos P, Volk RF, Miller CJ, Wright AT, Turnbaugh PJ. 2023. The global anaerobic metabolism regulator *fnr* is necessary for the degradation of food dyes and drugs by *Escherichia coli*. *mBio*, in press.

Reducing the Rainbow: The intestinal environment shapes gut bacterial azo dye depletion with impacts on drug absorption

Lindsey Michelle Pieper

Abstract

Dyes are a ubiquitous part of the human diet and, also added to pills as excipients. While traditionally considered inert, emerging literature suggests that these compounds can have far-reaching impacts on host health and disease. Moreover, many dyes contain azo-bonds ($R-N=N-R'$), which are subject to gut bacterial metabolism. However, the mechanisms through which gut bacteria impact azo dyes, their sensitivity to environmental factors, and their downstream consequences for host health and disease remain poorly understood. In this thesis, I demonstrate that azo dyes inhibit drug absorption by blocking intestinal OATP2B1 uptake transporters; gut bacterial metabolism of these azo dyes rescues this effect. Using the model gut bacterium *Escherichia coli*, I discovered that the canonical azoreductase gene is unnecessary for azoreduction due to an alternative pathway in which hydrogen sulfide produced from L-Cysteine depletes these dyes. Further, I dissect the signaling pathways that control this process, revealing that oxygen sensing via the fumarate nitrate reductase regulator (*fnr*) alters L-Cysteine import via the small regulatory RNA, *fnrS*. Consistent with these findings, the gut microbiota impacts hydrogen sulfide levels and the pharmacokinetics of the azo bonded anti-inflammatory drug sulfasalazine. Taken together, these results demonstrate the critical role of diet and environmental factors like oxygen in shaping the metabolic activity of human gut bacteria and add to the growing literature demonstrating that the gut

microbiome controls both drug metabolism and absorption. Our results provide a strong foundation to dissect how the microbiome impacts azo dyes and other compounds in the context of the complex microbial, host, and dietary pressures found in the gastrointestinal tract.

Table of Contents

Chapter 1: Introduction	1
1.1 Introduction.....	1
1.2 References.....	4
Chapter 2: Bacterial metabolism rescues the inhibition of intestinal drug absorption by food and drug additives	8
2.1 Introduction.....	8
2.2 Results.....	9
2.3 Discussion.....	26
2.4 Supplementary Figures.....	30
2.5 Supplementary tables.....	37
2.6 Materials and Methods.....	51
2.6 References.....	62
Chapter 3: The global anaerobic metabolism regulator <i>fnr</i> is necessary for the degradation of food dyes and drugs by <i>Escherichia coli</i>	70
3.1 Introduction.....	72
3.2 Results.....	73
3.2.1 <i>E. coli</i> AzoR is dispensable for azo dye depletion under Anaerobic conditions.....	73
3.2.2 FNR is necessary for the depletion of azo dyes by <i>E. coli</i>	77
3.2.3 Paired transcriptomics and proteomics reveals additional genes necessary for azo dye depletion.....	79
3.2.4 L-Cysteine metabolism is required for depletion of azo dyes.....	84

3.2.5 FNR indirectly regulates the <i>cyu</i> genes.....	89
3.3 Discussion.....	92
3.4 Supplementary Figures.....	97
3.5 Supplementary Tables.....	104
3.6 Materials and Methods.....	130
3.7 References.....	142
Chapter 4: Gut Microbiome Environment Impacts Sulfasalazine	
pharmacokinetics and Sulfide Levels.....	154
4.1 Introduction.....	154
4.2 Results.....	156
4.2.1 The gut microbiota impacts the pharmacokinetics of sulfasalazine.....	156
4.2.2 Sulfasalazine pharmacokinetics are not altered by the addition of <i>E. coli</i> or dietary interventions.....	162
4.2.3 Gastrointestinal sulfide levels are robust to diet and <i>E. coli</i> strain engraftment status.....	166
4.3 Discussion.....	171
4.4 Materials and Methods.....	174
4.5 References.....	178
Chapter 5: Conclusion.....	182
5.1 Conclusion.....	182
5.2 References.....	184

List of Figures

2.1	Multiple oral excipients inhibit OATP2B1-mediated transport.....	12
2.2	Excipient inhibitors of OATP2B1 uptake have distinct physicochemical and functional properties compared to non-inhibitors.....	13
2.3	Oral administration of the OATP2B1 transport excipient inhibitor FD&C Red No. 40 reduces fexofenadine bioavailability in mice.....	15
2.4	Human gut bacteria metabolize azo dye excipients.....	18
2.5	Microbial metabolism of azo dye excipients rescues OATP2B1 uptake inhibition.....	21
2.6	Interactions between the gut microbiota, azo dye excipients, and drug absorption in conventionally-raised (CONV-R) and gnotobiotic mice.....	24
3.1	<i>E. coli</i> AzoR is sufficient but dispensable for azoreductase activity.....	75
3.2	The <i>fnr</i> regulator is necessary for azoreductase activity in <i>E. coli</i>	78
3.3	Multi-omic analysis of the <i>fnr</i> regulon reveals a complex regulon and the importance of L-Cysteine metabolism.....	82
3.4	The <i>fnr</i> regulator is necessary for the metabolism of L-Cysteine to hydrogen sulfide, enabling the reduction of azo bonds.....	85
3.5	L-Cysteine import and metabolism genes impact bacterially produced sulfide and FD&C Red No. 40 depletion.....	87
3.6	FNR indirectly regulates sulfide production and FD&C Red No. 40 depletion predominantly via small RNA <i>fnrS</i>	90
3.7	Working model describing how the host environment shapes gut bacterial azo reduction.....	92

4.1	Sulfasalazine can be activated by enzymatic and non-enzymatic mechanisms.....	154
4.2	Gut microbiota presence impacts sulfasalazine activation and pharmacokinetics.....	157
4.3	Gut microbiota presence impacts sulfasalazine metabolite, sulphapyridine pharmacokinetics.....	158
4.4	Conventionalization leads to increased sulfide levels.....	159
4.5	Streptomycin interferes with gut bacterial azo reduction.....	161
4.6	Sulfasalazine is significantly depleted in <i>wt E. coli</i> compared to Δ <i>fnr E. coli</i>	162
4.7	Combinations of <i>E. coli</i> engraftment and/or diet does not impact sulfasalazine pharmacokinetics.....	165
4.8	Combinations of <i>E. coli</i> engraftment and/or diet does not impact sulphapyridine pharmacokinetics.....	166
4.9	Murine GI sulfide levels are robust to dietary intervention.....	169

List of Supplementary Figures

S2.1	Study design overview.....	30
S2.2	Dose-dependent inhibition of OATP2B1-mediated estrone sulfate (ES) uptake by FD&C Red No. 40 and FD&C Yellow No. 6.....	31
S2.3	Comparison of physicochemical properties between OATP2B1 excipient inhibitors and non-inhibitors.....	32
S2.4	Isolation of azo dye excipient metabolizing human gut bacterial isolates.....	33
S2.5	Metabolism of azo dye excipients by human gut bacterial isolates.....	34
S2.6	Measurement of residual FD&C Red No. 40 in conditioned media samples by liquid chromatography-high resolution mass spectrometry (LC-HRMS).....	35
S2.7	Gut bacterial metabolism rescues the inhibitory effects of intestinal drug transport by excipient azo dyes.....	36
S3.1	Validation of bacterial strains.....	97
S3.2	<i>E. coli</i> AzoR is sufficient to deplete azo dyes.....	98
S3.3	<i>fnr</i> is necessary for the depletion of multiple azo dyes.....	99
S3.4	Dye depletion corresponds to the growth phase of <i>wt E. coli</i>	100
S3.5	<i>ydeN</i> is not required for azo dye depletion.....	101
S3.6	Key transcripts and proteins in Δ <i>fnr E. coli</i> relative to wild-type controls.....	102
S3.7	Regulation of <i>cyuR</i> and its operon.....	103

List of Supplementary Tables

S2.1	Validation and inhibitory potencies of 24 OATP2B1 excipient inhibitors.....	37
S2.2	Aggregation analysis of OATP2B1 excipient inhibitors.....	39
S2.3	Physicochemical features of OATP2B1 excipient inhibitors versus non-inhibitors.....	40
S2.4	Pharmacokinetic parameters of fexofenadine with FD&C Red No. 40 in conventional and germ-free P-glycoprotein deficient mice.....	41
S2.5	Human gut bacterial isolates capable of clearing azo dyes	43
S2.6	Human gut bacterial isolate genome sequencing statistics.....	45
S2.7	Inhibitor constant (K_i) values for excipient azo dyes and their metabolites.....	47
S2.8	Estimated intestinal concentrations of excipients based on maximum allowable amount and estimated daily intake.....	50
S3.1	Keio collection strains screened for decreased azoreductase activity.....	104
S3.2	RNA sequencing metadata and sequencing statistics.....	113
S3.3	Peptide metadata and peptide counts	114
S3.4	Multivariate analysis of differentially abundant E. coli proteins and transcripts.....	115
S3.5	Genes with consistent changes in both RNA-seq and proteomics data	116
S3.6	Genes that were consistently affected at the transcript and protein level and significant in our knockout screen.....	125
S3.6	Strains used in Chapter 3.....	126
S3.8	Primers used in Chapter 3.....	128

Chapter 1: Introduction

1.1 Introduction

Trillions of microorganisms reside in the gastrointestinal (GI) tract of humans and animals. An impressive body of research demonstrates the importance of the complex interplay between diet, the gut microbiome, digestion, metabolism, and intestinal uptake of dietary macronutrients. However, the interactions between the microbiome and xenobiotics are less well understood. Xenobiotics are defined as any compound foreign, or unexpected, in an organism or ecological setting (1). One such group of xenobiotics commonly encountered by the human gut microbiome is azo bonded ($R-N=N-R'$) compounds (2, 3). Interestingly, previous studies have demonstrated the ability of intestinal microbes to reduce azo dyes at this bond (4). Most notably, these are commonly found in food and drug dyes such as the prolific, FD&C Red No. 40 (5, 6). Microbiota also encounter azo-bonded drugs, such as sulfasalazine (SSZ) and prontosil, which rely on microbial azo metabolism for activation (7–9).

While azo dyes have largely been considered safe, inert additions to food and drugs, their interactions with and on the gut microbiome are unknown. Additionally, while azo-bonded are known substrates of bacteria and bacterial enzymes (4, 6, 10), how this process occurs in a complex human gut microbiome community is unknown. In this thesis, I explore how gut microbiota interact with commonly consumed azo dyes (**Chapter 2**). I demonstrate the ability of microbial depletion of azo dyes and profile a sampling of azo bond-reducing bacterial species. With collaborators, we show that this depletion is sufficient to clear a negative effect of azo dyes, their inhibition of OATP2B1 uptake transporters.

Azo reductase activity is considered a core metabolic function of all human gut microbiotas (4, 11–14), the enzymatic and non-enzymatic mechanisms responsible for this activity and their sensitivity to environmental factors remain poorly understood. Azoreductase (AzoR) enzymes, found in a number of species across a variety of phyla, are generally cited and recognized as the main contributors to azo depletion activity (4, 10, 15, 16). While these enzymes have been well studied *in vitro* and provide important mechanistic insight (17–24), these studies do not address cellular and genetic mechanisms impacting azo depletion activity nor do they address the impact interactions of the larger GI microenvironment. In **Chapter 3**, I seek to address these larger picture questions using the model organism *Escherichia coli*. Surprisingly, I found that the canonical azoR gene is unnecessary for azo depletion. Additionally, we found this activity is broadly regulated by oxygen-sensing dual-transcriptional regulator, FNR, and more specifically through the small regulatory RNA *fnrS*. Interestingly, these regulate L-Cysteine (L-Cys) transport and metabolism which releases hydrogen sulfide (H₂S), which can react with and deplete azo dyes.

While a majority of my project focused on the most commonly consumed azo dye as a model chemical, pharmaceuticals with azo bonds are also commonly consumed. As previously mentioned, SSZ is a rationally designed pro-drug commonly prescribed for both inflammatory bowel disease (IBD) (2, 25) and Rheumatoid Arthritis (RA) (26). SSZ was designed such that it remains in pro-drug form until reaching the microbially rich colon where the azo bond would be microbially metabolized and the active inflammatory drug moiety, 5-aminosalicylic acid (5-ASA) would be released at the site of intestinal inflammation for IBD treatment (27). While this design relies on microbial activation of azo

bonds, little is known about how microbial community composition and microbiome environment impact SSZ activation and the pharmacokinetics of its metabolites. Additionally, the abiotic activation of azo compounds by H₂S was recently described (28, 29) and how sulfide levels correlate with and impact SSZ activation is unknown. In **Chapter 4** I explore how microbiota and diet impact SSZ pharmacokinetics and H₂S levels. I find that the presence of a microbiota and antibiotic treatment elicit SSZ PK changes and sulfide levels. However, dietary intervention did not impact sulfide levels or induce SSZ PK differences.

Taken together, my thesis provides a contextual understanding of how the gut microbiome interacts with and impacts one class of xenobiotic compounds. Our results highlight the importance of probing the environmental context and regulation of key microbial metabolisms at a mechanistic level. Moreover, dissecting these complexities could help explain the inter-individual variability in azo drug efficacy and provide an understanding of bacterial prodrug activation that could inform future studies and therapies.

References

1. Collins SL, Patterson AD. 2020. The gut microbiome: an orchestrator of xenobiotic metabolism. *Acta Pharm Sin B* 10:19–32.
2. Peppercorn MA. 1984. Sulfasalazine: Pharmacology, Clinical Use, Toxicity, and Related New Drug Development. *Ann Intern Med* 101:377.
3. Bastaki M, Farrell T, Bhusari S, Bi X, Scrafford C. 2017. Estimated daily intake and safety of FD&C food-colour additives in the US population. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess* 34:891–904.
4. Chung KT, Stevens SE Jr, Cerniglia CE. 1992. The reduction of azo dyes by the intestinal microflora. *Crit Rev Microbiol* 18:175–190.
5. Lehmkuhler A, Miller MD, Bradman A, Castorina R, Chen M-A, Xie T, Mitchell AE. 2022. Levels of FD&C certified food dyes in foods commonly consumed by children. *J Food Compos Anal* 112:104649.
6. Chung KT. 2016. Azo dyes and human health: A review. *Journal of Environmental Science and Health - Part C Environmental Carcinogenesis and Ecotoxicology Reviews* <https://doi.org/10.1080/10590501.2016.1236602>.
7. Spanogiannopoulos P, Bess EN, Carmody RN, Turnbaugh PJ. 2016. The microbial pharmacists within us: a metagenomic view of xenobiotic metabolism. *Nat Rev Microbiol* 14:273–287.

8. Gingell R, Bridges JW, Williams RT. 1971. The role of the gut flora in the metabolism of prontosil and neoprontosil in the rat. *Xenobiotica* 1:143–156.
9. Peppercorn MA, Goldman P. 1972. The role of intestinal bacteria in the metabolism of salicylazosulfapyridine. *J Pharmacol Exp Ther* 181:555–562.
10. Misal SA, Gawai KR. 2018. Azoreductase: a key player of xenobiotic metabolism. *Bioresources and Bioprocessing* <https://doi.org/10.1186/s40643-018-0206-8>.
11. Chung KT, Fulk GE, Egan M. 1978. Reduction of azo dyes by intestinal anaerobes. *Appl Environ Microbiol* 35:558–562.
12. Wang R-F, Chen H, Paine DD, Cerniglia CE. 2004. Microarray method to monitor 40 intestinal bacterial species in the study of azo dye reduction. *Biosens Bioelectron* 20:699–705.
13. Zahran SA, Ali-Tammam M, Hashem AM, Aziz RK, Ali AE. 2019. Azoreductase activity of dye-decolorizing bacteria isolated from the human gut microbiota. *Sci Rep* 9:5508.
14. Ryan A. 2017. Azoreductases in drug metabolism. *Br J Pharmacol* 174:2161–2173.
15. Sandhya S. 2010. Biodegradation of Azo Dyes Under Anaerobic Condition: Role of Azoreductase, p. 39–57. *In* Atacag Erkurt, H (ed.), *Biodegradation of Azo Dyes*. Springer Berlin Heidelberg, Berlin, Heidelberg.

16. Sarkar S, Banerjee A, Halder U, Biswas R, Bandopadhyay R. 2017. Degradation of Synthetic Azo Dyes of Textile Industry: a Sustainable Approach Using Microbial Enzymes. *Water Conservation Science and Engineering* 2:121–131.
17. Nakanishi M, Yatome C, Ishida N, Kitade Y. 2001. Putative ACP phosphodiesterase gene (*acpD*) encodes an azoreductase. *J Biol Chem* 276:46394–46399.
18. Ito K, Nakanishi M, Lee W-C, Sasaki H, Zenno S, Saigo K, Kitade Y, Tanokura M. 2005. Crystallization and preliminary X-ray analysis of AzoR (azoreductase) from *Escherichia coli*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 61:399–402.
19. Morrison JM, Wright CM, John GH. 2012. Identification, Isolation and characterization of a novel azoreductase from *Clostridium perfringens*. *Anaerobe* 18:229–234.
20. Crescente V, Holland SM, Kashyap S, Polycarpou E, Sim E, Ryan A. 2016. Identification of novel members of the bacterial azoreductase family in *Pseudomonas aeruginosa*. *Biochem J* 473:549–558.
21. Chen H, Wang R-F, Cerniglia CE. 2004. Molecular cloning, overexpression, purification, and characterization of an aerobic FMN-dependent azoreductase from *Enterococcus faecalis*. *Protein Expr Purif* 34:302–310.
22. Liu Z-J, Chen H, Shaw N, Hopper SL, Chen L, Chen S, Cerniglia CE, Wang B-C. 2007. Crystal structure of an aerobic FMN-dependent azoreductase (AzoA) from *Enterococcus faecalis*. *Arch Biochem Biophys* 463:68–77.

23. Nachiyar CV, Rajakumar GS. 2005. Purification and characterization of an oxygen insensitive azoreductase from *Pseudomonas aeruginosa*. *Enzyme Microb Technol* 36:503–509.
24. Wang C-J, Hagemeyer C, Rahman N, Lowe E, Noble M, Coughtrie M, Sim E, Westwood I. 2007. Molecular Cloning, Characterisation and Ligand-bound Structure of an Azoreductase from *Pseudomonas aeruginosa*. *J Mol Biol* 373:1213–1228.
25. Das KM. 1989. Sulfasalazine therapy in inflammatory bowel disease. *Gastroenterol Clin North Am* 18:1–20.
26. Plosker GL, Croom KF. 2005. Sulfasalazine: a review of its use in the management of rheumatoid arthritis. *Drugs* 65:1825–1849.
27. Sousa T, Yadav V, Zann V, Borde A, Abrahamsson B, Basit AW. 2014. On the Colonic Bacterial Metabolism of Azo-Bonded Prodrug of 5-Aminosalicylic Acid. *J Pharm Sci* 103:3171–3175.
28. Wolfson SJ, Hitchings R, Peregrina K, Cohen Z, Khan S, Yilmaz T, Malena M, Goluch ED, Augenlicht L, Kelly L. 2022. Bacterial hydrogen sulfide drives cryptic redox chemistry in gut microbial communities. *Nat Metab* 4:1260–1270.
29. Dai Q, Zhang S, Liu H, Huang J, Li L. 2020. Sulfide-mediated azo dye degradation and microbial community analysis in a single-chamber air cathode microbial fuel cell. *Bioelectrochemistry* 131:107349.

Chapter 2: Bacterial metabolism rescues the inhibition of intestinal drug absorption by food and drug additives

2.1 Introduction

One of the most notable aspects of life in the developed world is the routine exposure to chemicals through processed foods, pharmaceuticals, cosmetics, and the environment. Recent studies suggest that many of these small molecules have deleterious effects on human health (1), although the mechanisms through which they impact human pathophysiology remains poorly described. Even less is known about the reciprocal interactions between these compounds and the trillions of microorganisms that reside within the gastrointestinal tract (the gut microbiome) (2–4), despite emerging evidence that food and cosmetic additives can influence host-microbiome interactions (1, 5).

In this study, we focus on the bioactivity of pharmaceutical excipients — defined as substances other than the active pharmaceutical ingredients that are intentionally included in an approved drug delivery system or a finished drug product. On average, excipients make up 90% of a drug formulation and play crucial roles including assisting in stability, bioavailability, manufacturing, and patient acceptability (6). Excipients in oral drug products are present in intestinal fluid together with active ingredients; however, their impact on drug disposition is poorly understood. More broadly, many of the excipients used in pharmaceuticals are also widespread and abundant in processed foods and cosmetics and therefore provides a routine source of chemical exposures for most individuals in the developed world.

2.2 Results

We aimed to systematically screen the interactions between common pharmaceutical excipients and the intestinal transporter Organic Anion Transporting Polypeptide 2B1 (OATP2B1) by developing an *in vitro* assay to identify OATP2B1-inhibiting excipients. OATP2B1 (SLCO2B1) is localized to the apical membrane of the intestine and mediates the absorption of many oral prescription drugs, including the statins fluvastatin and rosuvastatin (7), the antihistamine fexofenadine (8, 9), and the β -adrenoceptor blocker atenolol (10). Compounds that inhibit OATP2B1 transport may potentially reduce the absorption of many drugs. To identify excipients that interact with OATP2B1, we analyzed a comprehensive collection of molecular excipients identified using the Centers of Excellence in Regulatory Science and Innovation (CERSI) Excipients Browser (11) and assembled a library of unique molecular excipients to screen for OATP2B1 transport inhibition. We excluded excipients that are no longer used, poorly soluble (solubility of <1 mM in H₂O, DMSO, or ethanol), commercially unavailable, formulated for delivery by inhalation, or alternative salt forms of other excipients (*e.g.* sodium acetate was tested while calcium/magnesium acetate were not). This collection totaled 136 oral excipients spanning multiple functional classes including dyes, buffering agents, antimicrobial agents, and flavoring agents (**Figure 2.1a, Supplemental Figure S2.1**). This excipient collection was not exclusive to drugs, as 65% of our collection are also found in food products.

We developed an *in vitro* assay to identify potential inhibitors of OATP2B1-mediated uptake. The human *OATP2B1* cDNA was cloned into the mammalian expression vector pcDNA™ 5/FRT and transfected into human embryonic kidney (HEK)

Flp-In™ cells to generate a stable OATP2B1-overexpressing cell line. We selected HEK cells for our screen because of their routine and robust use for functional transporter expression assays (12) as well as possessing a low innate rate of drug uptake. The fluorescent molecule 4',5'-dibromofluorescein (DBF) was used as a substrate of OATP2B1 uptake for screening. We determined a $K_m=4.7 \mu\text{M}$ for OATP2B1-mediated uptake of DBF (**Figure 2.1b**), which is in agreement with the literature (13). The role of OATP2B1 in DBF uptake was confirmed by comparing our overexpressing cell line to a control cell line (HEK Flp-In™ cells transfected with empty vector pcDNA™), revealing a significant (3-fold) increase in uptake under OATP2B1-overexpressing conditions (**Figure 2.1c**). DBF uptake was significantly inhibited by the endogenous OATP2B1 substrate estrone sulfate (**Figure 2.1c**). The statistical effect size (Z'-factor) was 0.79, indicating excellent assay quality (14). Together, these results demonstrate that our assay reliably and reproducibly measures the rate of OATP2B1-dependent DBF uptake.

Next, we used our validated assay to screen the full 136-member excipient library for inhibition of DBF uptake. Considering that the amount of excipients in an oral drug product can outweigh the active ingredient by 10- to 100-fold, we selected an excipient screening concentration of 200 μM (10-fold above the standard 20 μM used in drug screens (15)). We classified hits as excipients that inhibited OATP2B1 uptake by >50% in order to focus on a manageable number of excipients for follow-up dose-response studies. Using this criterion, we identified 24 inhibitors in our primary screen (17.6% of the library; **Figure 2.1d**). Next, we validated each of the 24 excipients by performing dose-response inhibition assays. This analysis validated 100% (24/24) of the inhibitors identified in our primary screen (**Figure 2.1e and Supplemental Table S2.1**), providing

additional support for our >50% inhibition threshold and suggesting that additional excipient inhibitors remain to be discovered. Six excipients were potent inhibitors of OATP2B1-mediated uptake with an inhibition constant (K_i) $\leq 1.0 \mu\text{M}$ (**Supplemental Table S2.1**). We further validated two excipient uptake inhibitors, FD&C Red No. 40 and FD&C Yellow No. 6, using the endogenous OATP2B1 substrate estrone sulfate (**Supplemental Figure S2.2**). An aggregation test on the 24 excipient inhibitors demonstrated that the K_i values of 9 excipients were >10 times their aggregation concentration, suggesting specific inhibition of OATP2B1 (**Supplemental Table S2.2**). These results indicate that the ability of oral food and drug excipients to inhibit drug uptake is far broader than previously appreciated (16).

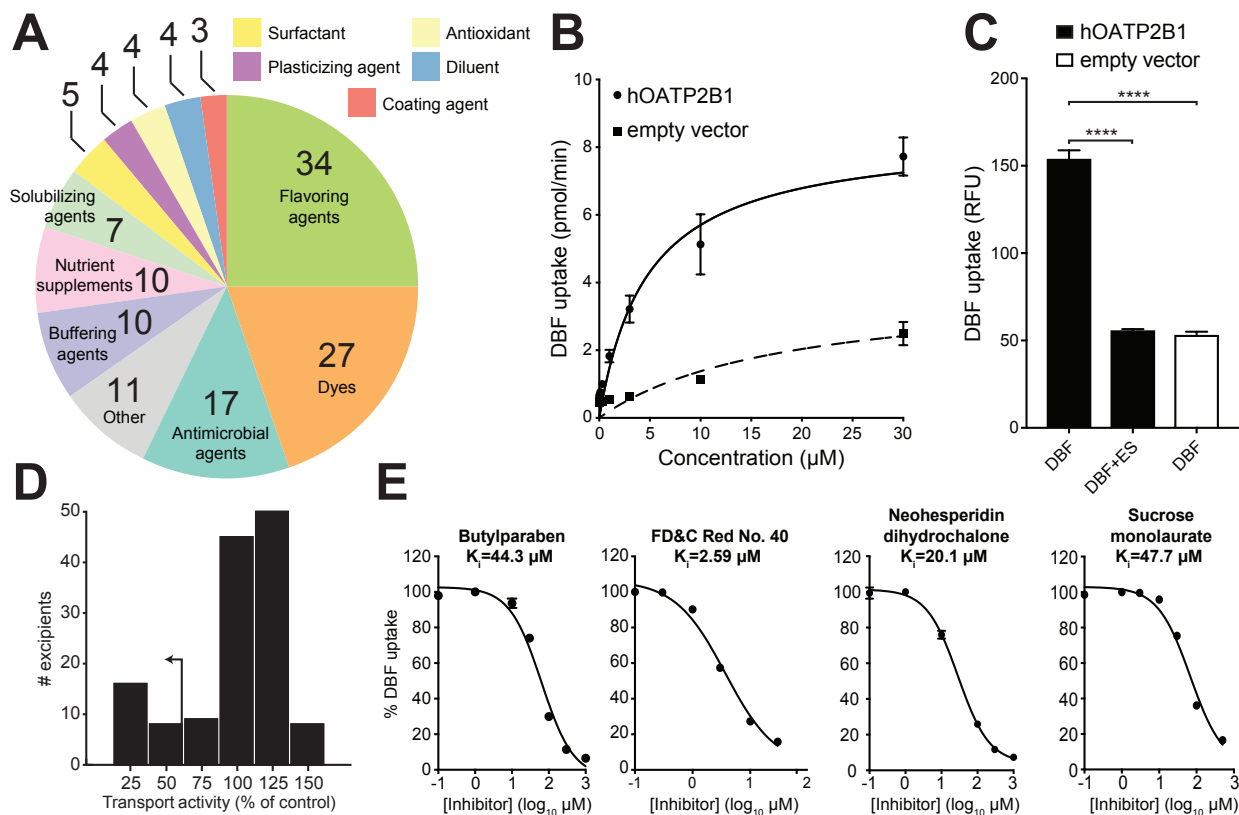


Figure 2.1. Multiple oral excipients inhibit OATP2B1-mediated transport. (A) Functional categories of the 136 oral molecular excipients included in our screen. (B) Kinetics of human OATP2B1-mediated 4',5'-dibromofluorescein (DBF) uptake. Human OATP2B1-overexpressing (circles) and empty vector-transfected Human Embryonic Kidney (HEK) cells (squares) were incubated with DBF from 0.01 μM to 30 μM for 2 min. Data points represent the mean \pm SD of DBF uptake from ≥ 3 replicate determinations in three experiments. (C) Human OATP2B1-overexpressing HEK cells (black bars) were incubated in Hank's Balanced Salt Solution (HBSS) uptake buffer containing 2 μM DBF for 3 min with or without 200 μM estrone sulfate (ES). Empty vector-transfected HEK cells (white bar) were assayed as above for background DBF uptake determination. Each column represents the mean \pm SD of DBF uptake from 8 replicate determinations. **** $p < 0.0001$, ANOVA with Dunnett's correction. (D) Histogram showing the inhibition results of screening 136 excipients against OATP2B1 uptake. Our $< 50\%$ transport activity cutoff is indicated by the arrow. (E) Dose-response curves of excipients against OATP2B1 transport. A representative excipient from each functional category is shown. Values represent the mean \pm SD of DBF uptake from three replicate determinations in a single experiment.

The identified OATP2B1 transport inhibitors were chemically and functionally diverse, which included dyes, surfactants, antimicrobials, and flavoring agents (**Figure 2.2a**). Despite this diversity, we were able to identify characteristic signatures of OATP2B1 inhibitors relative to the overall library. Multiple physicochemical properties of excipient inhibitors were distinct relative to non-inhibitors, including increased molecular weight ($p=3.8010^{-08}$, Student's t-test; **Figure 2.2B**) and lipophilicity ($p=7.3010^{-11}$, Student's t-test; **Figure 2.2C, Supplemental Figure S2.3 and Table S2.3**). We also found that OATP2B1 inhibitors were significantly enriched for dyes relative to the overall panel, representing 18/24 (75%) of the inhibitors ($p<0.0001$, Fisher's exact test; **Figure 2.2a and Supplemental Table S2.1**).

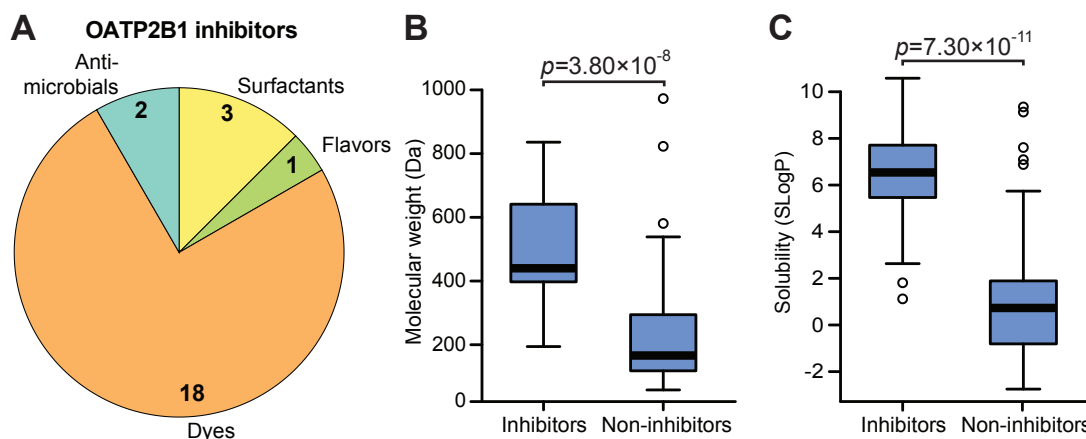


Figure 2.2. Excipient inhibitors of OATP2B1 uptake have distinct physicochemical and functional properties compared to non-inhibitors. (A) Composition of the OATP2B1-inhibiting excipients identified from our screen. Excipient inhibitors were significantly enriched for dyes ($p<0.0001$, Fisher's exact test). Excipient inhibitors display higher molecular weight (B) and solubility (SLogP; Crippen's log of the octanol/water partition coefficient) (C) compared to non-inhibitors. Box plots are shown for data in panels B, C ($n=24$ inhibitors versus 112 non-inhibitors) with the bold line representing the median. Points outside the bars are shown as open circles. p -values represent Student's t-tests.

OATP2B1 plays an important role in fexofenadine absorption (8, 17) and is a target of food-drug interactions (9, 18). We selected the azo dye excipient FD&C Red No. 40 for *in vivo* studies because it is the dye with the highest approved amount in the United States by the Food and Drug Administration (FDA) and widely used in both food and drug products (19). The inhibitory effect of FD&C Red No. 40 on fexofenadine bioavailability was examined in P-glycoprotein (Pgp) deficient (*mdr1a/b*^{-/-}) mice since fexofenadine is also a substrate of Pgp, which reduces absorption of fexofenadine (20). Concomitant administration of 25 mg/kg FD&C Red No. 40 (10 mM estimated intestinal concentration) significantly reduced fexofenadine area under the plasma concentration-time curve (AUC₀₋₃₆₀) (n=9) by 48% compared to administration of the vehicle (n=8 mice/group, *p*=0.0026, ANOVA with Tukey's correction; **Figure 2.3a and c, and Supplemental Table S2.4**). We also observed a trend towards decreased peak plasma concentration (**Figure 2.3d**). In contrast, administration of 2.5 mg/kg FD&C Red No. 40 (1 mM estimated intestinal concentration) (n=4) resulted in fexofenadine levels comparable to vehicle controls (**Figure 2.3a-d**).

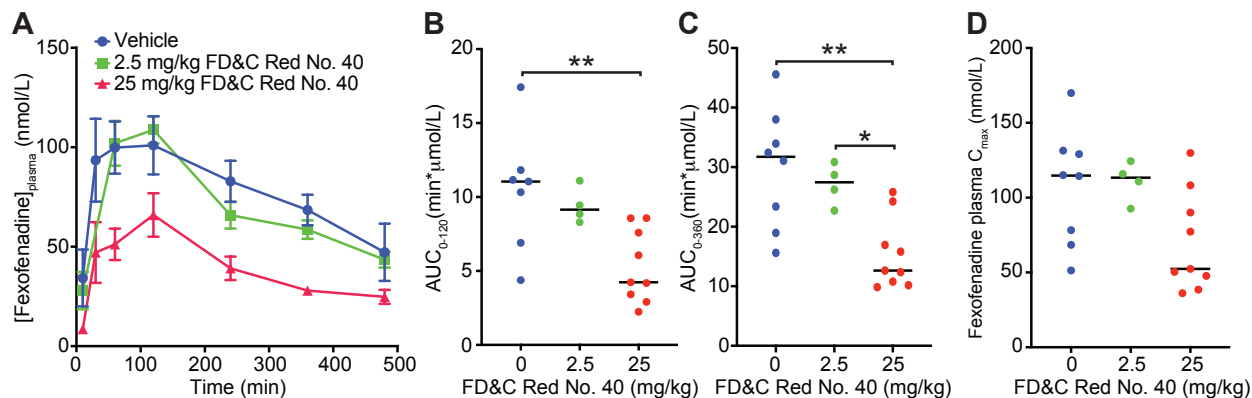


Figure 2.3. Oral administration of the OATP2B1 transport excipient inhibitor FD&C Red No. 40 reduces fexofenadine bioavailability in mice. (A) Plasma concentration-time profiles of fexofenadine after oral administration of 15 mg/kg fexofenadine with saline (blue circles, n=8), 2.5 mg/kg FD&C Red No. 40 (green squares, n=4), or 25 mg/kg FD&C Red No. 40 (red triangles, n=9) to P-glycoprotein (Pgp) deficient (*mdr1a/b*^{-/-}) mice. Each point represents the mean \pm sem. (B-D) Pharmacokinetic parameters calculated using Phoenix WinNonlin v8.1: (B) plasma fexofenadine AUC_{0-120min}; (C) plasma fexofenadine AUC_{0-360min}; and (D) maximum plasma concentration (C_{max}). Each point represents the data from an individual mouse administered with 15 mg/kg fexofenadine plus saline (blue; n=8), fexofenadine plus 2.5 mg/kg FD&C Red No. 40 (green; n=4), and fexofenadine plus 25 mg/kg FD&C Red No. 40 (red; n=9). **p*<0.05, ***p*<0.01; ANOVA with Tukey's correction.

Of the 18 excipient dyes identified as inhibitors of OATP2B1 transport, 8 belong to the azo dye family: synthetic dyes with one or more azo bonds (the functional group R-N=N-R'). The reductive cleavage of the azo bond is facilitated by azoreductases, which are encoded by phylogenetically diverse bacteria, including multiple bacterial taxa prevalent in the human gastrointestinal tract (21). Because azo dye excipients are orally administered, they have the opportunity of encountering and being cleaved by bacterial azoreductases, thus altering their chemical structure and potentially also their bioactivity.

Despite sharing an azo bond, azo dyes are structurally diverse and display variable susceptibility to reduction by bacteria (22, 23). Furthermore, the ability of gut bacteria to metabolize the specific azo dyes identified in our screen was poorly understood (22, 24, 25). To test the ability of complex human gut microbiotas to metabolize the 8 identified azo dye OATP2B1 inhibitors, we performed an *ex vivo* screen wherein each of the azo dye excipients were anaerobically incubated with human fecal samples from three unrelated healthy individuals. One dye was removed from this analysis for technical reasons. Following azo bond reduction, these dyes lose their chromogenic properties and therefore become colorless. All of the tested dyes were cleared by human gut bacteria (**Figure 2.4a**). The extent of dye elimination varied between dyes, ranging from 90.5 ± 2.9 to $48.2\pm 7.3\%$, consistent with prior data suggesting that these enzymes display some degree of substrate specificity (22, 23).

Next, we developed an agar plate-based assay to identify human gut bacterial isolates capable of excipient azo dye metabolism. Dilutions of human fecal suspensions were applied on agar plates supplemented with azo dyes and incubated anaerobically. We identified metabolizers by inspecting agar plates for colonies that produced a zone of dye clearance, indicative of azo bond cleavage (**Figure 2.4b and Supplemental Figure S2.4a**). Representative positive isolates were selected and restreaked on azo dye containing media to confirm their phenotype (**Figure 2.4b, Supplemental Figure. S2.4b and c**). A PCR-based fingerprinting method (26) was used to dereplicate strains from the same subject (26); the same bacterial fingerprint from a single human subject was observed across multiple dyes (**Supplemental Table S2.5**).

To more definitively test for the ability of each isolate to metabolize multiple dyes, we selected 22 unique azo dye excipient metabolizing bacterial isolates for 16S rRNA gene sequencing. We identified bacteria from the three major phyla found in the gut: Bacteroidetes (n=6), Firmicutes (n=10), and Actinobacteria (n=6) (**Supplemental Table S2.5**). All of the bacterial strains tested were capable of clearing multiple azo dyes (**Figure 4c**). The efficiency of dye elimination was primarily determined by strain not by dye (55% versus 19% of total variation; both factors $p < 0.0001$, two-way ANOVA). This analysis also revealed a phylogenetic signature of azo dye excipient clearance, with bacteria belonging to the Firmicutes and Bacteroidetes phyla significantly more active than Actinobacteria (both comparisons $p < 0.0001$, ANOVA with Tukey's correction; **Figure 2.4c and Supplemental Figure S2.5**). Together, these results indicate that human gut bacteria can metabolize multiple excipient dyes and the extent of metabolism is influenced by bacterial taxonomy.

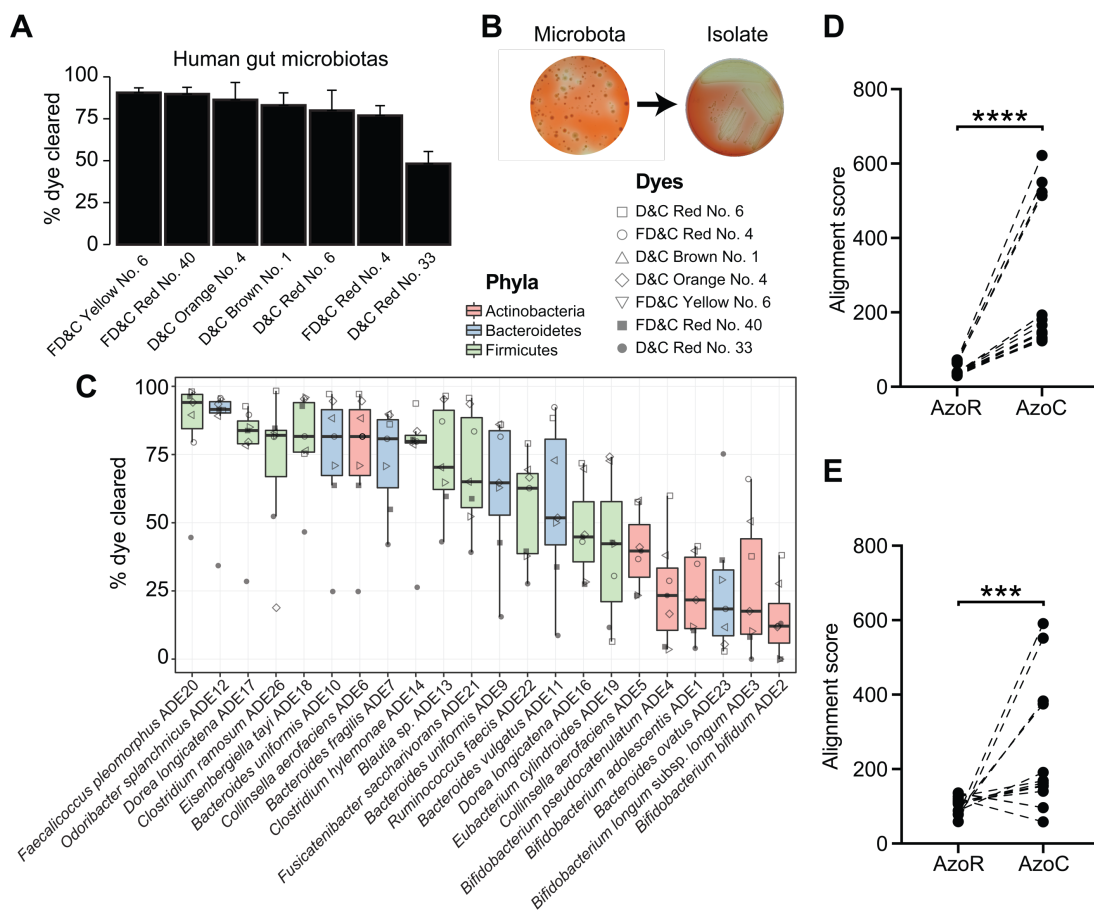


Figure 2.4. Human gut bacteria metabolize azo dye excipients. (A) Excipient azo dyes are cleared following the *ex vivo* incubation with human fecal samples. Dilutions of human fecal suspensions from 3 healthy unrelated individuals were incubated anaerobically for 24 hours in BHI⁺ media supplemented with various excipient azo dyes and conditioned media was sampled and analyzed for residual dye concentration spectrophotometrically. Data is normalized to uninoculated controls. Values represent mean±sem. (B) Isolation of human gut bacteria capable of clearing FD&C Red No. 40. Dilutions of human fecal suspensions were used to inoculate BHI⁺ agar media supplemented with various excipient azo dyes and incubated anaerobically, with FD&C Red No. 40 shown here. Azo dye depleting colonies were identified by a zone of decolorization. Isolates capable of azo dye clearance were picked and restreaked on BHI⁺ agar media supplemented with the same azo dye to confirm the phenotype. (C) Human gut bacterial isolates show variable abilities to deplete azo dye excipients. Human fecal bacterial isolates previously identified as azo dye metabolizers were incubated in triplicate in BHI⁺ media supplemented with various excipient azo dyes and incubated anaerobically for 24 hours. Residual dye was quantified spectrophotometrically from conditioned media samples and normalized to uninoculated controls. Each point represents the mean dye clearance, and the shape corresponds to the excipient azo dye. Colored boxes correspond to bacterial phyla. (D-E) Needleman-Wunsch alignment scores of the azoreductase homologs from the 9 newly sequenced genomes (D) and previously (Figure Caption continued on the next page)

(Figure caption continued from the previous page) sequenced reference genomes (E) when aligned to AzoR from *E. coli* or AzoC from *C. perfringens*. Each point represents one gene. Paired Student's t-test, *** $p < 0.001$, **** $p < 0.0001$.

Next, we sought to gain insight into the genes and enzymes responsible for gut bacterial azo dye excipient metabolism. We selected 9 human gut bacterial isolates from the two most active phyla (**Supplemental Figure S2.5**) for sequencing: Firmicutes (1 *Clostridium ramosum* and 1 *Ruminococcus lactaris*) and Bacteroidetes (2 *Bacteroides ovatus*, 2 *Bacteroides uniformis*, 2 *Bacteroides vulgatus*, and 1 *Odoribacter splanchnicus*). Illumina sequencing and assembly resulted in high quality draft genomes (>99% completion; <5% contamination) for all 9 isolates (**Supplemental Figure S2.6**). We then searched for predicted proteins with significant sequence homology to a set of 5 previously characterized gut bacterial azoreductases originating from the Firmicutes and Proteobacteria phyla (27–31). As expected, all 9 of the isolate genomes encoded at least one azoreductase homolog. Surprisingly, all of the top matches were to AzoC, a non-canonical azoreductase from the strict anaerobic *Clostridium perfringens* that uses flavin adenine dinucleotide (FAD) as a cofactor as opposed to flavin mononucleotide (FMN) (27) (**Figure 2.4d**).

This observation was generalizable across a broader panel of gut bacterial reference genomes. We retrieved 22 additional reference genomes from the Integrated Microbial Genome (IMG) Database (32) that belonged to 6 of the bacterial species identified in our screen, including *Bifidobacterium bifidum* (7 genomes), *Collinsella aerofaciens* (1 genome), *Bacteroides fragilis* (8 genomes), *B. ovatus* (2 genomes), *B. vulgatus* (2 genomes), and *Odoribacter splanchnicus* (2 genomes). Consistent with our

newly sequenced genomes, the top matches for all but one of the species were to AzoC from *Clostridium perfringens*. The one exception, *B. bifidum*, was the least active isolate of the 22 we characterized (**Figure 2.4c**) and was more closely related to AzoR from *E. coli* (**Figure 2.4e**). Taken together, these results suggest that the well-characterized AzoR from *Escherichia coli* may not be representative of the dominant class of azoreductases found within the human gut, even for other Gram-negative phyla like the Bacteroidetes.

We hypothesized that bacterial metabolism of azo dye excipients would decrease their ability to inhibit OATP2B1 transport based on our previous observation that lower molecular weight is a characteristic feature distinguishing OATP2B1 inhibitors versus non-inhibitors (**Figure 2.2b**). To test this hypothesis, we assayed conditioned media of human bacterial isolates grown in the presence of the azo dye FD&C Red No. 40 for OATP2B1 uptake inhibition. This representative azo dye excipient was selected due to our previous *in vivo* results (**Figure 2.3**) and because it is the dye with the highest amount certified in the United States by the FDA and widely used in both food and drug products (19). Samples of conditioned media that showed complete clearance of FD&C Red No. 40 failed to inhibit OATP2B1-mediated estrone sulfate uptake, whereas samples with high levels of FD&C Red No. 40 remaining inhibited uptake (**Figure 2.5a**). LC-HRMS analysis confirmed the depletion of FD&C Red No. 40 from conditioned media samples (**Supplemental Figure S2.6**). As expected, the level of clearance of FD&C Red No. 40 was significantly correlated with OATP2B1-mediated estrone sulfate uptake (**Figure 2.5b**).

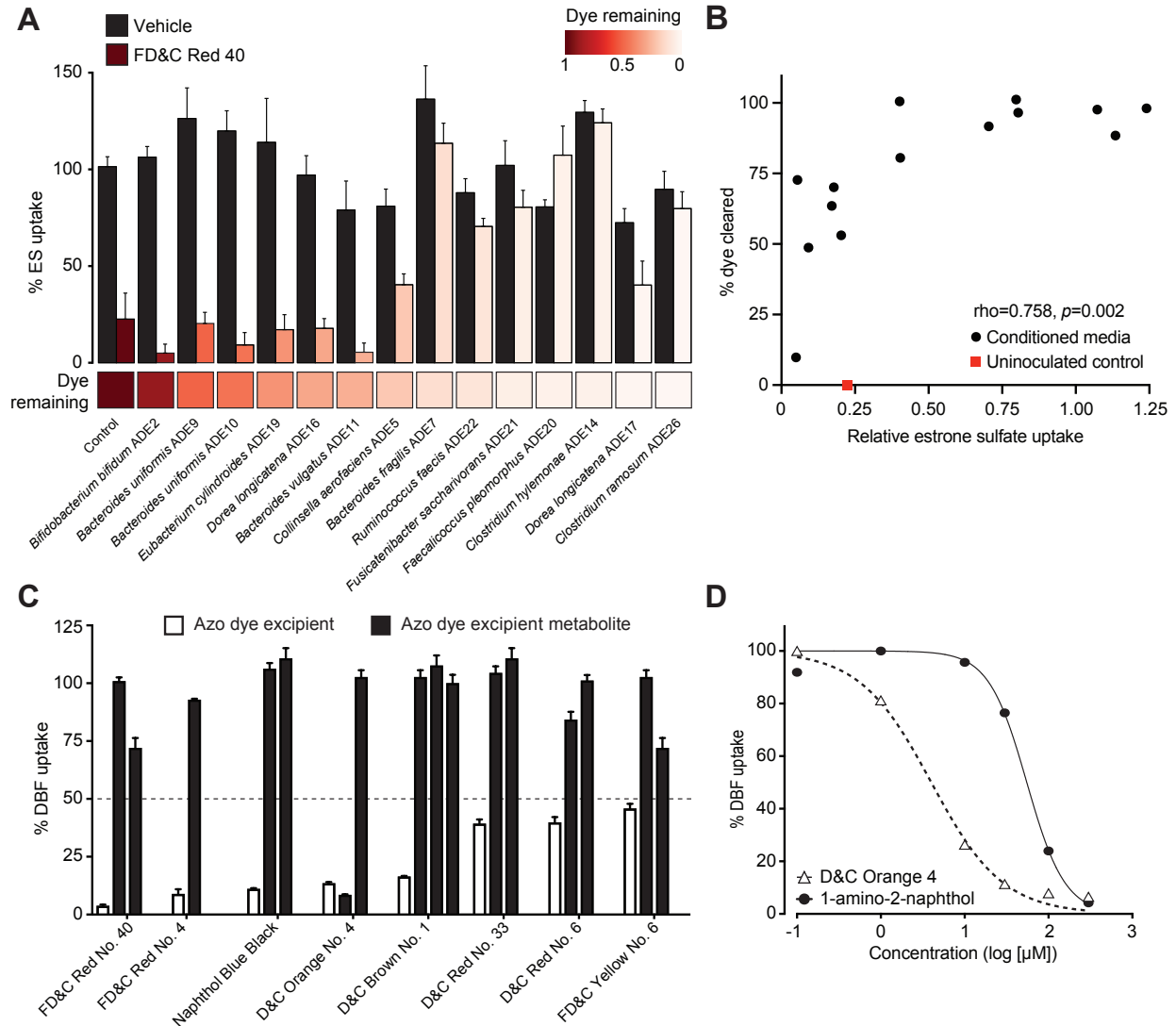


Figure 2.5. Microbial metabolism of azo dye excipients rescues OATP2B1 uptake inhibition. (A) Conditioned media samples of gut bacteria grown in the presence of 200 μM FD&C Red No. 40 or vehicle (DMSO) were assayed for the ability to inhibit OATP2B1 uptake. Human OATP2B1-overexpressing cells were incubated with Hank's Balanced Salt Solution (HBSS) uptake buffer containing 10 nM [^3H]-estrone sulfate and bacterial conditioned media supplemented with FD&C Red No. 40 or vehicle (DMSO). Values represent the normalized mean \pm SD of [^3H]-estrone sulfate uptake from three replicate determinations. The heat map below the bar graph shows relative FD&C Red No. 40 dye remaining in conditioned media samples, which was determined spectrophotometrically. (B) FD&C Red No. 40 levels are associated with OATP2B1-mediated [^3H]-estrone sulfate uptake (Spearman's Correlation). (C) Inhibition of OATP2B1-mediated DBF uptake by azo dyes and their corresponding reduced metabolites. Human OATP2B1-overexpressing HEK cells were incubated in HBSS uptake buffer containing 2 μM DBF for 3 min with the designated concentrations of each azo dye and their metabolites. Concentrations assayed are specified in **Dataset S1** and **SI Appendix, Table S7**. Values represent the mean \pm SD of normalized DBF (Figure Caption continued on the next page)

(Figure caption continued from the previous page) uptake from $n \geq 3$ replicate determinations. (D) Dose-response curves of D&C Orange No. 4 and its metabolite, 1-amino-2-naphthol against OATP2B1-mediated DBF uptake. Values represent the mean \pm SD of DBF uptake from three replicate determinations.

To more definitively test if this decrease in OATP2B1 inhibition was due to the biotransformation of azo dye excipients to their downstream microbial metabolites, we tested 12 unique metabolites from each of the 8 azo dye inhibitors. In contrast to the azo dye substrates, all but one of the corresponding reduced metabolites resulted in levels of DBF uptake above our 50% cutoff, demonstrating their decreased ability to inhibit OATP2B1 transport (**Figure 2.5c and Supplemental Table S2.7**). The one exception (a metabolite of D&C Orange No. 4) had a much higher inhibition constant: the K_i value for D&C Orange No. 4 was 2.11 μM and the K_i values were 62.5 μM and $>200 \mu\text{M}$ for 1-amino-2-naphthol and sulfanilic acid, respectively (**Figure 2.5d**). This *in vitro* data suggests that gut bacterial azoreductase activity could rescue the inhibition of OATP2B1 uptake by azo dye excipients.

Experiments in mouse models provided additional support for the physiological relevance of azo dye excipient FD&C Red No. 40 and the gut microbiome for the absorption of the OATP2B1 substrate fexofenadine. As expected, conventionally-raised (CONV-R) wild-type BALB/c mice excreted significantly less FD&C Red No. 40 (oral dose of 5 mg/mouse or $\sim 250 \text{ mg/kg}$) relative to germ-free controls (**Figure 2.6a**; $n=3-8$ mice/group). No dye was detected in untreated CONV-R mice. FD&C Red No. 40 was also detectable in stool samples at the more modest dose (25 mg/kg) that led to reduced fexofenadine bioavailability but not at the lowest dose tested (2.5 mg/kg) that had plasma fexofenadine levels indistinguishable from vehicle controls (**Figures 2.3 and 2.6b**). These

observations led us to hypothesize that gut bacterial azoreductase activity could be saturated at high azo dye levels, which would lead to OATP2B1 inhibition even in the presence of azo dye excipient metabolizing strains. Consistent with this hypothesis, colonization of germ-free mice with two distinct 3-member synthetic communities composed of the top 3 and bottom 3 FD&C Red No. 40 metabolizers (**Supplemental Figure S2.4c**) led to levels of dye excretion that were indistinguishable from germ-free controls (**Figure 2.6a**; n=3-8 mice/group). The redundancy of azoreductases across diverse bacterial taxa together with the higher levels of colonization (**Figures 2.6c and d**) may explain why CONV-R mice clear azo dyes better than the gnotobiotic mice tested here (**Figure 6a**). Alternatively, the broader differences in intestinal permeability (33) and gene expression (34, 35) in germ-free and gnotobiotic mice may disrupt the complex interactions between excipients, the gut microbiome, and drug transporters. We rederived the Pgp deficient mouse model under germ-free conditions (**Materials and Methods**) and repeated our fexofenadine pharmacokinetics experiment; however, unlike in CONV-R mice (**Figure 2.3**), the high dose of FD&C Red No. 40 had a more modest impact on fexofenadine bioavailability in germ-free Pgp deficient animals (**Figure 2.6e-h and Supplemental Table S2.4**). More work is needed to determine the mechanisms responsible, including differences in both drug and excipient disposition in the absence of a gut microbiome.

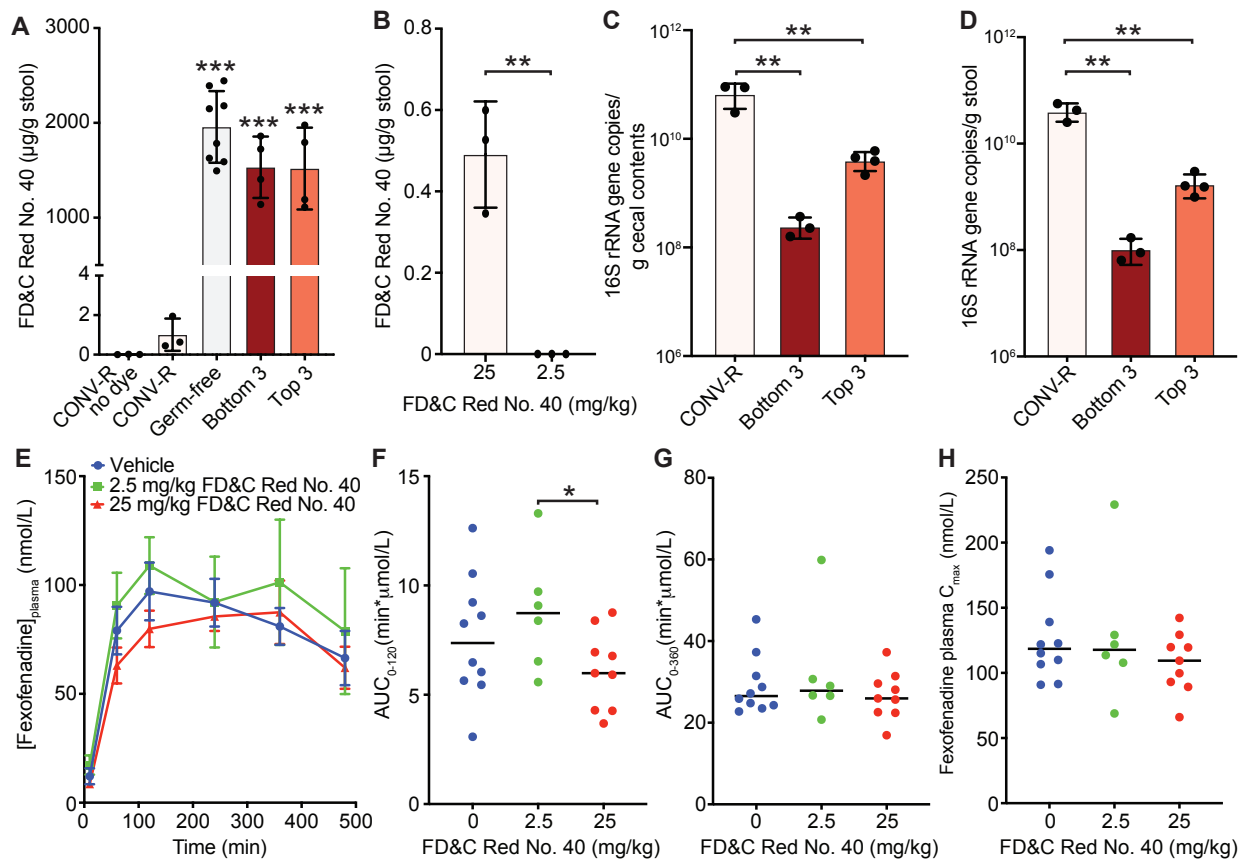


Figure 2.6. Interactions between the gut microbiota, azo dye excipients, and drug absorption in conventionally-raised (CONV-R) and gnotobiotic mice. (A,B) Recovery of residual FD&C Red No. 40 dye in fecal collections over 24 hours. (A) CONV-R, germ-free (GF), and gnotobiotic mice were orally dosed with 5 mg of FD&C Red No. 40 (~250 mg/kg; n=3-8 mice/group). CONV-R mice without dye are also included as a control (n=3). Bottom 3 indicates mice colonized with the 3 worst FD&C Red No. 40 metabolizers: *Bifidobacterium bifidum* ADE2, *Bifidobacterium pseudocatenulatum* ADE4, and *Bifidobacterium adolescentis* ADE1. Top 3 indicates mice colonized with the best 3 FD&C Red No. 40 metabolizers: *Faecalicoccus pleomorphus* ADE20, *Clostridium hylemonae* ADE14, and *Clostridium ramosum* ADE26. (B) Fecal FD&C Red No. 40 dye levels in CONV-R mice depend upon oral dose (n=3 mice/group). (C,D) Total colonization is significantly higher in the (C) cecum and (D) stool of CONV-R mice relative to both gnotobiotic groups. (E) Plasma concentration-time profiles of fexofenadine after oral administration of 15 mg/kg fexofenadine plus saline (blue circles, n=10), 2.5 mg/kg FD&C Red No. 40 (green squares, n=6), or 25 mg/kg FD&C Red No. 40 (red triangles, n=9) to P-glycoprotein (Pgp) deficient (*mdr1a/b*^{-/-}) germ-free mice. Each point represents the mean±sem. (F-H) Pharmacokinetic parameters calculated using Phoenix WinNonlin v8.1: (F) plasma fexofenadine AUC_{0-120min}; (G) plasma fexofenadine AUC_{0-360min}; and (H) maximum plasma concentration (C_{max}). Each point represents the data from an individual mouse administered with 15 mg/kg fexofenadine plus saline (blue; n=10), 2.5 mg/kg FD&C Red No. 40 (green; n=6), (Figure Caption continued on the next page)

(Figure caption continued from the previous page) or 25 mg/kg FD&C Red No. 40 (red; n=9). Statistics: (A,C,D) ANOVA with Tukey's correction, ** $p < 0.01$, *** $p < 0.001$, (B,F) Unpaired Student's t-test, * $p < 0.05$, ** $p < 0.01$. The reported contrasts in panel A are relative to CONV-R mice either with or without dye administration.

2.3 Discussion

Pharmaceutical excipients constitute, on average, 90% of a drug formulation and yet they are often assumed to be inactive despite not being explicitly tested. Nevertheless, the Similarity Ensemble Approach (SEA) (36) reveals that some excipients are structurally similar to bioactive molecules, including drugs, and may share their on-target activities (11). Our results confirm and extend several recent *in vitro* studies suggesting that excipients can inhibit intestinal drug absorption (37–39). By comprehensively screening a collection of 136 oral molecular excipients, we report the discovery of 24 inhibitors of the intestinal drug absorptive transporter OATP2B1. These excipients belong to several functional classes including dyes, surfactants, antimicrobial preservatives, and flavoring agents, 6 of which are predicted to achieve clinically relevant intestinal concentrations (**Supplemental Table S2.8**). Our study, along with others, suggest that pharmaceutical excipients are more bioactive than assumed, potentially affecting drug therapy, human biology, and pathophysiology, which are all important areas for future study. Though clinical trials are needed to determine the *in vivo* effects of these excipients, these findings have potential clinical implications due to the key role of OATP2B1 in drug absorption (8, 9, 17, 18). On the other hand, our data provides multiple strategies for mitigating the bioactivity of food and drug additives. These include identifying prescription drugs that are substrates of OATP2B1 and additives that are inhibitors of OATP2B1. Current formulation strategies can be modified accordingly to mechanistic-understanding-based formulation, which is of particular importance for the generic drug industry to develop bioequivalent drug product formulations (40).

To further assess the potential clinical relevance of the OATP2B1-inhibiting excipients, we estimated the maximal intestinal concentrations of excipients based on the maximum allowable amount from the World Health Organization and the Code of Federal Regulations Title 21. According to the FDA perspective about the role of transporters in drug-drug interactions (41), it is possible that excipients with $[I]/K_i \geq 10$ ($[I]$, estimated intestinal concentration; K_i , inhibitory constant) could inhibit OATP2B1-mediated drug absorption at clinically relevant concentrations. The $[I]/K_i$ values for six excipients exceed 10 if the maximum allowable amount is administered as a single bolus (**Supplemental Table S2.8**). In addition, a survey reported that the amounts of artificial food colors certified for use in the United States by the FDA was 62 mg/capita/day in 2010, of which FD&C Red No. 40 accounts for 40%, FD&C Yellow No. 6 for 24%, FD&C Blue No.1 for 4%, and FD&C Red No. 3 for 4% (19). Based on 62 mg/day, the $[I]/K_i$ for FD&C Red No. 40 is 84.9 (**Supplemental Table S2.8**), and thus could have clinically meaningful effects on drug absorption.

Most oral drug products do not contain sufficient concentrations of the identified excipients, especially dyes. Larger amounts are present in food; however, the dyes we consume as part of our diet are not usually delivered as a single bolus together with drugs, lowering the risk of food-drug interactions. High levels of single excipients (e.g., FD&C Red No. 40) can be found in baked food products; up to 1240 mg/kg (42). Considering a 100 g cupcake, a single serving of baked food products may contain as much as 124 mgs of FD&C Red No. 40. For an adult, the fasted small bowel contains a total volume of 43 ± 14 ml of water (43), suggesting that the potential intestinal concentration of FD&C Red No. 40 could reach as high as 10 mM following a single serving of a baked food

product. Taken together, these calculations suggest that while drug-dye interactions within a drug formulation may be below physiologically relevant concentrations, dyes consumed through select food products may reach high intestinal concentrations, resulting in the inhibition of OATP2B1-mediated absorption of prescription drugs.

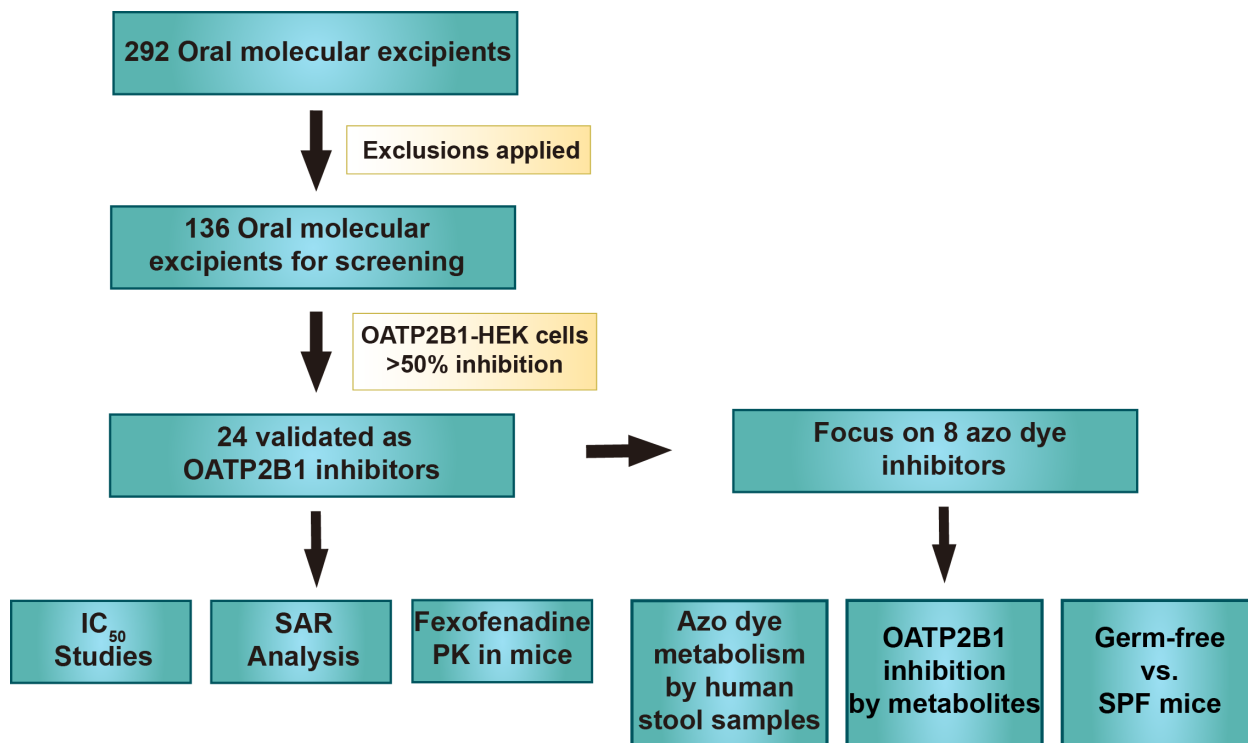
Despite these projections, we found that mice administered with 2.5 mg/kg (1 mM estimated intestinal concentration) of FD&C Red No. 40 did not show a reduction of fexofenadine plasma levels. A likely explanation for this lack of inhibition *in vivo* is that 2.5 mg/kg FD&C Red No. 40 falls within the metabolic capacity of the gut microbiome's azoreductase activity leaving no intact dye and therefore no OATP2B1 interaction. This is supported by a lack of dye detected in the stool (**Figure 2.6b**). On the other hand, 25 mg/kg (10 mM estimated intestinal concentration) of FD&C Red No. 40 appears to saturate the gut microbiome's azoreductase capacity with detectable levels of FD&C Red No. 40 in the stool, resulting in an inhibition of OATP2B1 transport and thus decreased fexofenadine absorption (**Figure 2.3**).

Our data shows that excipient inhibitors of OATP2B1 transport are enriched for azo dyes. These dyes are susceptible to metabolism by human gut bacterial azoreductases producing metabolites that no longer inhibit OATP2B1 transport (**Supplemental Figure S2.7**). While our bacterial isolate data provide some support for the possibility that differences in the gut microbiome could alter azo dye metabolism, the ubiquity and functional redundancy of this enzymatic activity across microbial communities and diverse bacterial taxonomic groups (44) suggests that microbiome shifts may not contribute in a meaningful way to inter-individual differences in excipient-transporter interactions.

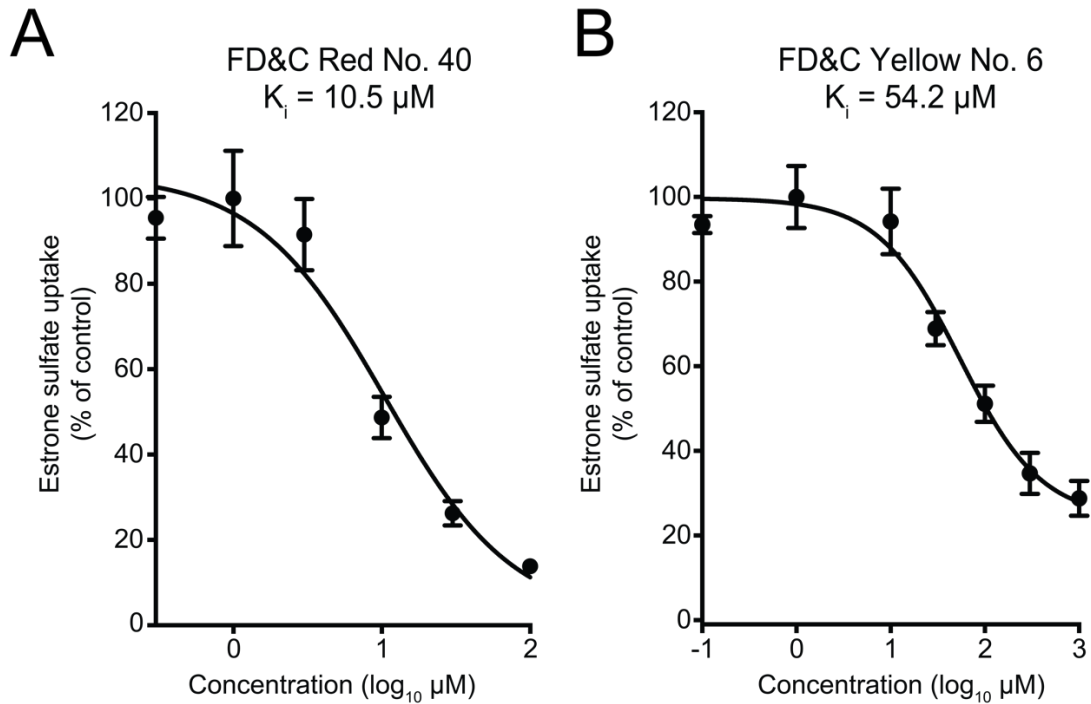
Interestingly, our pharmacokinetic data demonstrate that plasma fexofenadine levels are largely unaffected by FD&C Red No. 40 administration in germ-free mice, which lack bacterial azoreductase activity. We did observe a significant decrease in the fexofenadine AUC_{0-120} at 25 mg/kg versus 2.5 mg/kg FD&C Red No. 40 or vehicle. On the contrary, despite possessing bacterial azoreductases, CONV-R mice are sensitive to inhibition of fexofenadine absorption at higher FD&C Red No. 40 (25 mg/kg) doses. Despite serving as a powerful tool for investigating the microbial contributions towards the host, germ-free animals display gross developmental and physiological differences compared to their CONV-R counterparts, including broad differences in the expression of genes involved in drug metabolism and transport (34, 35). Although, OATP2B1 gene expression does not appear to be differentially expressed in germ-free versus CONV-R mice (35), other factors such as OATP2B1 protein levels and microbial metabolite substrates or inhibitors of OATP2B1 could also contribute to the observed fexofenadine transport inhibition at a high dose of FD&C Red No. 40. Further research is needed to address the mechanisms responsible for the observed differences in fexofenadine absorption during FD&C Red No. 40 administration in CONV-R versus germ-free.

In conclusion, these findings emphasize the importance of considering the chemical milieu in which drugs are taken and provide a mechanism through which the human gut microbiome ameliorates the impact of chemical exposures. This study also has clear translational implications, providing the first step towards the rational selection of excipients to minimize their off-target impacts on human and microbial cells.

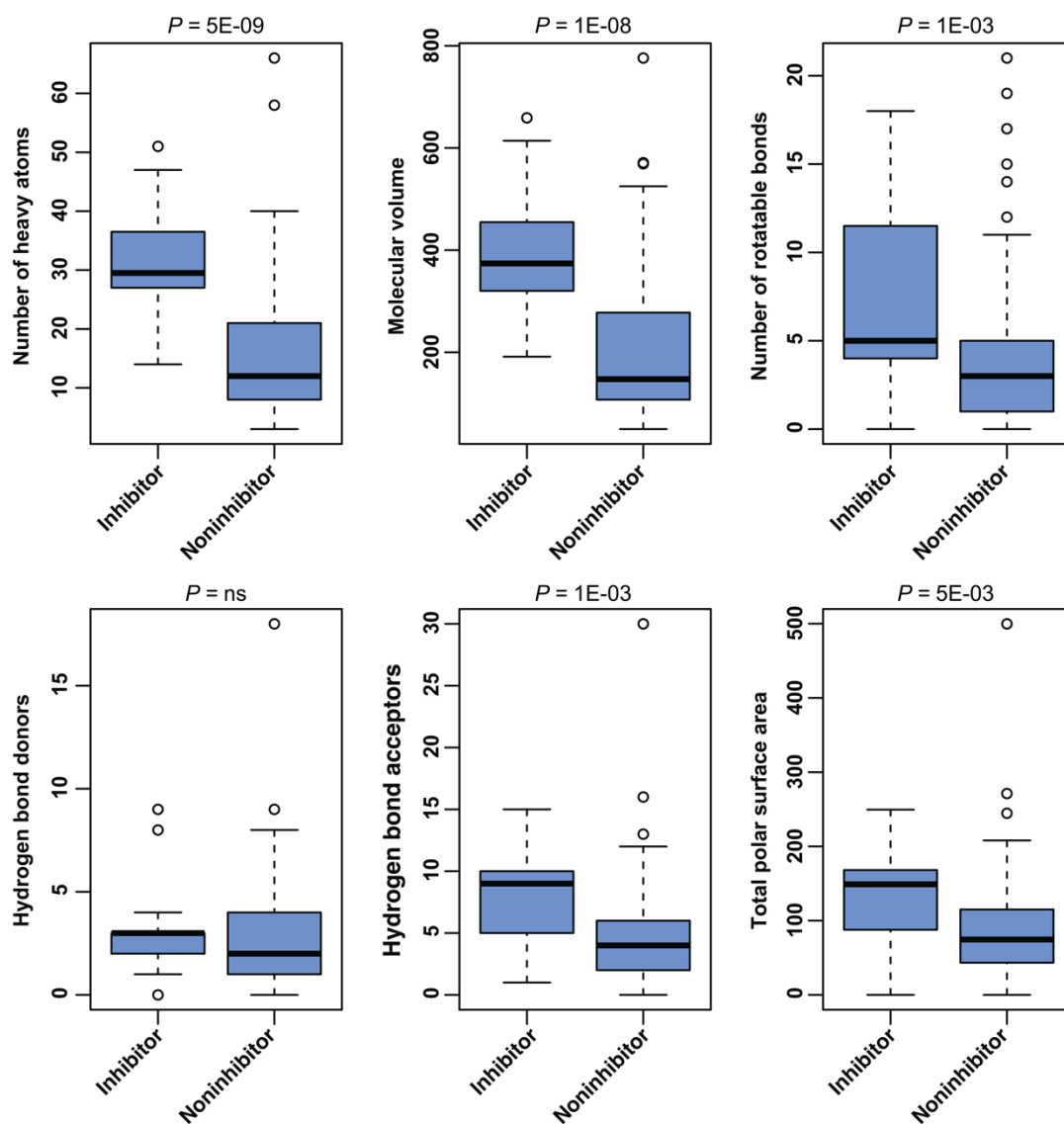
2.4 Supplementary Figures



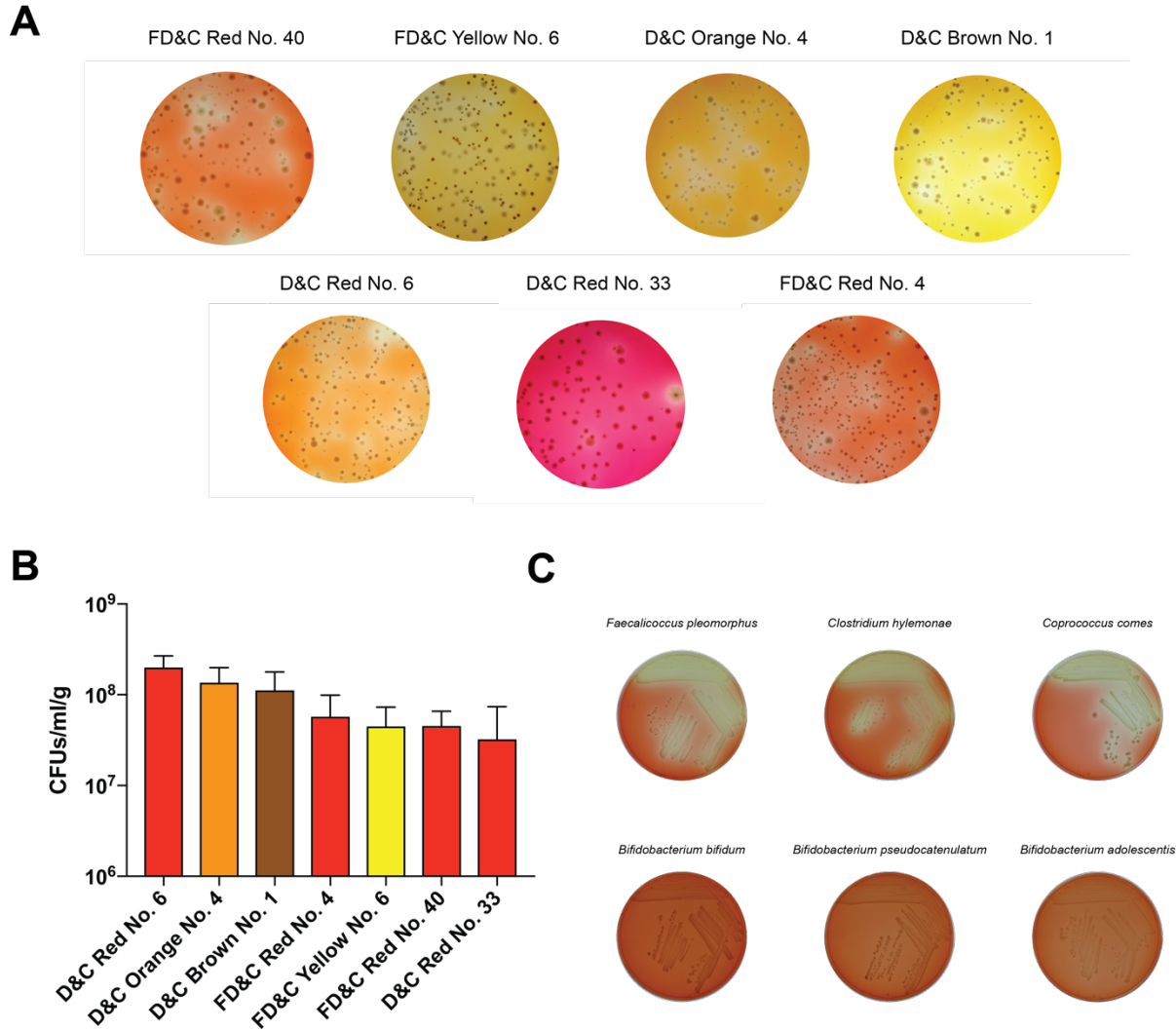
Supplemental Figure S2.1. Study design overview. In total a list of 136 curated oral molecular excipients were compiled and screened for inhibition of OATP2B1-mediated drug uptake in OATP2B1-overexpressing HEK293 cells. IC₅₀ studies and quantitative structure-activity relationship (SAR) were conducted on 24 identified inhibitors and fexofenadine pharmacokinetics (PK) were evaluated to examine the inhibitory effect of the representative OATP2B1 inhibitor, FD&C Red No. 40, in P-glycoprotein-deficient (*mdr1a/b*^{-/-}) mice. In addition, the metabolism of OATP2B1-inhibiting azo dyes by gut bacteria was assessed through *ex vivo* incubations of healthy human stool samples, *in vitro* analysis of gut bacterial isolates, and experiments in germ-free and conventionally-raised (SPF, specific pathogen free) mice. The inhibitory effect of azo dyes and their bacterial metabolites on OATP2B1 was also determined.



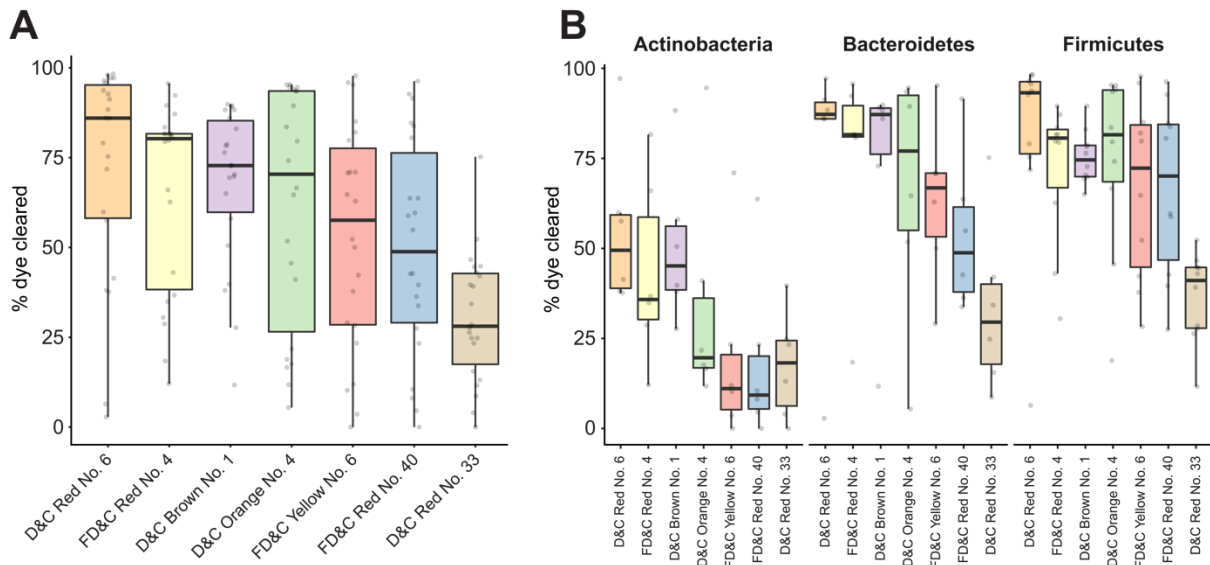
Supplemental Figure S2.2. Dose-dependent inhibition of OATP2B1-mediated estrone sulfate (ES) uptake by FD&C Red No. 40 and FD&C Yellow No. 6. Human OATP2B1-overexpressing cells were incubated with Hank's Balanced Salt Solution (HBSS) uptake buffer containing 10 nM [^3H]-estrone sulfate with designated concentrations of FD&C Red No. 40 (A) or FD&C Yellow No. 6 (B) for 3 min. Data points represent the mean \pm SD of [^3H]-estrone sulfate uptake from three replicate determinations.



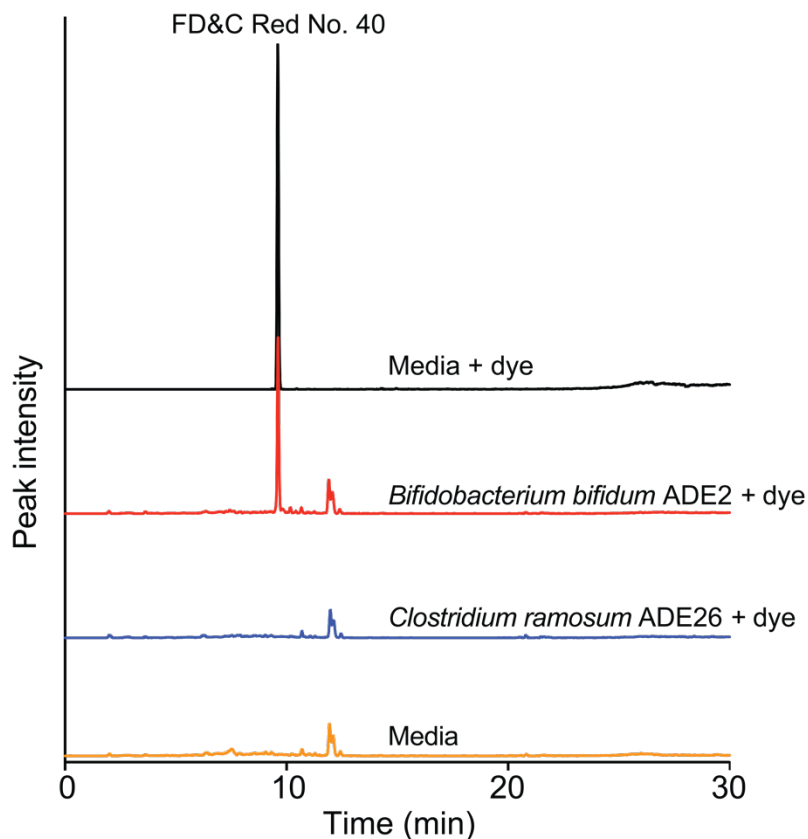
Supplemental Figure S2.3. Comparison of physicochemical properties between OATP2B1 excipient inhibitors and non-inhibitors. Box plots comparing physicochemical properties of 24 inhibitors versus 112 non-inhibitors. The bold line represents the median and the error bars the upper and bottom quartiles. Points outside the bars are shown as open circles. p -values represent Student's t -tests.



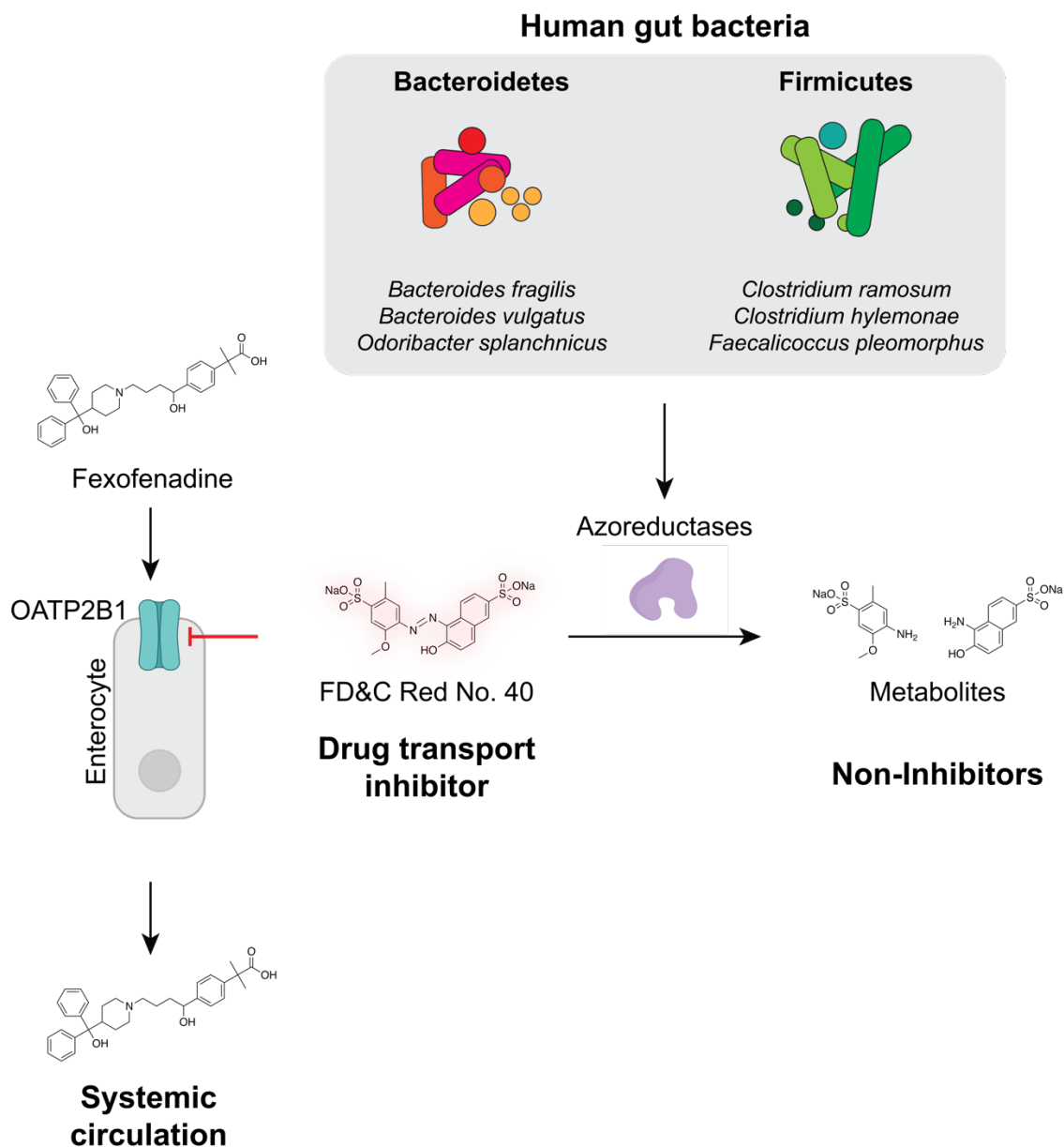
Supplemental Figure S2.4. Isolation of azo dye excipient metabolizing human gut bacterial isolates. (A) A dilution of human fecal suspensions were plated on BHI⁺ media supplemented with excipient azo dyes and incubated anaerobically. Metabolizing colonies were identified by a characteristic decolorization halo surrounding the colony. (B) Recovery numbers of azo dye metabolizing bacterial isolates from 3 unrelated healthy human fecal donors. Bar plots show mean±SD. (C) Azo dye metabolizing colonies were restreaked on BHI⁺ media supplemented with azo dye to confirm their phenotype. Top agar plates show isolates capable of FD&C Red No. 40 metabolism. The bottom plates show isolates unable to metabolize FD&C Red No. 40.



Supplemental Figure S2.5. Metabolism of azo dye excipients by human gut bacterial isolates. 22 unique human gut bacterial isolates recovered from 3 healthy human fecal samples were assayed for metabolism against 7 azo dye excipients. Bacteria isolates were inoculated in BHI⁺ media supplemented with azo dye and incubated anaerobically for 24 hours. Conditioned media was recovered and assayed for residual azo dye spectrophotometrically. All assays were performed in triplicate and normalized to an uninoculated control sample. (A) Boxplots displaying percent azo dye metabolized by all strains for each azo dye. Each dot represents the mean metabolism by a single bacterial isolate. (B) Boxplots displaying percent azo dye metabolized grouped according to phylum for each azo dye. Each dot represents the mean metabolism by a single bacterial isolate.



Supplemental Figure S2.6. Measurement of residual FD&C Red No. 40 in conditioned media samples by liquid chromatography-high resolution mass spectrometry (LC-HRMS). Extracted ion chromatograms (EICs) show the presence of FD&C Red No. 40 in the conditioned media sample of a non-metabolizing isolate (*Bifidobacterium bifidum* ADE2, red), but could not be detected in a sample from a metabolizing isolate (*Clostridium ramosum* ADE26, blue). Sterile media supplemented with FD&C Red No. 40 is shown in black. Sterile media is shown in yellow. A 20 ppm error was used for each trace.



Supplemental Figure S2.7. Gut bacterial metabolism rescues the inhibitory effects of intestinal drug transport by excipient azo dyes. Common food and drug additives, including azo dyes like FD&C Red No. 40, hinder drug absorption by inhibiting the intestinal drug transporter OATP2B1 and therefore decrease systemic drug levels. Azoreductases from human gut bacteria metabolize azo dyes and the downstream metabolites no longer inhibit OATP2B1 transport. Human gut bacteria belonging to the Bacteroidetes and Firmicutes phyla are the most active azo dye metabolizers, with representative species indicated.

2.5 Supplementary Tables

Supplemental Table S2.1. Validation and inhibitory potencies of 24 OATP2B1 excipient inhibitors.

Excipient	Functional Category	K _i (μM) (95% confidence intervals)
Benzalkonium chloride	Antimicrobial agent	62.1 (20.2-187)
Butylparaben	Antimicrobial agent	44.3 (33.8-58.0)
D&C Brown No. 1	Azo dye	3.11 (2.31-4.20)
D&C Orange No. 4	Azo dye	2.11 (1.86-2.41)
D&C Red No. 33	Azo dye	58.1 (49.2-68.7)
FD&C Red No. 4	Azo dye	8.10 (7.34-8.95)
D&C Red No. 6	Azo dye	11.3 (7.69-16.7)
FD&C Red No. 40	Azo dye	2.59 (1.92-3.50)
FD&C Yellow No. 6	Azo dye	68.4 (58.3-80.4)
Naphthol Blue Black	Azo dye	0.40 (0.32– 0.49)
D&C Green No. 5	Dye	1.54 (1.18-2.01)
D&C Red No. 27	Dye	1.0 (0.45-1.31)
D&C Red No. 28	Dye	1.0 (0.65-1.57)
EXT. D&C Yellow No. 7	Dye	24.5 (21.9-27.5)
FD&C Blue No. 1	Dye	13.0 (11.7-14.4)
FD&C Red No. 3	Dye	0.88 (0.69-1.11)
Guinea Green B	Dye	0.77 (0.64-0.92)
Light green CF Yellowish	Dye	0.80 (0.73-0.89)
Ponceau xylidine	Dye	3.45 (3.04-3.92)
Rhodamine B	Dye	44.4 (36.4-54.0)
Neohesperidin dihydrochalone	Flavoring agent	20.1 (16.2-24.8)

Excipient	Functional Category	K _i (μM) (95% confidence intervals)
Docusate sodium	Surfactant	2.76 (1.41-5.42)
Sodium lauryl sulfate	Surfactant	1.98 (1.37-2.85)
Sucrose monolaurate	Surfactant	47.7 (36.8-61.9)

Supplemental Table S2.2. Aggregation analysis of OATP2B1 excipient inhibitors.

Excipient	K_i (μM)	Aggregation
Naphthol Blue Black	0.40	No Aggregation @ 5 μM
Light Green CF Yellowish	0.80	No Aggregation @ 200 μM
FD&C Red No. 3	0.88	No Aggregation @ 500 μM
D&C Red No. 28	1.0	No Aggregation @ 10 μM
Sodium lauryl sulfate	1.98	No Aggregation @ 50 μM
D&C Orange No. 4	2.11	No Aggregation @ 100 μM
FD&C Red No. 40	2.59	No Aggregation @ 500 μM
D&C Brown No. 1	3.11	No Aggregation @ 50 μM
FD&C Blue No. 1	13.0	No Aggregation @ 200 μM
Neohesperidin dihydrochalone	20.1	No Aggregation @ 200 μM

Supplemental Table S2.3. Physicochemical features of OATP2B1 excipient inhibitors versus non-inhibitors.

Chemical feature	Inhibitors		Non-inhibitors		<i>p</i> -value ¹
	Mean	Standard deviation	Mean	Standard deviation	
Molecular weight (MW)	504.0	181.0	216.0	157.0	4×10 ⁻⁰⁸
No. of heavy atoms (HA)	31.6	9.6	14.7	10.9	5×10 ⁻⁰⁹
Molecular volume (MV)	401.0	124.0	195.0	132.0	1×10 ⁻⁰⁸
No. of rotatable bonds (RB)	7.0	4.7	4.0	4.3	1×10 ⁻⁰³
No. of hydrogen bond donors (HBD)	2.9	2.0	2.9	2.9	NS ²
No. of hydrogen bond acceptors (HBA)	8.2	3.5	5.2	4.3	1×10 ⁻⁰³
SLogP*	6.5	2.6	1.1	2.6	7×10 ⁻¹¹
Total polar surface area (TPSA)	133.0	61.5	90.5	72.6	5×10 ⁻⁰³

¹Student's t-test

²NS, not significant

*SLogP, Crippen's log of the octanol/water partition coefficient (including implicit hydrogens)

Supplemental Table S2.4. Pharmacokinetic parameters of fexofenadine with FD&C Red No. 40 in conventional and germ-free P-glycoprotein deficient mice.

Parameter	Experimental group		
	Fexofenadine + vehicle	Fexofenadine + 2.5 mg/kg FD&C Red No. 40	Fexofenadine + 25 mg/kg FD&C Red No. 40
Conventional P-glycoprotein (Pgp) deficient (<i>mdr1a/b</i>^{-/-}) mice			
n=	8	4	9
AUC ₀₋₁₂₀ (min*µmol/L)	10.4 ± 4.1**	9.42 ± 1.21	5.31 ± 2.45**
AUC ₀₋₃₆₀ (min*µmol/L)	29.0 ± 10**	27.1 ± 3.5*	15.4 ± 5.98*,**
C _{max} (nmol/L)	107 ± 38.9	111 ± 13.4	70.0 ± 33.26
T _{max} (min)	90 (30 - 240)	90 (60 - 120)	120 (30 - 240)
Germ-free P-glycoprotein (Pgp) deficient (<i>mdr1a/b</i>^{-/-}) mice			
n=	10	6	9
AUC ₀₋₁₂₀ (min*µmol/L)	7.60 ± 2.8	8.77 ± 2.72 [#]	6.13 ± 1.80 [#]
AUC ₀₋₃₆₀ (min*µmol/L)	29.1 ± 7.17	32.3 ± 13.9	26.7 ± 5.88

Parameter	Experimental group		
	Fexofenadine + vehicle	Fexofenadine + 2.5 mg/kg FD&C Red No. 40	Fexofenadine + 25 mg/kg FD&C Red No. 40
C_{max} (nmol/L)	127 ± 34.1	128 ± 53.6	108 ± 23.2
T_{max} (min)	180 (60 - 480)	120 (60 - 360)	360 (60 - 480)

* $p < 0.05$, ANOVA with Tukey's correction; 2.5 mg/kg versus 25 mg/kg FD&C Red No. 40.

** $p < 0.01$, ANOVA with Tukey's correction; vehicle versus 25 mg/kg FD&C Red No. 40.

$p < 0.05$, Unpaired Student's t-test; 2.5 mg/kg versus 25 mg/kg FD&C Red No. 40.

Supplemental Table S2.5. Human gut bacterial isolates capable of clearing azo dyes¹.

Isolate ID	Phylum	Species ID (RDP)	Human sample ID	FD&C Red No. 40	FD&C Yellow No. 6	D&C Orange No. 4	D&C Brown No. 1	D&C Red No. 6	D&C Red No. 33	FD&C Red No. 4
ADE 1	Actinobacteria	<i>Bifidobacterium adolescentis</i>	C				Isolated			
ADE 2		<i>Bifidobacterium bifidum</i>	B					Isolated		
ADE 3		<i>Bifidobacterium longum</i> subsp. <i>longum</i>	A					Isolated		
ADE 4		<i>Bifidobacterium pseudocatenulatum</i>	B				Isolated			
ADE 5		<i>Collinsella aerofaciens</i>	A					Isolated		
ADE 6		<i>Collinsella aerofaciens</i>	C							Isolated
ADE 7	Bacteroidetes	<i>Bacteroides fragilis</i>	C	Isolated						
ADE 23		<i>Bacteroides ovatus</i>	A						Isolated	
ADE 8		<i>Bacteroides ovatus</i>	A	Match						Isolated
ADE 9		<i>Bacteroides uniformis</i>	A	Isolated	Match	Match				
ADE 10		<i>Bacteroides uniformis</i>	B	Isolated		Match	Match			
ADE 24		<i>Bacteroides vulgatus</i>	B							Isolated
ADE 25		<i>Bacteroides vulgatus</i>	C			Isolated				

Isolate ID	Phylum	Species ID (RDP)	Human sample ID	FD&C Red No. 40	FD&C Yellow No. 6	D&C Orange No. 4	D&C Brown No. 1	D&C Red No. 6	D&C Red No. 33	FD&C Red No. 4
ADE 11		<i>Bacteroides vulgatus</i>	B							Isolated
ADE 12		<i>Odoribacter splanchnicus</i>	B				Isolated			
ADE 13	Firmicutes	<i>Blautia sp.</i>	C	Match					Isolated	Match
ADE 14		<i>Clostridium hylemonae</i>	B			Isolated				
ADE 26		<i>Clostridium ramosum</i>	A				Isolated	Match		
ADE 15		<i>Ruminococcus lactaris</i>	B			Match		Isolated		
ADE 16		<i>Dorea longicatena</i>	B					Isolated		
ADE 17		<i>Dorea longicatena</i>	C				Isolated			
ADE 18		<i>Eisenbergiella tayi</i>	B		Match	Match				
ADE 19		<i>Eubacterium cylindroides</i>	A				Isolated		Match	Match
ADE 20		<i>Faecalicoccus pleomorphus</i>	A		Isolated				Isolated	
ADE 21		<i>Fusicatenibacter saccharivorans</i>	C		Match					
ADE 22		<i>Ruminococcus faecis</i>	C							

¹Isolated indicates the strains we characterized. Match indicates strains that were dropped during dereplication (identical fingerprints to the characterized strain).

Supplemental Table S2.6. Human gut bacterial isolate genome sequencing statistics.

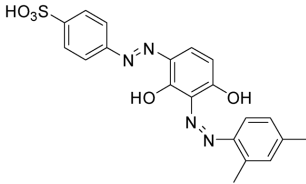
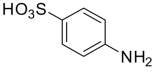
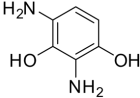
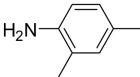
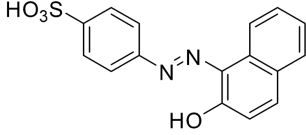
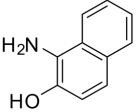
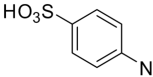
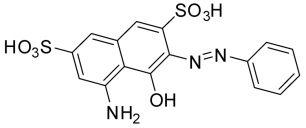
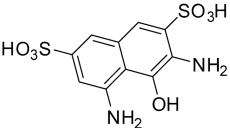
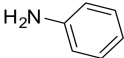
Isolate ID	Species: RDP ¹ / MiGA ²	Human sample ID	No. of contigs	N50	Length (bp)	G C %	Completeness ³	Contamination %*
ADE23	<i>Bacteroides ovatus</i> / <i>Bacteroides ovatus</i>	A	490	37,526	6,993,105	42.02	99.5	0.4
ADE8	<i>Bacteroides ovatus</i> / <i>Bacteroides ovatus</i>	A	155	96,019	7,007,536	41.98	99.5	0.4
ADE9	<i>Bacteroides uniformis</i> / <i>Bacteroides uniformis</i>	A	61	203,982	4,849,103	46.49	99.3	0.3
ADE26	<i>Clostridium ramosum</i> / <i>Erysipelotrichaceae</i> bacterium GAM147	A	81	87,080	2,799,720	29.95	99.1	0.0
ADE10	<i>Bacteroides uniformis</i> / <i>Bacteroides uniformis</i>	B	479	25,346	5,098,547	47.04	99.2	0.4
ADE11	<i>Bacteroides vulgatus</i> / <i>Bacteroides vulgatus</i>	B	173	66,034	5,320,796	42.12	99.3	0.4
ADE15	<i>Ruminococcus lactaris</i> / Family <i>Lachnospiraceae</i>	B	188	40,242	3,326,352	42.69	99.4	0.1
ADE12	<i>Odoribacter splanchnicus</i> / <i>Odoribacter splanchnicus</i>	B	154	59,843	4,600,354	43.39	99.5	0.0
ADE25	<i>Bacteroides vulgatus</i> / <i>Bacteroides vulgatus</i>	C	317	39,866	5,339,661	42.23	99.2	4.1

¹Species identification determined using RDP SeqMatch.

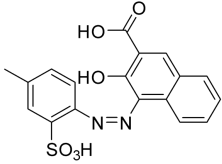
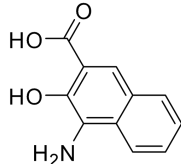
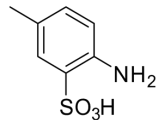
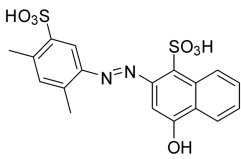
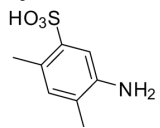
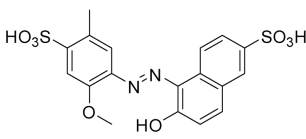
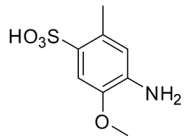
²Species identification determined using Microbial Genomes Atlas (MiGA) NCBI Prok Search (58).

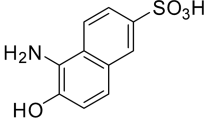
³Metrics determined using CheckM based on clade-specific marker gene occurrence (60).

Supplemental Table S2.7. Inhibitor constant (K_i) values for excipient azo dyes and their metabolites.

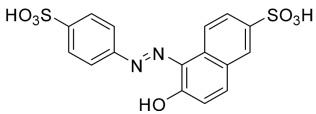
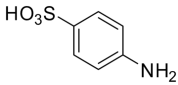
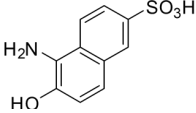
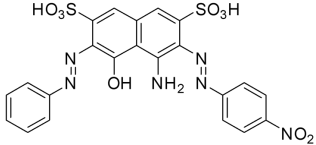
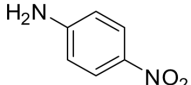
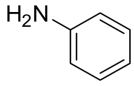
Excipient azo dyes	K_i (μM)	Azo dye metabolite	Max screening concentration (μM)	K_i (μM)
<p>D&C Brown No. 1</p> 	3.11	<p>Sulfanilic acid</p> 	200	>200
		<p>2,4-Diaminobenzene-1,3-diol</p> 	200	>200
		<p>2,4-Dimethylaniline</p> 	200	>200
<p>D&C Orange No. 4</p> 	2.11	<p>1-Amino-2-naphthol</p> 	200	62.5
		<p>Sulfanilic acid</p> 	200	>200
<p>D&C Red No. 33</p> 	58.1	<p>3,5-Diamino-4-hydroxy-naphthalene-2,7-disulfonic acid</p> 	50	>50
		<p>Aniline</p> 	200	>200

Supplemental Table S2.7. Inhibitor constant (K_i) values for excipient azo dyes and their metabolites cont'd.

Excipient azo dyes	K_i (μM)	Azo dye metabolite	Max screening concentration (μM)	K_i (μM)
<p>D&C Red No. 6</p> 	11.3	<p>4-Amino-3-hydroxy-[2]naphthoic acid</p> 	200	>200
		<p>4-Aminotoluene-3-sulfonic acid</p> 	200	>200
<p>FD&C Red No. 4</p> 	8.10	<p>5-Amino-2,4-xylenesulfonic acid</p> 	125	>125
		<p>Unknown metabolite</p>	NA*	NA*
<p>FD&C Red No. 40</p> 	2.59	<p>4-Amino-5-methoxy-2-methyl-benzenesulfonic acid</p> 	50	>50

Excipient azo dyes	K _i (μM)	Azo dye metabolite	Max screening concentration (μM)	K _i (μM)
		5-Amino-6-hydroxy-2-naphthalene-2-sulfonic acid 	200	>200

Supplemental Table S2.7. Inhibitor constant (K_i) values for excipient azo dyes and their metabolites cont'd.

Excipient azo dyes	K_i (μM)	Azo dye metabolite	Max screening concentration (μM)	K_i (μM)
<p>FD&C Yellow No. 6</p> 	68.4	<p>Sulfanilic acid</p> 	200	>200
		<p>5-Amino-6-hydroxy-2-naphthalene-2-sulfonic acid</p> 	200	>200
<p>Naphthol Blue Black</p> 	0.40	<p>4-Nitroaniline</p> 	200	>200
		<p>Aniline</p> 	200	>200
		<p>Unknown metabolite</p>	NA*	NA*

*NA, not available

Supplemental Table S2.8. Estimated intestinal concentrations of excipients based on maximum allowable amount and estimated daily intake.

Excipient	K _i (μM)	Max allowable amount				Estimate daily intake [†]		
		Amount	mg/day [#]	[I] (μM) ^{##}	[I]/K _i	Intake (mg)	[I] (μM) ^{##}	[I]/K _i
FD&C Red No. 40	2.59	7 mg/kg bw [*]	490	3950	1520	24.80	220	84.9
FD&C Blue No. 1	13.0	6 mg/kg bw [*]	420	2117	163	NA ¹	NA ¹	NA ¹
Sodium Lauryl Sulfate	1.98	25 mg/L ^{**}	NA	87	44	NA ¹	NA ¹	NA ¹
FD&C Red No. 3	0.88	0.1 mg/kg bw [*]	7	32	36	NA ¹	NA ¹	NA ¹
FD&C Yellow No. 6	68.4	4 mg/kg bw [*]	280	2480	36	16.70	165	2.41
Docosate Sodium	2.76	0.1 mg/kg bw [*]	7	63	23	NA ¹	NA ¹	NA ¹

*ADI, Accepted daily intake: <http://apps.who.int/food-additives-contaminants-jecfa-database/search.aspx>

**Code of Federal Regulations title 21 (CFR 21):

<https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?fr=172.822>

#Assuming adult body weight (bw) of 70 kg

##Intestinal concentration assuming excipient dissolved in 250 ml liquid (41).

¹NA, not available

[†]Data from Stevens et. al., 2013. (19).

[I], estimated intestinal concentration

2.6 Materials and Methods

Oral molecular excipient library and reagents. As of March 2017, 292 oral molecular excipients listed on the UCSF-Stanford Centers of Excellence in Regulatory Science Innovation (CERSI) Excipients Browser (<http://excipients.ucsf.bkslab.org/excipients/molecular/?route=oral>) were considered for screening (11). This collection was curated by excluding excipients that were duplicates, no longer used, experienced solubility issues, commercially unavailable, or formulated for delivery by inhalation. In total, 136 excipients were screened for OATP2B1-mediated transport inhibition. 4',5'-Dibromofluorescein (DBF) was purchased from Sigma-Aldrich (St. Louis, MO). Tritium labeled estrone sulfate was purchased from Perkin Elmer (Waltham, MA) with a specific activity of 54 Ci/mmol.

Establishment of stable OATP2B1-overexpressing cell line and cell culture. The human *OATP2B1* cDNA was cloned into Thermo Fisher Scientific pcDNA™ 5/FRT Mammalian Expression Vector (Waltham, MA) and transfected into Life Technologies Flp-In™ 293 cells using Lipofectamine LTX (Carlsbad, CA) according to the manufacturer's protocol. Stably transfected cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum, penicillin (100 U/ml), streptomycin (100 µg/ml), sodium pyruvate (110 µg/ml), and hygromycin B (100 µg/ml) at 37°C in a humidified incubator with 5% CO₂.

4',5'-Dibromofluorescein uptake kinetics assay. The assay was based on a previously reported protocol (45) with minor modifications. Cells (8x10⁴/well) were seeded in black wall poly-D-lysine-coated 96-well plates 24 hours prior to experiments. The 4',5'-dibromofluorescein (DBF) uptake kinetics study was conducted by incubating cells in

Hank's Balanced Salt Solution (HBSS) containing 0.001 μM to 30 μM DBF at 37°C for 2 min. Fluorescence in cells was measured using a fluorescence microplate reader with excitation and emission wavelengths at 485 nm and 560 nm, respectively. OATP2B1-mediated uptake of DBF was determined by subtracting background uptake from the empty vector-transfected control cells at each substrate concentration. The Z' assay factor was calculated according to the equation: $Z' = 1 - 3(\delta_{\text{control}+} + \delta_{\text{control}-}) / (\mu_{\text{control}+} - \mu_{\text{control}-})$, where δ and μ are the standard deviation and the mean, respectively (14).

Screening excipient library for OATP2B1 uptake inhibition and IC₅₀ determinations.

A total of 136 excipients were screened at 200 μM with 1% (v/v) of vehicle (DMSO), except for several surfactants and those with limited solubility, which were tested at lower concentrations. Downstream reduced metabolites of selected azo dye excipients were synthesized by Advinus Therapeutics (Pune, India) and tested at 200 μM , except for those with limited solubility, which were tested at lower concentrations (**Supplemental Table S7**). The screen was performed in triplicate with 2 μM DBF. The estrone sulfate uptake assay was conducted by incubating cells in HBSS containing 0.02 μM [³H]-estrone sulfate for three minutes at 37°C. The cells were then washed twice with ice-cold HBSS followed by lysis using lysis buffer (0.1 N NaOH and 0.1% SDS). [³H]-estrone sulfate uptake was determined by radioactivity in the lysate using a liquid scintillation counter. IC₅₀ values for compounds were determined by fitting the uptake results to the Hill equation by non-linear regression using GraphPad Prism. The inhibition constant (K_i) was estimated using the following equation (12):

$$K_i = \frac{IC_{50}}{1 + \frac{S}{K_m}} \quad (1)$$

The K_m value of DBF for OATP2B1 is 4.7 μM and S is 2 μM .

Analysis of physicochemical properties of OATP2B1 excipient inhibitors. Molecular structure files in structure-data format (SDF) were obtained from the CERSI Excipients Browser at <http://excipients.ucsf.bkslab.org/> and the PubChem database (46). MayaChemTools software (47) was used to compute eight one-dimensional and two-dimensional descriptors, such as molecular weight (MW), number of heavy atoms (HA), number of rotatable bonds (RB), molecular volume (MV), number of hydrogen bond donors (HBD), number of hydrogen bond acceptors (HBA), Crippen's log of the octanol/water partition coefficient (including implicit hydrogens) (SLogP), and total polar surface area (TPSA). Boxplots were prepared using R software. Differences between distributions of molecular descriptors were assessed by the Student's t-test using R software. The R package FSelector (48) was used to select two non-redundant and informative physicochemical properties from the eight descriptors (e.g. molecular weight and SLogP).

Aggregation test. The test for aggregation was performed according to a previously reported protocol (49). Dynamic light scattering experiments were carried out in a 384-well plate using a modified Wyatt Technology DynaPro Plate Reader II system (Santa Barbara, CA); the width of the 60 mW laser beam (at ~830 nm) was expanded to be appropriate for detecting large colloidal particles. Version 1.7 of Wyatt Technology Dynamics software was used to process the data, as well as to automatically adjust laser power and detector angle (158 degrees). The compounds were each concentrated in DMSO, subsequently diluted with filtered 50 mM potassium phosphate, pH 7.0, leading to a final concentration of 1% DMSO, and then added to the plates for measurement.

Fexofenadine pharmacokinetic studies. All animal experiments were approved by the University of California, San Francisco (UCSF) IACUC. All pharmacokinetic experiments were performed using 8-10 week-old male P-glycoprotein (Pgp) deficient (*mdr1a/b^{-/-}*) mice (Taconic Biosciences, Model#: 1487-M). Germ-free Pgp deficient mice were derived by cesarean delivery into the UCSF Gnotobiotic Core Facility and maintained in germ-free isolators. Germ-free status of mice were confirmed by monthly screening of fecal pellets by PCR for amplification of the 16S rRNA (for detection of bacteria) and Internal transcribed spacer (ITS; for detection of fungi) gene sequences. Mice were fasted overnight prior to drug and excipient oral administration. For the vehicle control group, saline was administered by oral gavage immediately followed by 15 mg/kg fexofenadine by oral administration. For test groups, FD&C Red No. 40 was administered at a concentration of 2.5 mg/kg or 25 mg/kg by oral gavage immediately followed by 15 mg/kg fexofenadine by oral administration. Blood samples (~25 µl per time point) were obtained from the mouse tail vein at 10, 30, 60, 120, 240, 360, and 480 min post drug administration. Pharmacokinetic experiments with germ-free mice were conducted in a Class II biosafety cabinet to ensure sterility and blood samples were obtained at 10, 60, 120, 240, 360, and 480 min post drug administration. Blood samples were collected in Fisherbrand heparinized glass microhematocrit capillary tubes (Waltham, MA) and centrifuged at 3500g for 15 min to recover blood plasma, which was stored at -80°C until analysis. Methods describing the analytical quantification of fexofenadine from plasma can be found below. Pharmacokinetic parameters were computed via noncompartmental analysis using Phoenix WinNonlin v8.1 (Certara, Princeton, NJ). Area under the curve

(AUC) was calculated using the linear-up, log-down method. Exact sample collection times were used for pharmacokinetic analysis.

Screening human fecal samples and bacterial isolates for azo dye metabolism in liquid cultures. Our full study protocol for the human fecal samples used in this work received approval from the University of California, San Francisco (UCSF) Institutional Review Board. Following informed consent, a single stool sample was collected from three unrelated healthy adult volunteers and frozen immediately at -80°C . On the day of the assay for bacterial dye metabolism, the fecal sample was thawed on ice. All downstream procedures and incubations were performed in a COY Laboratory Products Inc. anaerobic chamber (Grass Lake, MI) with the following atmosphere: 10% H_2 , 5% CO_2 , and 85% N_2 . All reagents were equilibrated in the anaerobic chamber for at least 24 hours before using. Three independent segments of a single stool specimen were sampled and added to supplemented BD™ Brain-Heart Infusion (BHI⁺) medium (Franklin Lakes, NJ) at a ratio of 15 ml/g wet weight. BHI⁺ is composed of: L-cysteine hydrochloride (0.05% w/v), hemin (5 $\mu\text{g}/\text{ml}$), and vitamin K (1 $\mu\text{g}/\text{ml}$). Samples were vortexed vigorously for 5 min and then allowed to settle for 10 min. Fecal suspensions were then serially diluted 10-fold in BHI⁺ medium and 70 μl of the 10^{-3} dilution was used to inoculate Hungate tubes with 7 ml of anaerobically equilibrated BHI⁺ medium supplemented with azo dye at a final concentration of 250 μM . Azo dye stocks were prepared in DMSO at a concentration of 25 mM and used at a final concentration of 1% (v/v). Naphthol blue black was excluded due to solubility issues. Cultures were incubated at 37°C . A conditioned media sample was taken at 24 hours, centrifuged to remove cells, and analyzed for residual dye concentration by measuring absorbance at 450 nm ($A_{450\text{ nm}}$). A sterile control

with no added dye (BHI⁺ medium only) was used as the background $A_{450\text{ nm}}$ and subtracted from each sample. Percent dye metabolized was calculated by dividing $A_{450\text{ nm}}$ from conditioned media samples by the $A_{450\text{ nm}}$ of the sterile control supplemented with dye (uninoculated BHI⁺ medium plus azo dye) and multiplying by 100. Human fecal bacterial isolates were screened in a 96-deep-well plate using the same media and dye concentrations described above. Overnight cultures of bacterial isolates were grown in Hungate tubes and adjusted to an $OD_{600\text{ nm}}$ of 0.1. A 15 μl aliquot of this bacterial suspension was used to inoculate 1.5 ml of dye containing BHI⁺ medium and cultures were incubated anaerobically at 37°C for 24 hours.

Isolation and characterization of azo dye metabolizing human gut bacteria. BD™ BHI⁺ agar (13.5 g/L of agar; Franklin Lakes, NJ) medium supplemented with 250 μM of azo dye were inoculated with 1 ml of a 10^{-6} fecal suspension dilution prepared as described above. The human fecal samples ranged from $1-5 \times 10^9$ CFUs/g of stool (**Fig. S4B**). Agar plates were incubated anaerobically at 37°C for 3-7 days. Colonies that showed dye metabolism (decolorized halo) were picked and restreaked on BHI⁺ agar supplemented with azo dye to confirm the phenotype.

Genomic DNA from each strain was prepared using the Qiagen DNeasy Blood & Tissue Kit (Hilden, Germany) following the manufacturer's protocol for bacteria. Dereplication of isolated bacterial strains was accomplished using Enterobacterial repetitive intergenic consensus PCR (ERIC-PCR) with the following primers: ERIC2 (5'-AAGTAAGTGAAGTGGGGTGAGCG-3') and ERIC1R (5'-ATGTAAGCTCCTGGGGATTAC-3') (26). PCR products were run on an agarose gel, and strains with identical ERIC-PCR product fingerprints, determined by eye, were

deemed replicates. The 16S rRNA gene was amplified from representative isolates using Thermo Fisher Scientific AmpliTaq Gold DNA polymerase (South San Francisco, CA) with the following primers: 8F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1391R (5'-GACGGGCGGTGWGTRCA-3) (50). Bidirectional Sanger sequencing of PCR products was performed by GENEWIZ (South Plainfield, NJ) using primers 8F and 1391R used in the PCR amplification. Species were determined by matching 16S rRNA gene sequences using BLASTN (51) searching the NCBI nucleotide collection database, Ribosomal Database Project's SeqMatch match tool (default parameters, both type and non-type strains), and SILVA Alignment, Classification and Tree Service (default settings) (52).

Genomic DNA was prepared for sequencing using Illumina Nextera XT reagents (San Diego, CA) and sequenced via the Illumina NovaSeq platform according to the manufacturer's instructions. Quality control and adapter trimming was carried out using fastp (version 0.19.5) (53), including sliding window filtering and a minimum length of 60 nucleotides. PhiX reads were removed using bowtie2 v2.3.4.1 (54). Overlaps were identified using vsearch v2.4 (55), supplied as a separate library for assembly using SPAdes v3.13.0 (56) using default coverage cutoffs, and annotated with PROKKA v1.12 using a minimum contig length of 200 bp (57). Genome sequencing statistics can be found in **Table S6**. Genome taxonomy was determined using Microbial Genome Atlas (MiGA) online NCBI Prok query. The most specific taxonomy at P-value <0.05 was chosen (58). Proteins homologous to confirmed bacterial azoreductases were identified with BLASTP (51) using previously characterized azoreductases from gut bacteria (27–31) (E-value <10⁻⁵; **Dataset S2**). We also analyzed previously sequenced reference genomes for the azo dye metabolizing bacterial species identified in this study. Reference genomes

were selected from the Joint Genome Institute (JGI) Integrated Microbial Genomes (IMG) Database (32) based on the species names of the isolates in **Table S5**. In instances where more than one reference genome existed in the IMG database for a given species, all genomes were searched and included in the results table (**Dataset S2**). Identified putative azoreductase homologs from newly sequenced and reference genomes were aligned to previously characterized azoreductase proteins using the Needleman-Wunsch algorithm with Emboss' needleall using default parameters (EMBOSS package version 6.6.0) (59).

Assaying bacterial conditioned azo dye containing media for inhibition of OATP2B1 uptake. Human gut bacterial isolates were grown as described above in the *Screening human stool samples and bacterial isolates for azo dye metabolism in liquid cultures* section. Following 48 hours of incubation, 96-well plates were centrifuged to pellet bacterial cells and the resulting conditioned media was recovered. Inhibition of OATP2B1-mediated [³H]-estrone sulfate uptake in OATP2B1-transfected HEK cells was performed as described in the above section *Screening excipient library for OATP2B1 uptake inhibition and IC₅₀ determinations* with the addition of conditioned media samples at 10% (v/v). A detailed description of the analytical analysis of conditioned media samples by LC-HRMS can be found below.

Quantification of FD&C Red No. 40 from murine fecal samples. All animal experiments were approved by the UCSF IACUC and conducted with support from the UCSF Gnotobiotic Core Facility (<http://gnotobiotics.ucsf.edu>). Conventionally-raised (CONV-R) mice were housed in a specific pathogen free (SPF) facility. Female BALB/c mice (Taconic Biosciences, Model#: BALB-F) were used between 8-10 weeks of age.

Germ-free female BALB/c mice between 8-10 weeks of age were bred in-house in sterile breeding isolators (Class Biologically Clean, USA) and transferred to experimental isolators prior to FD&C Red No. 40 dosing. Mice were fed a standard autoclaved chow diet *ad libitum* (Lab Diet 5021) with free access to water at room temperature ($24\pm 1^{\circ}\text{C}$) and a 12/12-hour light/dark cycle (7:00 a.m.-7:00 p.m.). Gnotobiotic mice were colonized by gavage of a pool of overnight cultures of the top 3 (*Faecalicoccus pleomorphus* ADE20, *Clostridium hylemonae* ADE14, and *Clostridium ramosum* ADE26) or bottom 3 (*Bifidobacterium bifidum* ADE2, *Bifidobacterium pseudocatenulatum* ADE4, and *Bifidobacterium adolescentis* ADE1) FD&C Red No. 40 metabolizers. Following colonization, gnotobiotic mice acclimated for a week before FD&C Red No. 40 dosing. Metabolic chambers were housed in experimental isolators for the germ-free and gnotobiotic animals to ensure sterility. Mice were orally dosed (0.2 ml) with 5 mg of FD&C Red No. 40 dissolved in saline and filter (0.2 μm) sterilized and immediately individually housed in a metabolic chamber for mice (Harvard Apparatus, 72-7060) for 24 hours. A control group of CONV-R mice were dosed with saline instead of FD&C Red No. 40. In addition to the 5 mg of FD&C Red No. 40 dose, CONV-R mice were also orally dosed (0.2 ml) with 25 mg/kg or 2.5 mg/kg of FD&C Red No. 40. Fecal pellets were collected and stored at -80°C . Fecal pellets were lyophilized overnight to dryness and accurately weighed. 300 μl of methanol was added and FD&C Red No. 40 was extracted for 40 min by sonication at room temperature. The supernatant was analyzed by LC-HRMS as described in the below.

Bacterial colonization levels from gnotobiotic and CONV-R mice were determined using quantitative PCR (qPCR). Total DNA was extracted from cecal and fecal samples

using the QIAamp DNA Fast Stool Mini Kit (Hilden, Germany) according to the manufacturer's protocol. qPCR of total 16S rRNA gene copies was carried out in triplicate in 10 µl reactions with 200 nM of 340F (5'-ACTCCTACGGGAGGCAGCAGT-3') and 514R (5'-ATTACCGCGGCTGCTGGC-3') primers. Reactions were performed using a BioRad CFX384 thermocycler with Life Technologies SYBRSelect for CFX Master Mix (Carlsbad, CA) following the manufacturer's protocol and using a 60°C annealing temperature. Absolute quantification was determined against a standard curve of *Escherichia coli* MG1655 genomic DNA. Mean Ct values from each sample were used for analysis. Absolute 16S rRNA gene copies were calculated by adjusting for dilutions and dividing by sample weight used for DNA extraction.

Estimation of maximum intestinal concentrations of excipients. The Accepted Daily Intake (ADI), maximum allowable amount or estimated daily intake of each excipient added for oral consumption was used to estimate the maximum intestinal concentration [I] as shown in equation 2. The FDA perspective about the role of transporters in drug-drug interactions indicates that drugs exhibiting [I]/IC₅₀ (or K_i) ≥10 should be evaluated for potential clinical drug-drug interactions (41). [I] is the theoretical maximal intestinal concentration of a drug after oral administration calculated as the highest clinical dose (mg) in a volume of 250 ml:

$$\text{Assumed maximum intestinal concentration, [I]} = \frac{\text{Amount (g)} \div \text{Molar Mass (g/mol)}}{250 \text{ ml}} \quad (2)$$

Liquid chromatography-mass spectrometry (LC-MS) analyses. Bacteria cultures grown in the presence of FD&C Red No. 40 were pelleted by centrifugation (4,000×g for

15 min) and the conditioned media was extracted 1:1 by volume with ethyl acetate. The ethyl acetate was removed by rotary evaporation and the dry extract was dissolved in 100 μ l of methanol. Liquid chromatography-high resolution mass spectrometry (LC-HRMS) analysis was performed using an Agilent G6430 triple quadrupole (QQQ) instrument with an Agilent Eclipse Plus C18 column (4.6 \times 100 mm). A linear gradient of 2-98% acetonitrile (v/v) over 30 min in H₂O with 0.1% formic acid (v/v) at a flow rate of 0.5 ml/min was used. Data was acquired under negative mode and 20 ppm mass error tolerance was used for each trace.

For analysis of fexofenadine in plasma, 4 times the volume of acetonitrile was added to plasma samples, vigorously vortexed, and centrifuged at 12,000 rpm for 5 min at room temperature. The supernatant was recovered and used for analysis. LC-HRMS analysis was performed using an Agilent Technologies 6520 Accurate-Mass Q-TOF LC-MS instrument and an Agilent Eclipse Plus C18 column (4.6 \times 100 mm). A linear gradient of 2-98% acetonitrile (v/v) over 30 min in H₂O with 0.1% formic acid (v/v) at a flow rate of 0.5 ml/min was used. Data was acquired under positive mode and 20 ppm mass error tolerance was used for each trace.

2.6 References

1. Yang H, Wang W, Romano KA, Gu M, Sanidad KZ, Kim D, Yang J, Schmidt B, Panigrahy D, Pei R, Martin DA, Ozay EI, Wang Y, Song M, Bolling BW, Xiao H, Minter LM, Yang G-Y, Liu Z, Rey FE, Zhang G. 2018. A common antimicrobial additive increases colonic inflammation and colitis-associated colon tumorigenesis in mice. *Sci Transl Med* 10.
2. Spanogiannopoulos P, Bess EN, Carmody RN, Turnbaugh PJ. 2016. The microbial pharmacists within us: a metagenomic view of xenobiotic metabolism. *Nat Rev Microbiol* 14:273–287.
3. Koppel N, Maini Rekdal V, Balskus EP. 2017. Chemical transformation of xenobiotics by the human gut microbiota. *Science* 356.
4. Bisanz JE, Spanogiannopoulos P, Pieper LM, Bustion AE, Turnbaugh PJ. 2018. How to Determine the Role of the Microbiome in Drug Disposition. *Drug Metab Dispos* 46:1588–1595.
5. Chassaing B, Koren O, Goodrich JK, Poole AC, Srinivasan S, Ley RE, Gewirtz AT. 2015. Dietary emulsifiers impact the mouse gut microbiota promoting colitis and metabolic syndrome. *Nature* 519:92–96.
6. Pharmaceutical excipients – where do we begin? NPS MedicineWise. <https://www.nps.org.au/australian-prescriber/articles/pharmaceutical-excipients-where-do-we-begin>. Retrieved 26 August 2019.
7. Varma MV, Rotter CJ, Chupka J, Whalen KM, Duignan DB, Feng B, Litchfield J, Goosen TC, El-Kattan AF. 2011. pH-sensitive interaction of HMG-CoA reductase

- inhibitors (statins) with organic anion transporting polypeptide 2B1. *Mol Pharm* 8:1303–1313.
8. Ming X, Knight BM, Thakker DR. 2011. Vectorial transport of fexofenadine across Caco-2 cells: involvement of apical uptake and basolateral efflux transporters. *Mol Pharm* 8:1677–1686.
 9. Akamine Y, Miura M, Komori H, Saito S, Kusuhara H, Tamai I, Ieiri I, Uno T, Yasui-Furukori N. 2014. Effects of one-time apple juice ingestion on the pharmacokinetics of fexofenadine enantiomers. *Eur J Clin Pharmacol* 70:1087–1095.
 10. Jeon H, Jang I-J, Lee S, Ohashi K, Kotegawa T, Ieiri I, Cho J-Y, Yoon SH, Shin S-G, Yu K-S, Lim KS. 2013. Apple juice greatly reduces systemic exposure to atenolol. *Br J Clin Pharmacol* 75:172–179.
 11. Irwin JJ, Pottel J, Zou L, Wen H, Zuk S, Zhang X, Sterling T, Shoichet BK, Lionberger R, Giacomini KM. 2017. A Molecular Basis for Innovation in Drug Excipients. *Clin Pharmacol Ther* 101:320–323.
 12. International Transporter Consortium, Giacomini KM, Huang S-M, Tweedie DJ, Benet LZ, Brouwer KLR, Chu X, Dahlin A, Evers R, Fischer V, Hillgren KM, Hoffmaster KA, Ishikawa T, Keppler D, Kim RB, Lee CA, Niemi M, Polli JW, Sugiyama Y, Swaan PW, Ware JA, Wright SH, Yee SW, Zamek-Gliszczynski MJ, Zhang L. 2010. Membrane transporters in drug development. *Nat Rev Drug Discov* 9:215–236.
 13. Izumi S, Nozaki Y, Komori T, Takenaka O, Maeda K, Kusuhara H, Sugiyama Y. 2016. Investigation of Fluorescein Derivatives as Substrates of Organic Anion Transporting Polypeptide (OATP) 1B1 To Develop Sensitive Fluorescence-Based OATP1B1 Inhibition Assays. *Mol Pharm* 13:438–448.

14. Zhang JH, Chung TD, Oldenburg KR. 1999. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J Biomol Screen* 4:67–73.
15. Khuri N, Zur AA, Wittwer MB, Lin L, Yee SW, Sali A, Giacomini KM. 2017. Computational Discovery and Experimental Validation of Inhibitors of the Human Intestinal Transporter OATP2B1. *J Chem Inf Model* 57:1402–1413.
16. Panakanti R, Narang AS. 2012. Impact of excipient interactions on drug bioavailability from solid dosage forms. *Pharm Res* 29:2639–2659.
17. Medwid S, Li MMJ, Knauer MJ, Lin K, Mansell SE, Schmerk CL, Zhu C, Griffin KE, Yousif MD, Dresser GK, Schwarz UI, Kim RB, Tirona RG. 2019. Fexofenadine and Rosuvastatin Pharmacokinetics in Mice with Targeted Disruption of Organic Anion Transporting Polypeptide 2B1. *Drug Metab Dispos* 47:832–842.
18. Dresser GK, Bailey DG, Leake BF, Schwarz UI, Dawson PA, Freeman DJ, Kim RB. 2002. Fruit juices inhibit organic anion transporting polypeptide-mediated drug uptake to decrease the oral availability of fexofenadine. *Clin Pharmacol Ther* 71:11–20.
19. Stevens LJ, Kuczek T, Burgess JR, Stochelski MA, Arnold LE, Galland L. 2013. Mechanisms of behavioral, atopic, and other reactions to artificial food colors in children. *Nutr Rev* 71:268–281.
20. Tahara H, Kusuhara H, Fuse E, Sugiyama Y. 2005. P-glycoprotein plays a major role in the efflux of fexofenadine in the small intestine and blood-brain barrier, but only a limited role in its biliary excretion. *Drug Metab Dispos* 33:963–968.
21. Haiser HJ, Turnbaugh PJ. 2013. Developing a metagenomic view of xenobiotic metabolism. *Pharmacol Res* 69:21–31.

22. Chung KT, Fulk GE, Egan M. 1978. Reduction of azo dyes by intestinal anaerobes. *Appl Environ Microbiol* 35:558–562.
23. Chung KT, Stevens SE Jr, Cerniglia CE. 1992. The reduction of azo dyes by the intestinal microflora. *Crit Rev Microbiol* 18:175–190.
24. Wang R-F, Chen H, Paine DD, Cerniglia CE. 2004. Microarray method to monitor 40 intestinal bacterial species in the study of azo dye reduction. *Biosens Bioelectron* 20:699–705.
25. Rafii F, Hall JD, Cerniglia CE. 1997. Mutagenicity of azo dyes used in foods, drugs and cosmetics before and after reduction by *Clostridium* species from the human intestinal tract. *Food Chem Toxicol* 35:897–901.
26. Versalovic J, Koeuth T, Lupski JR. 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Res* 19:6823–6831.
27. Morrison J, Dai S, Ren J, Taylor A, Wilkerson M, John G, Xie A. 2014. Structure and stability of an azoreductase with an FAD cofactor from the strict anaerobe *Clostridium perfringens*. *Protein Pept Lett* 21:523–534.
28. Chen H, Xu H, Kweon O, Chen S, Cerniglia CE. 2008. Functional role of Trp-105 of *Enterococcus faecalis* azoreductase (AzoA) as resolved by structural and mutational analysis. *Microbiology* 154:2659–2667.
29. Macwana SR, Punj S, Cooper J, Schwenk E, John GH. 2010. Identification and isolation of an azoreductase from *Enterococcus faecium*. *Curr Issues Mol Biol* 12:43–48.

30. Nakanishi M, Yatome C, Ishida N, Kitade Y. 2001. Putative ACP phosphodiesterase gene (*acpD*) encodes an azoreductase. *J Biol Chem* 276:46394–46399.
31. Wang C-J, Hagemeyer C, Rahman N, Lowe E, Noble M, Coughtrie M, Sim E, Westwood I. 2007. Molecular cloning, characterisation and ligand-bound structure of an azoreductase from *Pseudomonas aeruginosa*. *J Mol Biol* 373:1213–1228.
32. Chen I-MA, Chu K, Palaniappan K, Pillay M, Ratner A, Huang J, Huntemann M, Varghese N, White JR, Seshadri R, Smirnova T, Kirton E, Jungbluth SP, Woyke T, Elie-Fadrosh EA, Ivanova NN, Kyrpides NC. 2019. IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res* 47:D666–D677.
33. Kennedy EA, King KY, Baldrige MT. 2018. Mouse Microbiota Models: Comparing Germ-Free Mice and Antibiotics Treatment as Tools for Modifying Gut Bacteria. *Front Physiol* 9:1534.
34. Björkholm B, Bok CM, Lundin A, Rafter J, Hibberd ML, Pettersson S. 2009. Intestinal microbiota regulate xenobiotic metabolism in the liver. *PLoS One* 4:e6958.
35. Larsson E, Tremaroli V, Lee YS, Koren O, Nookaew I, Fricker A, Nielsen J, Ley RE, Bäckhed F. 2012. Analysis of gut microbial regulation of host gene expression along the length of the gut and regulation of gut microbial ecology through MyD88. *Gut* 61:1124–1131.
36. Keiser MJ, Setola V, Irwin JJ, Laggner C, Abbas AI, Hufeisen SJ, Jensen NH, Kuijjer MB, Matos RC, Tran TB, Whaley R, Glennon RA, Hert J, Thomas KLH, Edwards DD, Shoichet BK, Roth BL. 2009. Predicting new molecular targets for known drugs. *Nature* 462:175–181.

37. Sjöstedt N, Deng F, Rauvala O, Tepponen T, Kidron H. 2017. Interaction of Food Additives with Intestinal Efflux Transporters. *Mol Pharm* 14:3824–3833.
38. Gurjar R, Chan CYS, Curley P, Sharp J, Chiong J, Rannard S, Siccardi M, Owen A. 2018. Inhibitory Effects of Commonly Used Excipients on P-Glycoprotein in Vitro. *Mol Pharm* 15:4835–4842.
39. Engel A, Oswald S, Siegmund W, Keiser M. 2012. Pharmaceutical excipients influence the function of human uptake transporting proteins. *Mol Pharm* 9:2577–2581.
40. Lionberger RA. 2008. FDA critical path initiatives: opportunities for generic drug development. *AAPS J* 10:103–109.
41. Zhang L, Zhang YD, Strong JM, Reynolds KS, Huang S-M. 2008. A regulatory viewpoint on transporter-based drug interactions. *Xenobiotica* 38:709–724.
42. Bastaki M, Farrell T, Bhusari S, Bi X, Scrafford C. 2017. Estimated daily intake and safety of FD&C food-colour additives in the US population. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess* 34:891–904.
43. Mudie DM, Murray K, Hoad CL, Pritchard SE, Garnett MC, Amidon GL, Gowland PA, Spiller RC, Amidon GE, Marciani L. 2014. Quantification of gastrointestinal liquid volumes and distribution following a 240 mL dose of water in the fasted state. *Mol Pharm* 11:3039–3047.
44. Clarke G, Sandhu KV, Griffin BT, Dinan TG, Cryan JF, Hyland NP. 2019. Gut Reactions: Breaking Down Xenobiotic–Microbiome Interactions. *Pharmacol Rev* 71:198–224.

45. Erdman AR, Mangravite LM, Urban TJ, Lagpacan LL, Castro RA, de la Cruz M, Chan W, Huang CC, Johns SJ, Kawamoto M, Stryke D, Taylor TR, Carlson EJ, Ferrin TE, Brett CM, Burchard EG, Giacomini KM. 2006. The human organic anion transporter 3 (OAT3; SLC22A8): genetic variation and functional genomics. *Am J Physiol Renal Physiol* 290:F905–12.
46. Kim S, Thiessen PA, Bolton EE, Chen J, Fu G, Gindulyte A, Han L, He J, He S, Shoemaker BA, Wang J, Yu B, Zhang J, Bryant SH. 2016. PubChem Substance and Compound databases. *Nucleic Acids Res* 44:D1202–13.
47. Sud M. 2016. MayaChemTools: An Open Source Package for Computational Drug Discovery. *J Chem Inf Model* 56:2292–2297.
48. Kotthoff L. fselector. Github. <https://github.com/larskotthoff/fselector>. Retrieved 27 August 2019.
49. Coan KED, Shoichet BK. 2008. Stoichiometry and physical chemistry of promiscuous aggregate-based inhibitors. *J Am Chem Soc* 130:9606–9612.
50. Turner S, Pryer KM, Miao VP, Palmer JD. 1999. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *J Eukaryot Microbiol* 46:327–338.
51. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410.
52. Pruesse E, Peplies J, Glöckner FO. 2012. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 28:1823–1829.
53. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890.

54. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359.
55. Rognes T, Flouri T, Nichols B, Quince C, Mahé F. 2016. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4:e2584.
56. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477.
57. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069.
58. Rodriguez-R LM, Gunturu S, Harvey WT, Rosselló-Mora R, Tiedje JM, Cole JR, Konstantinidis KT. 2018. The Microbial Genomes Atlas (MiGA) webserver: taxonomic and gene diversity analysis of Archaea and Bacteria at the whole genome level. *Nucleic Acids Res* 46:W282–W288.
59. Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* 16:276–277.
60. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055.

Chapter 3: The global anaerobic metabolism regulator *fnr* is necessary for the degradation of food dyes and drugs by *Escherichia coli*

3.1 Introduction

While it has long been appreciated that the gut microbiota, the trillions of microorganisms found in the gastrointestinal (GI) tract, and its aggregate genomes (the gut microbiome) contribute to the digestion and metabolism of dietary macronutrients, the broader role of the microbiome in the metabolism of xenobiotics (diet-derived and pharmaceutical small molecules) is less well understood. Recently, work in cell culture, mice, and humans have emphasized that both excipients (food and drug additives) and pharmaceuticals are extensively metabolized by the gut microbiota (1–6). The presence of an azo bond (R-N=N-R') is shared between both pharmaceuticals and excipients (7, 8), which is notable given the ability of diverse human gut bacteria to reduce azo bonds (9). While azo reductase activity may be a core metabolic function of all human gut microbiotas (10–14), the enzymatic and non-enzymatic mechanisms responsible for this activity and their sensitivity to environmental factors remain poorly understood.

Mechanistic insights into this process are essential given the broad impact of azo reduction for antibiotics (15) and anti-inflammatory drugs (7, 16). Furthermore, the consumption of food, drug, and cosmetic (FD&C) dyes is increasing (17), providing additional substrates for gut bacterial metabolism. We recently discovered that FD&C dyes are potent inhibitors of the mammalian influx transporter OATP2B1, interfering with drug absorption in mice (18). This effect was rescued by human gut bacterial metabolism due to an inability of the downstream microbial metabolites to inhibit OATP2B1 (18). Additional work in rodent models has implicated FD&C dyes in carcinogenesis (19–21)

and inflammatory bowel disease (22). Thus, the ability to predict or control azo reductase potential within the human gut microbiome could have broad implications for host health and disease.

The canonical enzyme implicated in this metabolic activity is azo reductase (AzoR) (9, 11, 14, 23). Work on the model human gut bacterium *Escherichia coli* has demonstrated that the purified AzoR protein is sufficient to reduce azo bonds (24). Other bacteria, for example, *Pseudomonas aeruginosa* (25–28) and *Enterococcus faecalis* (29–31) encode multiple azo reductases. Alternative mechanisms have been proposed for azo reductase activity, including the electron transport chain of *Shewanella oneidensis*, nicotinamide adenine dinucleotide (NADH), and hydrogen sulfide (H₂S) (32–34). While these studies provide valuable mechanistic insight, a major limitation is their focus on *in vitro* biochemistry, neglecting to address the cellular and genetic mechanisms that impact azo reductase activity or the potential confounding effects of the complex physiological and microbiological interactions within the GI tract.

We sought to address this knowledge gap through the mechanistic dissection of a representative member of the human gut microbiota. Surprisingly, we found that the *azoR* gene is dispensable for the azo reductase activity of *E. coli*. In our search for alternative azoreductase enzymes, we discovered that *E. coli* azo reduction is regulated by the oxygen-sensing dual-transcriptional regulator, FNR. In turn, we show that FNR, via small regulatory RNA *fnrS*, regulates L-Cysteine metabolism, which produces H₂S. Surprisingly, *E. coli*-produced H₂S is sufficient to reduce azo dyes and drugs. Taken together, we demonstrate that environmental oxygen regulates the metabolism of L-Cysteine to H₂S which degrades commonly consumed azo-bonded compounds. Our

results highlight the importance of studying and developing an understanding of the environmental context and regulation of key microbial metabolisms beyond annotated enzyme function.

3.2 Results

3.2.1 *E. coli* AzoR is dispensable for azo dye depletion under anaerobic conditions.

Prior work on azoreduction in the context of the human gut and industrial settings has focused on anaerobic conditions (11, 23). We compared the depletion of a representative azo bond-containing food coloring (FD&C Red No. 40) by *E. coli* K-12 BW25113 under aerobic and anaerobic conditions, revealing that oxygen interferes with this activity (**Figure 3.1a**). L-Cysteine, a commonly used supplement in anaerobic microbiology, was also required for FD&C Red No. 40 depletion by *E. coli* (**Figure 3.1b**). Unless otherwise noted, these optimized culture conditions [Luria Broth (LB) medium plus 0.05% L-Cysteine] were used for our whole cell azo dye depletion experiments.

Previously, we (18, 35) and others (36) had identified genes homologous to *E. coli* azo reductase (AzoR) in the genomes of diverse human gut bacteria capable of depleting drugs and other azo bond containing compounds. We had originally sought to establish a tractable system to study these heterologously expressed genes by first abolishing this activity in *E. coli*. Much to our surprise, the *E. coli* Keio collection azoreductase knockout strain (Δ *azoR::kan^R*) had activity indistinguishable from wild-type (*wt*) during growth in both solid (**Figure 3.1c**) and liquid (**Figure 3.1d**) media containing FD&C Red No. 40. We validated this finding by constructing and testing a clean deletion of *azoR* (**Figure 3.1e**), which was confirmed by gel electrophoresis (**Supplemental Figure S3.1a**) and Sanger sequencing (**Supplemental Figure S3.1b**). The Δ *azoR* strain did not differ from *wt* in whole cell dye depletion kinetics (**Figure 3.1f**).

Given these unexpected results, we wondered if AzoR might be inactive in the tested *E. coli* strains, potentially due to mutations that could have occurred during the

creation of our lab's culture collection (37). Sanger sequencing of the *azoR* gene verified that the coding sequence was 100% identical to the deposited *E. coli* BW25113 genome. To further validate the activity of this enzyme, we purified the Histidine-tagged AzoR protein from *E. coli* AG1 with the pCA24N plasmid for overexpression (ASKA collection) (38). Cells were grown to mid-exponential growth phase, *azoR* expression was induced with isopropyl β -d-1-thiogalactopyranoside (IPTG), then cells were harvested and protein was purified using a nickel-charged affinity column. The fraction containing the most protein was determined using an SDS-PAGE gel and the size of the protein was consistent with AzoR (**Supplemental Figure S3.2a**). As expected, the purified protein was sufficient to clear a panel of azo bond containing dyes and drugs (**Supplemental Figure S3.2b**). Additionally, we were able to create a Michaelis-Menten curve describing the enzymatic activity against FD&C Red No. 40, where we calculated a K_M of 67.7 μ M (**Figure 3.1g**). Finally, we tested our whole cell assay using a panel of 10 food additives and pharmaceutical compounds containing azo bonds (**Figure 3.1h**), demonstrating that *azoR* is unnecessary for the depletion of multiple dyes using the growth conditions established in **Figure 3.1a-b**. These results confirm that the canonical AzoR enzyme in *E. coli* (24) is sufficient for azo reductase activity; however, *azoR* is not required for azo bond reduction by *E. coli* cells, emphasizing that biochemical activity of purified proteins is not necessarily predictive of metabolism in whole cells and motivating a renewed search for the mechanism(s) responsible.

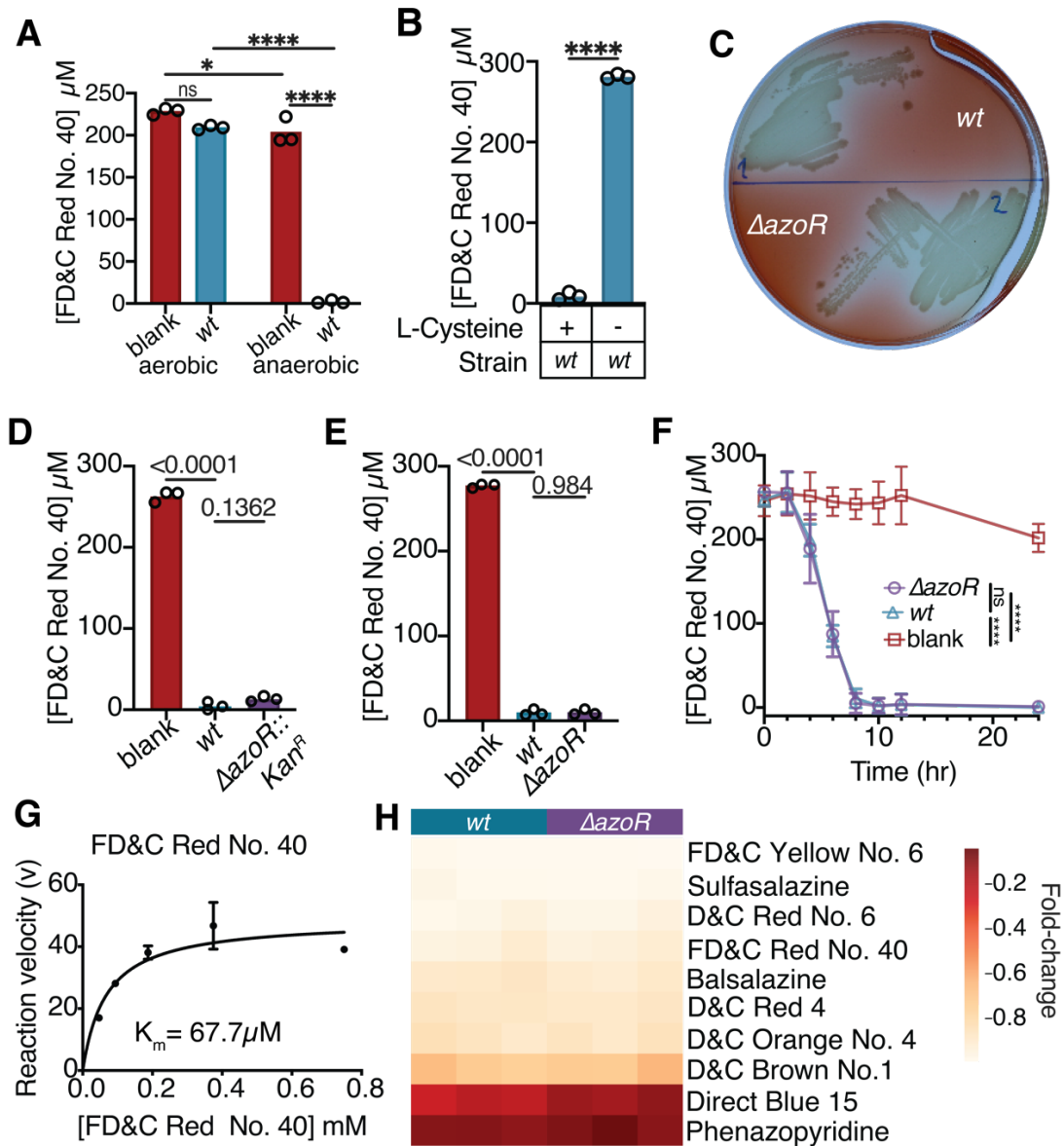


Figure 3.1 *E. coli* AzoR is sufficient but dispensable for azoreductase activity. (A) Azoreduction activity is significantly greater under anaerobic conditions. (B) Significant differences in FD&C Red No. 40 dye depletion is observed when *E. coli* is grown with or without L-Cysteine in the media. *E. coli* *azoR* is dispensable for azo reductase activity during anaerobic growth on (C) solid and (D) liquid media. (E) The knockout phenotype holds with a clean *azoR* deletion. (A-C) 250 μM FD&C Red No. 40. (C) 72 hours growth on LB agar with 4.13 mM L-Cysteine. (D,E) 24 hours growth in LB with 4.13 mM L-Cysteine. (A-B, D-E) Supernatant was removed from samples and analyzed for residual dye concentration spectrophotometrically. Concentrations calculated based on a standard curve, limit of detection=1.2 μM . (F) Azoreduction activity over a 24 hour time course. Media with 4.13 mM L-Cysteine and 250 μM FD&C Red No. 40 were inoculated with *wt* or ΔazoR and dye depletion was (Figure Caption continued on the next page)

(Figure caption continued from the previous page) monitored over time. (G) Michaelis-Menten curve for purified AzoR enzyme with FD&C Red No. 40. Michaelis-Menten curve fit using PRISM. (H) Dye depletion in liquid media by *wt* or Δ *azoR* across a panel of azo drugs and dyes. Grown with 250 μ M azo dye/drug, for 24 hours in LB with 4.13 mM L-Cysteine. Fold-changes are relative to average blank concentration. (A-B, D-H) Bars are mean. (A-B) **** $p < 0.0001$, (A, F) two-way ANOVA, (B) Mann-Whitney test, **** $p < 0.0001$. (D-E) one-way ANOVA. (A-B, D-H) $n = 3$ biological replicates/strain.

3.2.2 FNR is necessary for the depletion of azo dyes by *E. coli*.

To identify genes necessary for azo dye depletion, we screened 113 candidate gene deletions from the Keio collection using a high-throughput liquid media assay for the depletion of the widely used food coloring FD&C Red No. 40. We focused on anaerobic reductases based on the requirement for anaerobic growth (**Figure 3.2a**). We also included multiple genes based on prior studies linking them to azo dyes: genes upregulated by azo dyes in the presence of a quinone (39) or *Shewanella oneidensis* MR-1 electron transport chain that is involved in azo dye depletion (40). Gene annotations were cross referenced on EcoCyc (41) (**Supplemental Table S1**). Consistent with our previous data (**Figure 3.1**), Δ *azoR* was comparable to *wt* (**Figure 3.2a, Supplemental Table S1**). A stringent cutoff of 95% confidence interval of all genes tested revealed 22 gene deletions that had significantly impaired activity (**Figure 3.2a**). The most dramatic loss-of-function was found for Δ *fnr::Kan^R* (fumarate and nitrate reduction regulator) resulting in a 2.3-fold decrease in azo dye depletion. We validated the role of *fnr* in FD&C Red No. 40 depletion by constructing a clean deletion (**Figure 3.2b**), and confirming the deletion by gel electrophoresis (**Supplemental Figure S1C**) and Sanger sequencing (**Supplemental Figure S1D**). Activity was restored through complementation of *fnr* on the pCA24N plasmid from the ASKA collection (38) (**Figure 3.2c**).

FNR plays a key role in oxygen sensing and the transition to anaerobic metabolism (42, 43), explaining the previously observed requirement for anaerobic conditions (**Figure 3.1a**). The requirement for *fnr* was not unique to FD&C Red No. 40; significantly decreased dye depletion was observed across a panel of 7 food dyes and 3 azo bond-containing drugs (**Supplemental Figure S3**). Overexpression of AzoR was insufficient to

rescue Δfnr (Figure 3.2c), suggesting that other aspects of the vast *fnr* regulon (44–47) are responsible for the observed phenotype and/or permeability issues result in inadequate dye reaching AzoR in the cytosol (48).

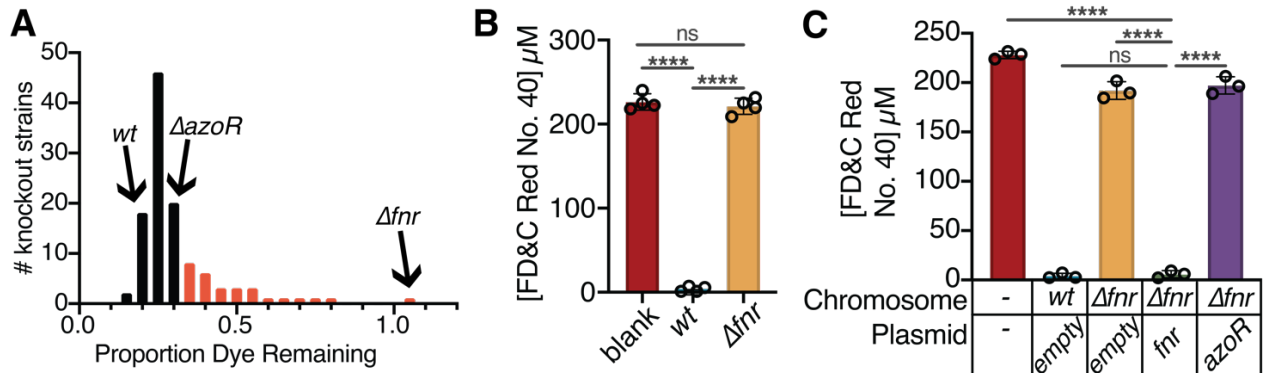


FIG 3.2 The *fnr* regulator is necessary for azoreductase activity in *E. coli*. (A) A targeted screen of 115 deletions of genes that encode enzymes and regulators revealed 21 loss-of-function strains (red bars, more dye remaining than the 95% confidence interval of the mean; $n=6$ biological replicates/strain across 2 independent experiments). The Δfnr (fumarate and nitrate reduction regulator) strain had the most extreme phenotype (complete loss-of-function). No significant gain-of-function phenotypes were observed. (B) Removal of the kanamycin cassette from Δfnr still leads to a complete loss-of-function ($n=4$ biological replicates/strain). (C) Complementation of *fnr* but not *azoR* rescues the azoreduction phenotype. Samples all induced with 0.1 mM IPTG at mid-exponential growth phase ($n=3$ biological replicates/strain). All experiments: 24 hours of growth in LB media, with 4.13 mM L-Cysteine and 250 μM FD&C Red No. 40, under anaerobic conditions. (B,C) Values are mean \pm stdev. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$, one-way ANOVA. Supernatant was removed from samples and analyzed for residual dye concentration spectrophotometrically. Concentrations calculated based on a standard curve. Limit of detection=1.2 μM .

3.2.3 Paired transcriptomics and proteomics reveals additional genes necessary for azo dye depletion.

FNR has a large regulon that has been extensively characterized by transcriptomics (RNA-seq) and Chromatin Immunoprecipitation Sequencing (ChIP-Seq) (43, 44, 46, 49). However, only 14 of the 55 genes found within the FNR regulon included in our genetic screen led to a significant loss-of-function at our 95% confidence interval (**Figure 3.2a, Supplemental Table S1**), indicating that only a subset of the FNR regulon is necessary for azo dye depletion. Prior to testing additional candidate genes, we sought to confirm and extend prior work on the FNR regulon. Our culture conditions differ from past experiments in three ways that likely alter gene expression and/or metabolic activity: (i) the presence of L-Cysteine; (ii) the inclusion of the azo dye FD&C Red No. 40; and (iii) the use of the vehicle DMSO, which is metabolized by *E. coli* under anaerobic conditions (50, 51). We also wanted to assess gene expression during stationary phase, given that FD&C Red No. 40 levels do not decrease until the exponential phase of growth (**Supplemental Figure S4**). Finally, we sought to build upon prior work on the FNR regulon by also assessing protein levels using proteomics.

We grew *wt* and Δfnr *E. coli* in LB media with 0.05% L-Cysteine and inoculated the cultures with 250 μ M FD&C Red No. 40 or DMSO vehicle at either mid-exponential or stationary growth phases (**Figure 3.3a**). After a 40 minute incubation period, samples were collected and split for paired transcriptomics (RNA-seq) and proteomics analysis. In total, we generated 12.4 ± 6.2 million high quality reads/sample (RNA-seq; **Supplemental Table S2**) and 881 ± 23 unique peptides after matching spectra through the MSGF+

database (52) and adjusting for false discovery rate of 0.01 against the MSGF generated decoy database (proteomics; **Supplemental Table S3**).

Principal coordinates analysis revealed a clear separation of profiles between strains at the RNA (**Figure 3.3b**) and protein (**Figure 3.3c**) levels. As expected (53–55), we also saw a clear difference in expression profile between growth phases; however, the separation between strains was maintained at both timepoints (**Figure 3.3b,c**). There was a more minimal impact of FD&C Red No. 40 (**Figure 3.3b,c**).

The number of differentially expressed genes was consistent with the overall trends in these datasets, revealing marked strain- and growth-phase differences in transcript and protein levels, with a minimal impact of FD&C Red No. 40 (**Figure 3.3d-e**, **Supplemental Table S4**). Only a single gene was significantly upregulated by FD&C Red No. 40 (\log_2 fold-change > 1 and p_{adj} < 0.1): *ydeN*, a putative sulfatase (56, 57). The $\Delta ydeN$ strain had comparable azo reductase activity to *wt* (**Supplemental Figure S5**). In the proteomics dataset 3 features were enriched in the vehicle group (**Figure 3.3e**).

Given the marked difference in azo dye depletion between the *wt* and Δfnr strains, we next focused on the 588 transcripts and 216 proteins significantly different between strains with a \log_2 fold-change > |1| (p_{adj} < 0.1; **Figure 3.3f-h**). 83 features were both significant and had a fold-change in the same direction in both datasets (**Figure 3.3h**). Notably, these consistent features included 62 enzymes and 18 [oxido]reductases (**Supplemental Table S5**). *azoR* was slightly significantly upregulated in the *wt* strain (**Figure S6A**), with a trend towards increased peptide intensity (**Supplemental Figure S6B**). As expected, *fnr* transcripts were undetectable in the Δfnr strain (**Supplemental**

Figure S6C). Furthermore, 96/113 genes used in our azo dye depletion screen (**Figure 3.2a**) had detectable expression (**Supplemental Table S1**).

Given the size of the FNR regulon in both prior studies (43, 44, 46, 49) and our datasets, including numerous enzymes that could feasibly impact azo dyes, we sought to further narrow down our candidate gene list by focusing on FNR-dependent regulators. Features that were significantly increased in both RNA-seq and proteomics in *wt* compared to Δfnr (\log_2 fold-change $>|1|$, $p_{adj} < 0.1$) were analyzed for direct regulators using EcoCyc (see *Methods*). This analysis revealed 11 regulators whose regulons were significantly enriched in our gene set (**Figure 3.3i**). As expected, *fnr* was the most significantly enriched regulator; however, the proportion of the FNR regulon represented was lower than some of the other regulators due to the much larger size of the FNR regulon and our stringent criteria for identifying FNR-dependent genes.

We went on to test the azoreduction phenotype of Keio collection knockouts of all 11 regulators (**Figure 3.3j**). Two regulators that were not in our original screen, *cyuR* (detoxification of L-Cysteine regulator) and *crp* (cAMP-activated global transcriptional regulator), demonstrated a significant loss-of-function compared to *wt* ($p < 0.0001$, one-way ANOVA). Crp regulates >100 genes during both aerobic and anaerobic growth (58–61). In contrast, *cyuR* regulates just 2 genes involved in L-Cysteine metabolism: *cyuA* (L-Cysteine desulfidase) and *cyuP* (L-Cysteine utilization permease) (62–64). Given the tractable size of the *cyuR* regulon and emerging evidence that L-Cysteine metabolism can enable degradation of azo bonds (33, 34, 65), we opted to focus on *cyuA* and *cyuP*.

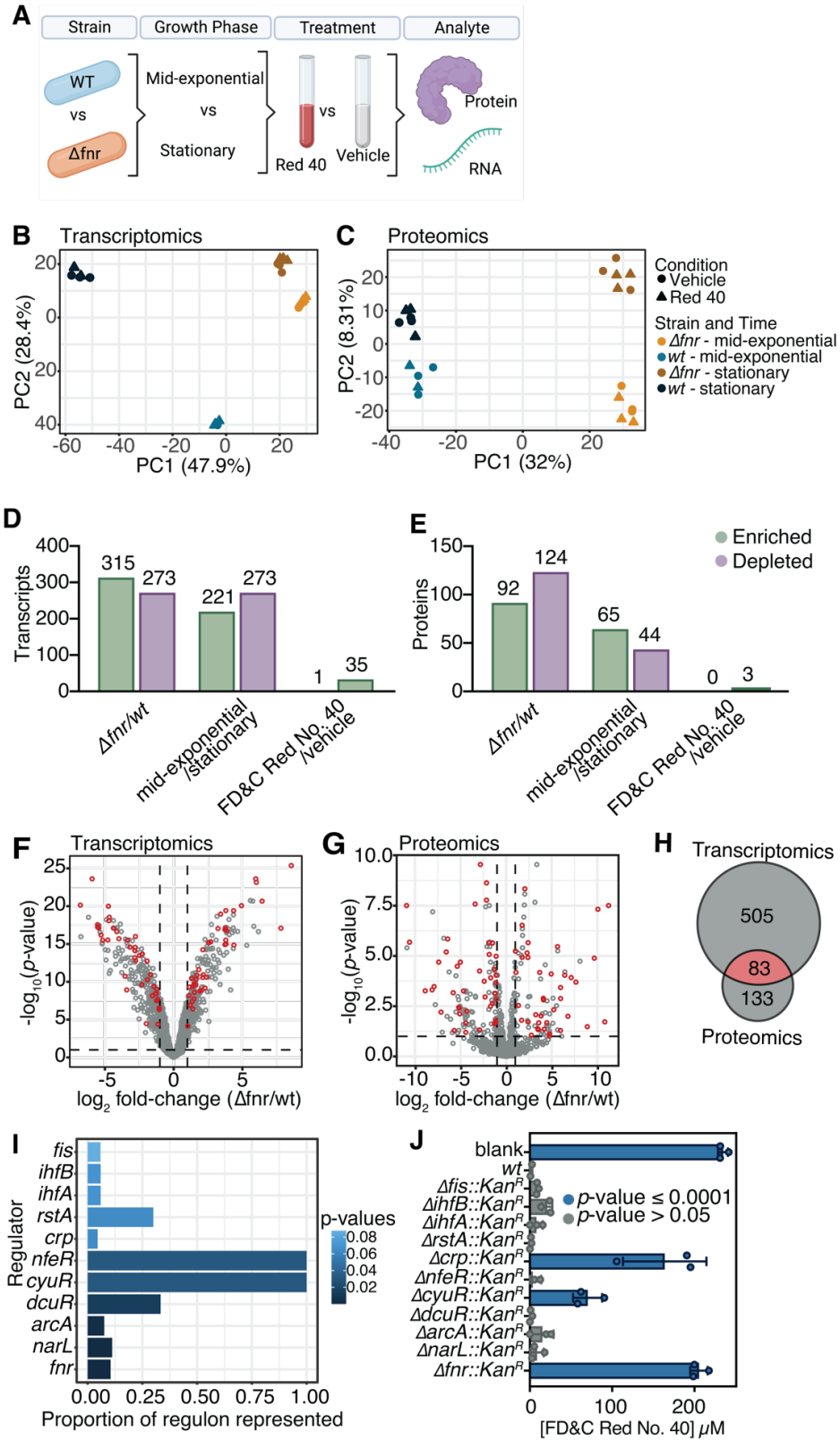


Figure 3.3 Multi-omic analysis of the *fnr* regulon reveals a complex regulon and the importance of L-Cysteine metabolism. (Figure Caption continued on the next page)

(Figure caption continued from the previous page) (A) Experimental design. *wt* and Δfnr strains were grown to mid-exponential or stationary phase, then dosed with 250 μ M FD&C Red No. 40 or vehicle control. After 45 mins, cultures were spun down and fractions taken for proteomics or transcriptomics processing. Created with BioRender.com. (B-C) Principal components analysis (PCA) of Euclidean distances for Transcriptomics (B) and Proteomics (C) datasets. Each point represents one sample from the experiment. (D,E) Barplots representing the number of features enriched or depleted ($p_{adj} < 0.1$, \log_2 fold-change $> |1|$) for each comparison for Transcriptomics (D) and Proteomics (E) datasets. (F,G) Volcano plots for Transcriptomics (F) and Proteomics (G) datasets. Each point represents the average expression of one transcript or protein, respectively. Differential expression is based on the strain-level comparison. Significantly differentially expressed features ($p_{adj} < 0.1$, \log_2 fold-change $\geq |1|$) are indicated by dashed lines, comparing Δfnr to *wt*. (H) Venn diagram of the number of differentially expressed features ($p_{adj} < 0.1$, \log_2 fold-change $> |1|$) in the two datasets. (I) Regulons enriched in the overlapping features of transcriptomics and proteomics datasets ($p_{adj} < 0.1$, with Benjamini-Hochberg correction) of differentially expressed genes ($p_{adj} < 0.1$, \log_2 fold-change $> |1|$). (J) Azo depletion activity by enriched regulator knockout strains from the Keio Collection. Values are mean \pm stdev.

3.2.4 L-Cysteine metabolism is required for depletion of azo dyes.

L-Cysteine metabolism by *E. coli* generates hydrogen sulfide (H₂S) as a dead-end metabolite (**Figure 3.4a**) (63, 66), which has been previously implicated in azo dye degradation (33, 34, 65). Thus, we hypothesized that *fnr* controls H₂S production and thus azo bond degradation by increasing the expression of the *cyuR*-regulated *cyuA* and *cyuP*. First, we tested the requirements for L-Cysteine and H₂S in azoreduction. To facilitate rapidly establishing anaerobic conditions, we routinely supplement our media with L-Cysteine (67–69). Removal of L-Cysteine impaired the ability of *E. coli* to clear FD&C Red No. 40 (**Figure 3.1b**). L-Cysteine promoted FD&C Red No. 40 depletion in a dose-dependent manner (**Figure 3.4b**), suggesting that L-Cysteine is a key substrate for azo reductase activity during *in vitro* growth under anaerobic conditions. Anaerobic conditions in LB media lacking L-Cysteine, equilibrated in the anaerobic chamber for one week were confirmed by eye using the indicator resazurin. Notably, L-Cysteine alone was unable to deplete FD&C Red No. 40 (**Figure 3.4c**), while addition of sodium sulfide demonstrated a dose-dependent reduction in sterile media (**Figure 3.4d**). To test the relationship between L-Cysteine and H₂S, we quantified H₂S in whole cell culture using the methylene blue assay (see *Methods*). H₂S levels were below our limit of detection (1.3 μM) in the Δfnr strain, significantly less than the triple-digit μM levels produced by *wt* (**Figure 3.4e**). We also quantified the amount of FD&C Red No. 40 depleted in these samples and saw the inverse trend (**Figure 3.4f**). Consistent with the importance of L-Cysteine metabolism for dye depletion, peak H₂S concentration corresponded to the decrease in FD&C Red 40 for *wt* with Δfnr stable over time (**Figure 3.4g**).

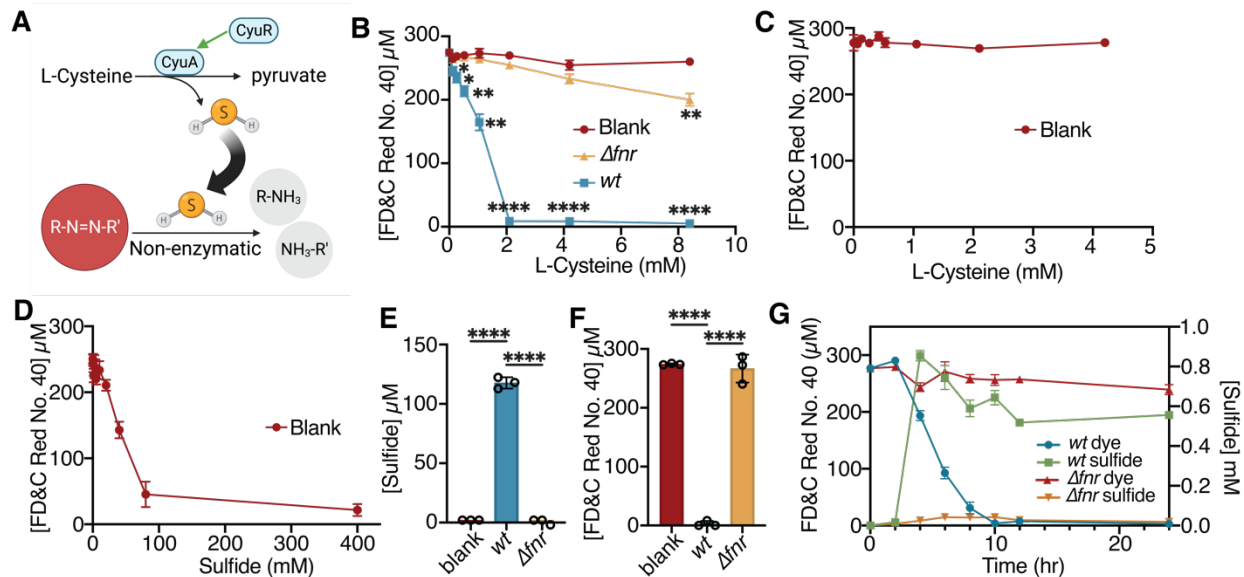


Figure 3.4. The *fnr* regulator is necessary for the metabolism of L-Cysteine to hydrogen sulfide, enabling the reduction of azo bonds. (A) Schematic depicting L-Cysteine metabolism pathway in *E. coli* which metabolizes L-Cysteine to pyruvate, releasing hydrogen sulfide. Created with BioRender.com. (B) *wt* and Δfnr *E. coli* was incubated in decreasing concentrations of L-Cysteine. Statistics relative to blanks. (C) Increasing L-Cysteine concentrations in LB with 250 μ M FD&C Red No. 40. (D) Increasing Sodium Sulfide concentrations in LB with 250 μ M FD&C Red No. 40 with 0.05% L-Cysteine. (E) Hydrogen sulfide production by *wt* and Δfnr *E. coli*. (F) FD&C Red No. 40 depletion by *wt* and Δfnr *E. coli*. (G) Red 40 depletion and sulfide measurements over time in *wt* (blue/green) and Δfnr (red/orange) *E. coli*. (B, E, F) * p <0.05, ** p <0.01, *** p <0.001, **** p <0.0001, one-way ANOVA. All experiments: values are mean \pm stdev. 24 hours of growth in LB media, with 4.13 mM L-Cysteine and 250 μ M FD&C Red No. 40 unless otherwise noted. Supernatant was removed from samples and analyzed for residual dye concentration spectrophotometrically. Concentrations calculated based on a standard curve. Limit of detection = 1.2 μ M (n=3 biological replicates).

Next, we investigated if H₂S alone was the *fnr*-driven azoreductase mechanism for FD&C Red No. 40 depletion. We utilized single and multi-gene knockouts for the *cyuR* regulon. As previously mentioned, CyuR regulates both the L-Cysteine transporter (*cyuP*) and L-Cysteine desulfidase (*cyuA*) (63) (**Figure 3.5a**). Consistent with our hypothesis, *cyuA*, *cyuP*, and *cyuR* were significantly decreased at the transcript level in Δfnr *E. coli* relative to *wt* controls (**Figure 3.5b**, **Supplemental Figure S6d**). We validated these findings using RT-qPCR (**Supplemental Figure S6e-g**). The *cyuP* and *cyuR* knockout strains produced significantly less sulfide (**Figure 3.5c**) and depleted significantly less FD&C Red No. 40 (**Figure 3.5d**) in L-Cysteine-containing media. The combined deletion of *cyuR* and *azoR* did not significantly impact sulfide levels (**Figure 3.5e**); however, there was a slight but statistical significance impairment in azo dye depletion (**Figure 3.5f**). The addition of 0.05% L-Cysteine at mid-exponential phase increased the expression of *cyuR*, *cyuA*, and *cyuP*; *cyuR* and *cyuP* were induced significantly higher in *wt* relative to Δfnr (**Supplemental Figure S7a-d**).

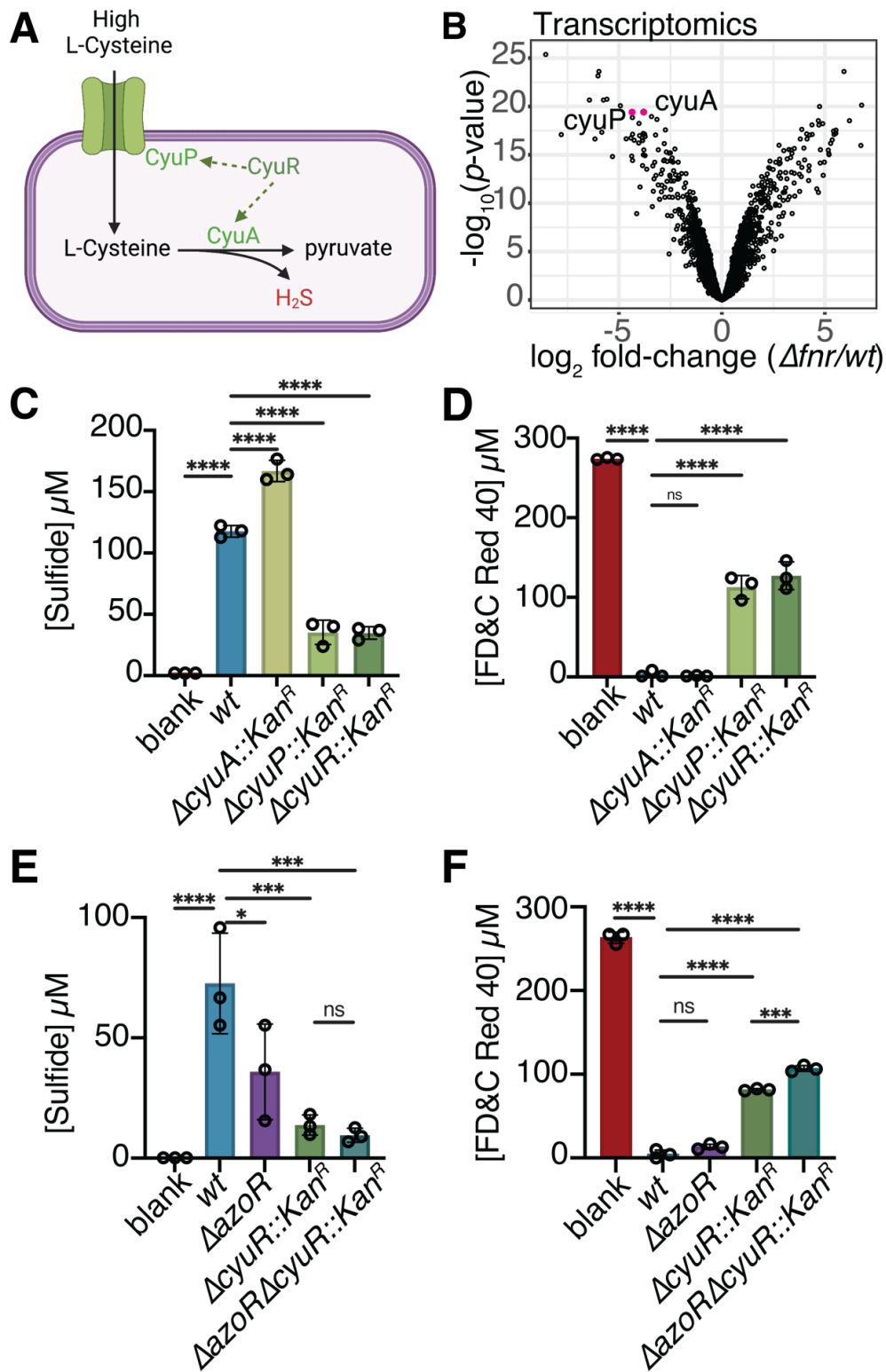


Figure 3.5 L-Cysteine import and metabolism genes impact bacterially produced sulfide and FD&C Red No. 40 depletion. (Figure Caption continued on the next page)

(Figure caption continued from the previous page) (A) Model of the *cyuR* regulon's role in L-Cysteine metabolism. CyuP acts as a L-Cysteine transporter into the cell, while CyuA is a L-Cysteine desulfidase. Created with BioRender.com. (B) Volcano plot for transcriptomics data: *cyuA* and *cyuP* are highlighted in pink. Each point represents the average expression of one transcript. Differential expression is based on the strain-level comparison. (C-D) Paired sulfide (C) and FD&C Red No. 40 (D) were measured from *cyuR* regulon knockouts and *cyuR* (Keio collection). (E-F) Paired sulfide (E) and FD&C Red No. 40 (F) were measured from $\Delta azoR$, $\Delta cyuR::Kan^R$, and $\Delta azoR\Delta cyuR::Kan^R$ strains. Values are mean \pm stdev. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$, one-way ANOVA. (C-F) $n = 3$ replicates/strain.

3.2.5 FNR indirectly regulates the *cyu* genes.

Finally, we sought to better understand the mechanism through which FNR impacts the expression of *cyuR* and its target genes *cyuAP*. FNR was not detected at the *cyuR* promoter in chromatin immunoprecipitation experiments (44), consistent with the lack of the FNR binding motif TTGATnnnnATCAA (44). These results led us to hypothesize that FNR indirectly regulates the *cyu* genes. A review of the prior literature and the EcoCyc database identified multiple potential intermediates that are both under the control of FNR and affect *cyu*: the MarA and SoxS transcription factors (70) and the regulatory small non-coding RNA *fnrS* (71, 72) (**Figure 3.6a**). MarA and SoxS both regulate *cyuR* (70), while *fnrS* is a proposed positive regulator of *cyuP* (72). Neither *marA* nor *soxS* complemented the Δfnr phenotype (**Supplemental Figure S7e-f**). Furthermore, the $\Delta marA::Kan^R$ and $\Delta soxS::Kan^R$ strains were not significantly different from *wt* (**Supplemental Figure S7g**). Together, these experiments suggest that *marA* and *soxS* are not involved in the *fnr*-dependent FD&C Red 40 depletion or H₂S production.

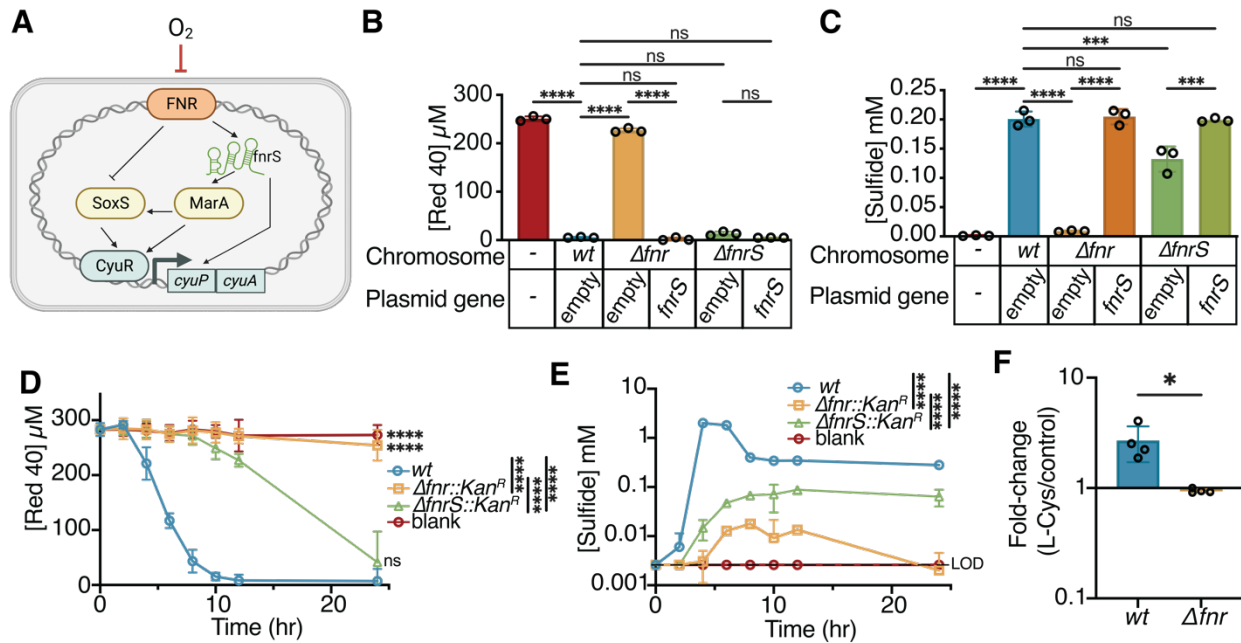


Figure 3.6 FNR indirectly regulates sulfide production and FD&C Red No. 40 depletion predominantly via small RNA *fnrS*. (A) Model of the potential regulation cascade from FNR to the expression of *cyuAP*. Created with BioRender.com. (B-C) Paired FD&C Red No. 40 (B) and sulfide (C) were measured from *wt*, $\Delta fnr::Kan^R$ complemented with pBD-lac or pBD-lac-*fnrS*, and $\Delta fnrS::Kan^R$ complemented with pBD-lac or pBD-lac-*fnrS*. (D-E) Paired FD&C Red No. 40 (D) and sulfide (E) were measured from *wt*, $\Delta fnr::Kan^R$, and $\Delta fnrS::Kan^R$ overtime. Samples inoculated at time point zero. (E) Plotted on a log₁₀ scale y-axis. (F) Fold-change of *fnrS* in the presence of L-Cysteine compared to vehicle at mid-exponential growth. All values are mean \pm stdev. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. (B-C) One-way ANOVA (n=4 replicates/strain). (D-E) Two-way ANOVA for overall comparisons and Two-way ANOVA comparing *wt* to each group to final time point (n=3 replicates/strain). (F) Mann-Whitney test (n=4 biological replicates averaged from 3 technical replicates/strain).

In contrast, we found that the small non-coding RNA *fnrS* is critical for these phenotypes. We were able to complement the Δfnr dye and sulfide phenotypes with *fnrS* (Figure 3.6b-c). While complementation was sufficient to rescue the Δfnr phenotype, the $\Delta fnrS::Kan^R$ carrying the empty plasmid did not show a loss-of-function phenotype at our usual sampling time of 24 hrs (Figure 3.6b-c). We hypothesized that this is likely due to

alternative regulators activating *cyuAP* and other L-Cysteine metabolism genes. Thus, we measured FD&C Red 40 degradation and sulfide production in the knockout strain over time. *fnrS::Kan^R* demonstrated a delayed onset of dye degradation and lower sulfide production compared to *wt* (**Figure 3.6d-e**). Furthermore, *fnrS* expression was induced significantly more in response to L-Cysteine in *wt* relative to Δ *fnr* *E. coli* (**Figure 3.6f**).

3.3 Discussion

Our results provide mechanistic insight into the degradation of azo dyes by human gut bacteria. Surprisingly, the canonical azoreductase enzyme of *E. coli* is dispensable for whole cell azoreductase activity. Instead, the shift to anaerobic growth is critical. We dissected a pathway through which oxygen is sensed by FNR, leading to upregulation of the *cyu* operon, increased uptake of L-Cysteine, enhanced production of H₂S, and non-enzymatic azo bond reduction (**Figure 3.7**).

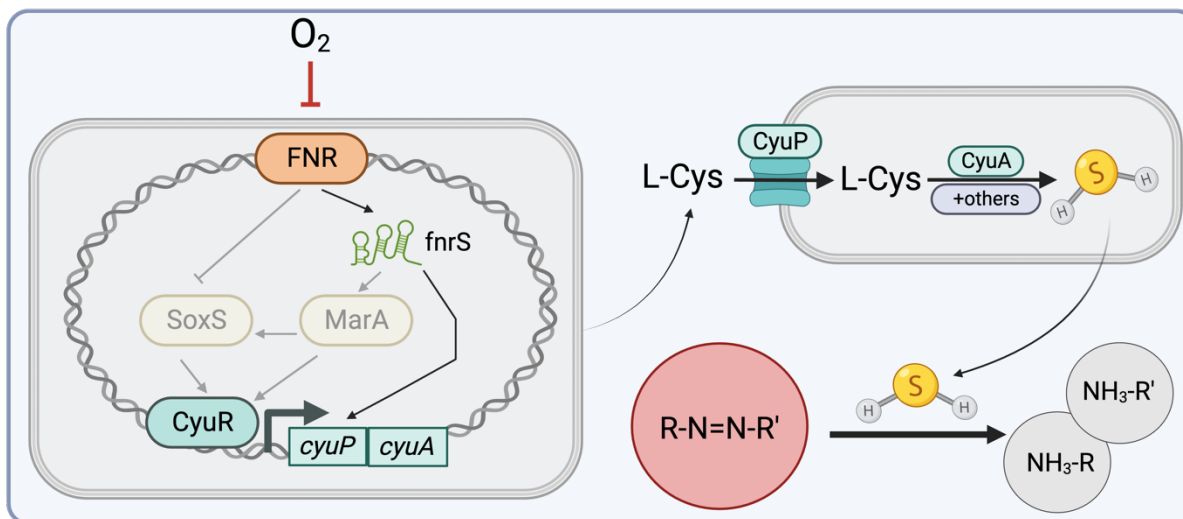


Figure 3.7 Working model describing how the host environment shapes gut bacterial azo reduction. Under anaerobic conditions, FNR activates the transcription of *fnrS*, a small non-coding regulatory RNA. *fnrS* in turn activates expression of the *cyuAP* operon. Alternatively, *fnrS* may upregulate the expression of *marA* which can drive expression of both *cyuR* and *soxS*. *SoxS* can also induce *cyuR* expression. *CyuP* transports L-Cysteine into the cell, where it is metabolized to H₂S through the *CyuA* desulfidase. Released H₂S can reduce the azo bond of azo compounds. Created with BioRender.com.

These findings provide a cautionary tale for the ability to extrapolate results from biochemical studies of purified proteins to infer the metabolic activity of cells and complex microbial communities. Despite a rich literature on *E. coli* AzoR (24, 48, 73), we report that the *azoR* gene is dispensable for azo reduction by *E. coli* during anaerobic growth. These findings emphasize the importance of pairing biochemical and heterologous expression studies of enzyme function with methods for gene deletion in the original strain. Future work should extend this approach to the other bacterial species with biochemically characterized azoreductase enzymes (25, 29, 74, 75). For species where genetic tools do not exist (the vast majority of human gut bacteria), one could leverage natural strain-level variation to test the association between gene presence and sequence variation with metabolic activity (37).

The global regulator FNR is essential for *E. coli* azoreduction, highlighting the importance of environmental oxygen for the gut bacterial metabolism of drugs and excipients. Confirming the prior work on FNR in different growth conditions (42, 61), we confirmed that the FNR regulon is vast, including hundreds of genes involved in different aspects of anaerobic metabolism. Interestingly, efforts to utilize bacterial azoreduction to degrade dye runoff in water systems also require anaerobic conditions, suggesting that our findings may be generalizable to microbial communities outside of the human gut (23, 76, 77). Importantly, *E. coli* exhibits features of both aerobic and anaerobic growth within the GI tract (61, 78), depending upon its physical niche and host pathophysiology (79, 80). Our results suggest that inter-individual variations in the environment in which *E. coli* grows could have downstream consequences for its ability to act upon azo bond containing compounds. More work is necessary to assess how the redox potential of the

gut lumen and other environmental factors shape gut bacterial azo reduction in the context of host health and disease.

Surprisingly, FNR controls azo reductase potential through the reduction of azo bonds by H₂S. Given that L-Cysteine is the major source of H₂S for *E. coli*, these results provide an additional layer of environmental control for azoreduction. Differences in dietary protein, host proteoglycans, and/or L-Cysteine biosynthesis by other members of the gut microbiota could all potentially influence the activity of *E. coli* and its ability to act on azo bonds. Future studies leveraging gnotobiotic mice, defined dietary interventions, and isotope tracing could be informative in this regard.

Our results indicate that the L-Cysteine utilization regulator *cyuR*, among others, is under indirect FNR control, emphasizing the complex multi-level regulatory network contributing to bacterial azoreduction. The *cyu* operon is the main source for L-Cysteine import and metabolism by *E. coli* (63, 81). We also found that H₂S production requires *cyuR*. Interestingly, the L-Cysteine uptake transporter, CyuP, was more important than the enzyme CyuA, suggesting that the import of L-Cysteine could provide substrate for additional pathways for H₂S production. CyuA only functions anaerobically (63), consistent with our observations that azoreduction only occurs anaerobically and is regulated by FNR and CyuR. Moreover, the key role of L-Cysteine import suggests that amino acid availability from diet and/or host tissues could be critical for gut bacterial azoreduction. Future studies with mouse models and various dietary amino acid interventions could provide an opportunity to modulate microbiome azoreductase activity.

Deletion of both *cyuR* and *azoR* did not fully abolish *E. coli* azoreduction. Thus, additional pathways for L-Cysteine metabolism and/or as-of-yet undiscovered enzyme(s)

may contribute to the metabolism of azo dyes by *E. coli*. In our initial screen, we found 21 deletion strains with a partial loss-of-function phenotype, all of which encode enzymes theoretically capable of azo reduction. Of these, 4 were also identified in our paired proteomic and transcriptomic dataset (**Table S6**), which could be viable candidates for further biochemical characterization and/or complementation screens in the Δfnr background.

H₂S production is widespread throughout gut bacterial species (82), suggesting that far more bacterial taxa are capable of azo reduction than previously appreciated. L-Cysteine can be made from a variety of different sources, thus methionine and other sulfur starting points could impact the gut microbiota's H₂S production capacity. Multiple *Desulfovibrio* species found in the GI tract can reduce sulfate to H₂S (83–86). Thus, GI levels of H₂S represent the net effect of multiple distinct pathways, which will be important to consider when translating these results into humans. As a first step, studies in gnotobiotic mice with defined microbial communities representing distinct assemblages of H₂S producers would help to better understand the relative importance of each species for azo dye degradation under different dietary and host selection pressures.

There are multiple limitations of this study. We focused entirely on the model gut Proteobacterium *E. coli* due to its genetic tractability and long history of molecular research; however, the generalizability of the pathway we described to other gut bacteria remains unclear. The development of genetic tools that permit similar mechanistic dissections into these more exotic gut bacteria will be essential for studying the conservation of enzymatic and non-enzymatic azo bond degradation in diverse gut bacterial species. The physiological relevance of our findings remains unclear, requiring

more work in mice or other preclinical models. It will also be important to test the clinical relevance of these findings in the context of diseases intentionally treated with azo bond containing drugs (e.g., inflammatory bowel disease) or diseases like cancer that could potentially be exacerbated by dietary exposures to azo dyes. Our results emphasize that in addition to metagenomics, it will be critical to measure key environmental factors in human clinical cohorts like oxygen and amino acid levels.

Additionally, our results favor the hypothesis that FNR indirectly regulates *cyuR* and in turn *cyuAP*. While our data indicates that *fnrS* is responsible for the majority of this indirect regulation, these results do not address the combinatorial impact of all intermediate regulators. Ideally, a quadruple knockout of *marA*, *soxS*, *fnrS*, and *cyuR* would be necessary to confirm these are the only pathways leading towards *cyuAP* expression and thus H₂S production. These results again highlight the complex regulation that underlies metabolic pathways and the importance of taking a nuanced approach to understanding not only gene product function but regulatory factors.

Despite these limitations, our results clearly emphasize the importance of environmental factors in controlling a clinically relevant metabolic activity of the human gut microbiome. These experiments are particularly timely given the rapid increase in food dye intake due to their incorporation in processed foods (87) and emerging evidence that they have unintended consequences for gut epithelial (18) and immune (22) cells. Continued progress towards dissecting the mechanisms responsible for this metabolic activity in *E. coli* and other experimentally tractable gut bacterial species is a key step towards predicting and controlling the location and extent of azo dye metabolism.

3.4 Supplemental Figures

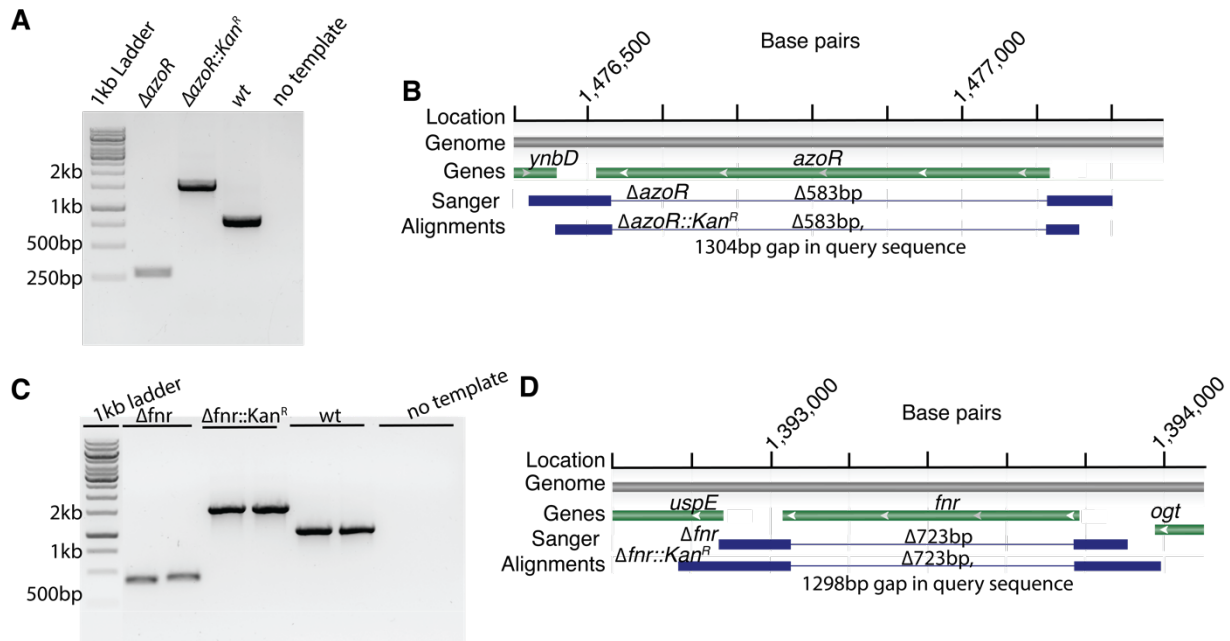


Figure S3.1 Validation of bacterial strains. (A) Gel electrophoresis of PCR product from $\Delta azoR$ strains used: $\Delta azoR$ and $\Delta azoR::Kan^R$ compared to wild-type (*wt*). (B) Alignment of the $\Delta azoR$ and $\Delta azoR::Kan^R$ Sanger sequencing results to the *E. coli* K12 BW21153 genome. (C) Gel electrophoresis of PCR product from Δfnr strains used: Δfnr and $\Delta fnr::Kan^R$ compared to wild-type (*wt*). (D) Alignment of the Δfnr and $\Delta fnr::Kan^R$ Sanger sequencing results to the *E. coli* K12 BW21153 genome.

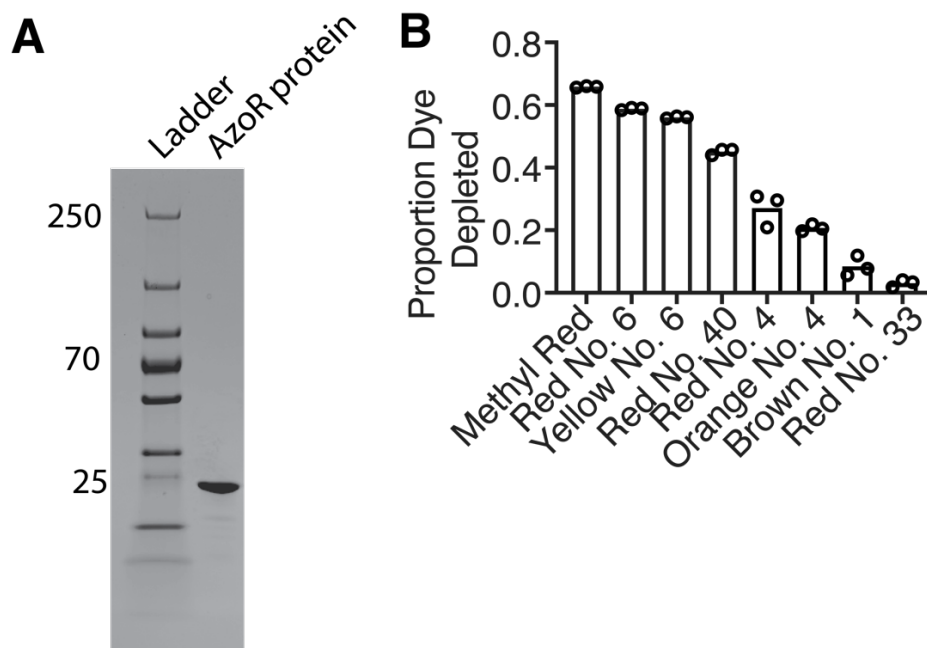


Figure S3.2 *E. coli* AzoR is sufficient to deplete azo dyes. (A) Purified AzoR protein from *E. coli* AG1 has the expected size (23 kDa) on an SDS-PAGE gel. (B) AzoR protein has variable impacts on the depletion of a panel of FD&C dyes. Incubations were conducted in PBS buffer with 2000 μ M NADH, and 20 μ M FMN at 37°C for 15 min, then proportion of the dye depleted was measured by spectrophotometric and calculated relative to no enzyme controls (n=3 biological replicates/strain).

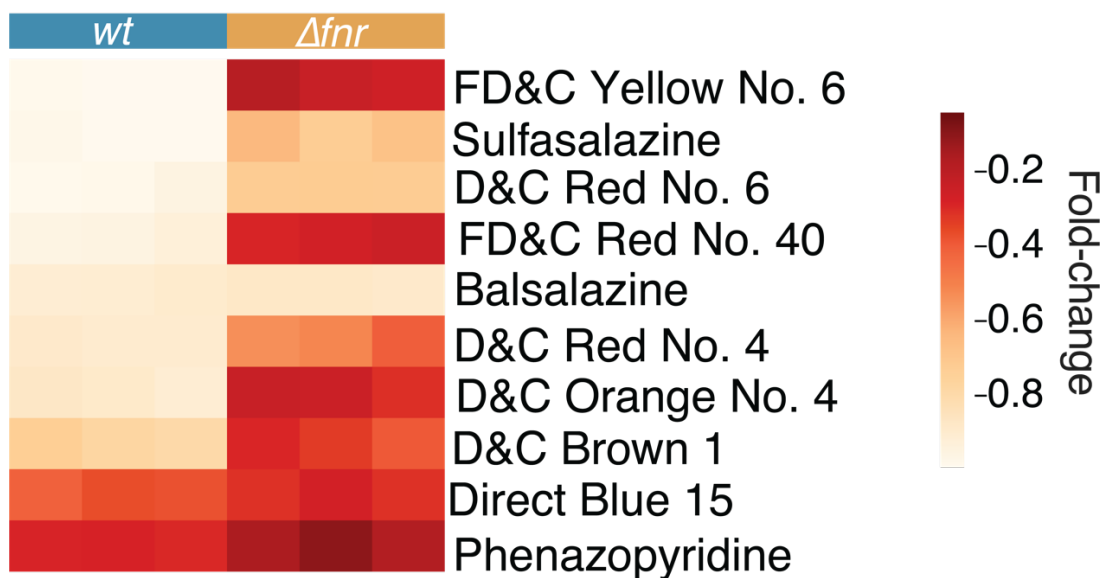


Figure S3.3 *fnr* is necessary for the depletion of multiple azo dyes. Dye depletion was assessed by spectrophotometry at appropriate wavelengths following 24 hours of incubation in LB media with 4.13 mM L-Cysteine and 250 μ M azo dye/drug. Values are fold-change relative to the mean blank concentration (n=3 biological replicates/strain).

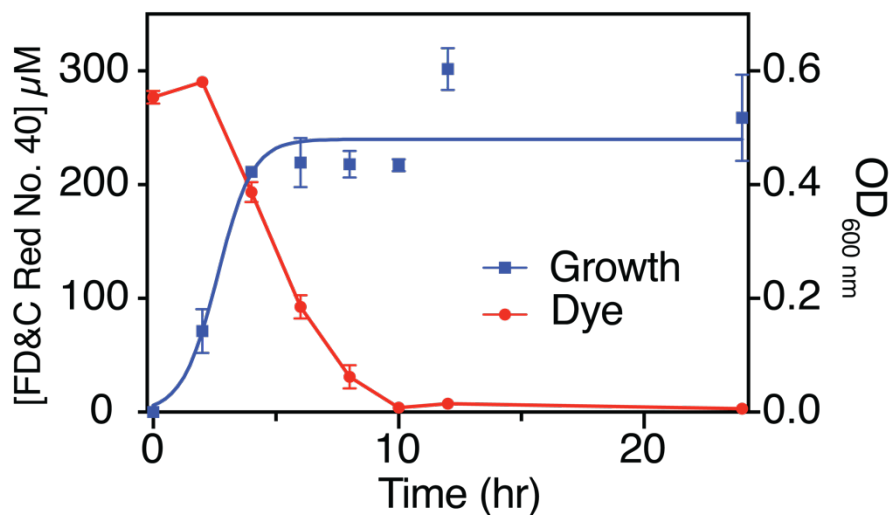


Figure S3.4 Dye depletion corresponds to the growth phase of *wt E. coli*. FD&C Red No. 40 depletion measured over time (red) and corresponding cell density (blue, OD_{600 nm}). Blue squares represent mean OD_{600 nm} measurement, while the blue line represents the growth curve modeled off individual measurements. Samples were grown in LB with 250 μM FD&C Red No. 40 and 4.2 mM L-Cysteine under anaerobic conditions. n=3 biological replicates.

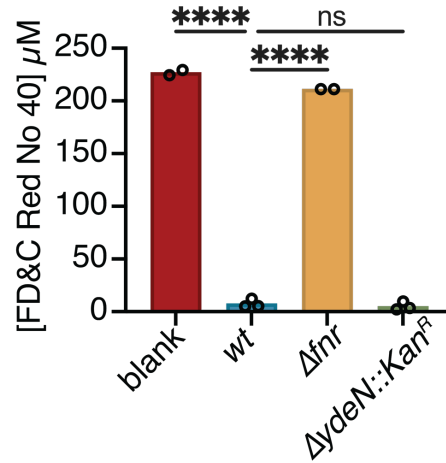


Figure S3.5. *ydeN* is not required for azo dye depletion. Samples grown in LB with 4.2 mM L-Cysteine and 250 μM FD&C Red No. 40 for 24 hours under anaerobic conditions. Values are mean \pm stdev. **** p <0.0001, one-way ANOVA (n=2-3 biological replicates).

-omics datasets:

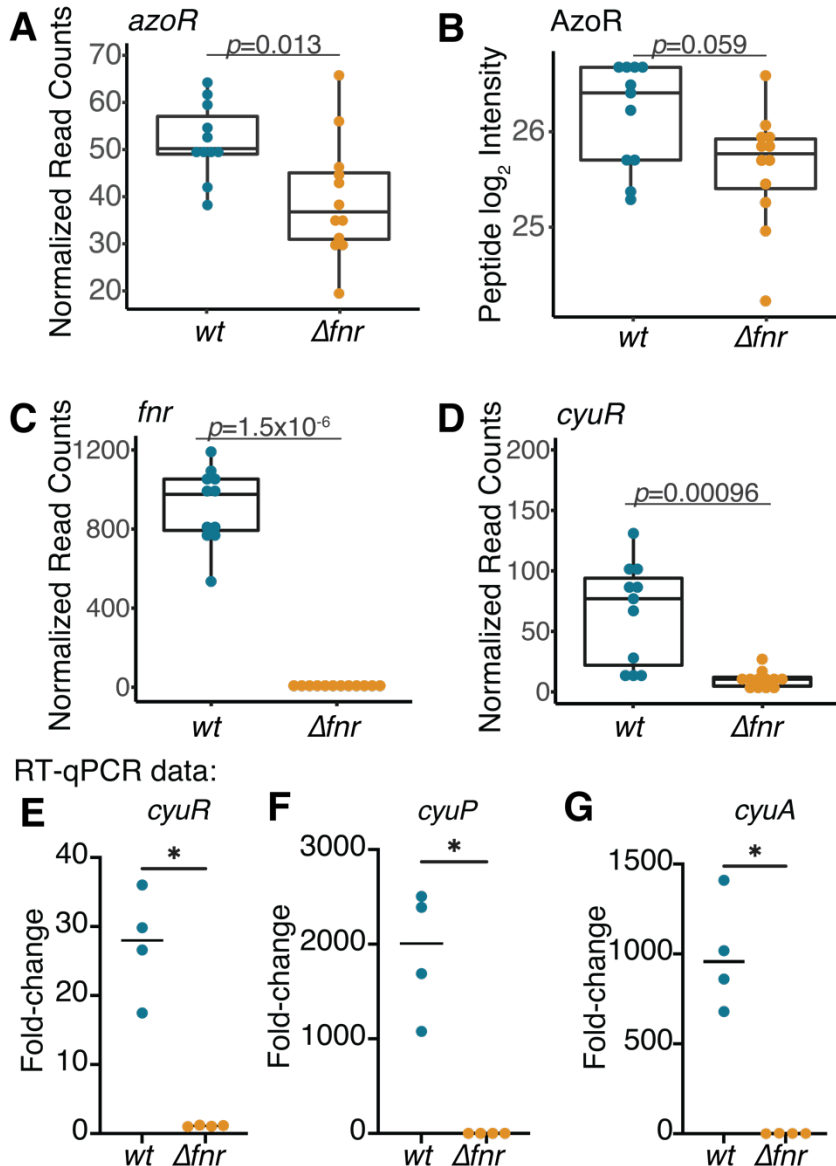


Figure S3.6 Key transcripts and proteins in Δfnr *E. coli* relative to wild-type controls. (A) *azoR* transcript levels are significantly depleted in the Δfnr strain. (B) AzoR peptide levels trend towards lower levels in the Δfnr strain. (C,D) *fnr* and *cyuR* transcript levels are significantly lower in the Δfnr strain. FNR and *cyuR* peptides were below the limit of detection. All timepoints and conditions are included within each genotype. p -value, Mann-Whitney test. (A,C,D) Normalized read counts are normalized by size factor and have an added pseudo count of 0.5 for \log_2 transformation plotting. (E-G) Fold-change of sample ΔC_t minus the maximum ΔC_t for each gene across all samples in *wt* or Δfnr *E. coli*. Measured in the presence of 4.2 mM L-Cysteine added at mid-exponential growth (Table S8). Anaerobic growth in LB with 4.2 mM L-Cysteine, * $p < 0.05$, Mann-Whitney test (n=4 biological replicates averaged from 3 technical replicates/strain).

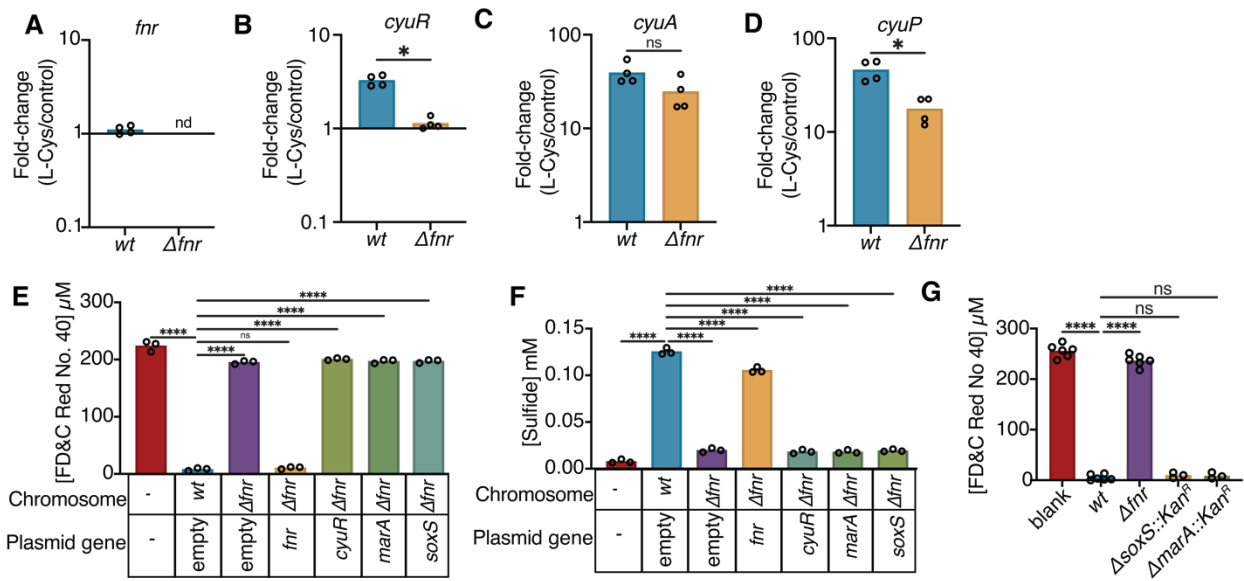


Figure S3.7 Regulation of *cyuR* and its operon. (A-D) Fold-change of transcripts in the presence of L-Cysteine compared to vehicle at mid-exponential growth measured by RT-qPCR (see **Table S8** for primer sequences). Paired FD&C Red No. 40 (E) and sulfide (F) were measured from *wt* or Δ *fnr* strains complemented either with an empty pCA24N vector or pCA24N carrying *fnr*, *cyuR*, *marA*, or *soxS*. (G) FD&C Red No. 40 was measured from knockout strains of *soxS* and *marA* (Keio collection). (A-D) Values are mean. * $p < 0.05$, Mann-Whitney test ($n = 4$ biological replicates averaged from 3 technical replicates/strain). (E-G) Bars are mean.. **** $p < 0.0001$, one-way ANOVA. (E-F) $n = 3$, (G) $n = 3-6$ replicates/strain.

3.5 Supplementary Tables

Supplemental Table S3.1. Keio collection strains screened for decreased azoreductase activity*.

*NA=not applicable; ND=not detected. **Per EcoCyc, ***(*adjusted p-value*<0.1)

Gene KO	Reason for screening	Proportion Dye Remaining (avg)	Genes above the 95% CI?	Gene annotation	FNR regulon?	Detected by RNA-seq?	log2 fold-change Δ fnr vs. wt	adjusted p-value	Altered by FD&C Red No. 40***
fnr	anaerobic/reductase	1.069	Yes	DNA-binding transcriptional dual regulator	-	Yes	NA	NA	No
yqiJ	anaerobic/reductase	0.789	Yes	DUF1449 domain-containing inner membrane protein YqiJ	No	No	ND	ND	No
cydD	anaerobic/reductase	0.755	Yes	glutathione/L-cysteine ABC exporter subunit CydD	Yes	Yes	0.69	1.1E-05	No
hcr	anaerobic/reductase	0.708	Yes	NADH oxidoreductase	Yes	Yes	-1.85	5.1E-08	No
fur	Brige	0.670	Yes	DNA-binding transcriptional dual regulator Fur	No	Yes	-0.69	4.5E-04	No
frdC	Zhang	0.611	Yes	fumarate reductase membrane protein FrdC	Yes	Yes	-3.65	3.0E-14	No
nuoC	anaerobic/reductase	0.552	Yes	NADH:quinone oxidoreductase subunit CD	Yes	Yes	1.23	1.0E-09	No
ydhH	anaerobic/reductase	0.550	Yes	(anmK) anhydro--N-acetylmuramic acid kinase	No	Yes	-0.03	8.5E-01	No
hybO	anaerobic/reductase	0.541	Yes	hydrogenase 2 small subunit	No	Yes	0.84	2.1E-07	No
narJ	Zhang	0.515	Yes	nitrate reductase 1 molybdenum cofactor assembly chaperone	Yes	Yes	-6.43	2.2E-21	No
frdD	Zhang	0.486	Yes	fumarate reductase	Yes	Yes	-3.20	1.3E-13	No

Gene KO	Reason for screen-ing	Proportion Dye Remaining (avg)	Genes above the 95% CI?	Gene annotation	FNR regulon?	Detected by RNA-seq?	log2 fold-change Δ fnr vs. wt	adjusted p-value	Altered by FD&C Red No. 40***
				membrane protein FrdD					
yeiQ	anaerobic/reductase	0.479	Yes	putative oxidoreductase	No	Yes	1.34	8.4E-05	No
hybC	anaerobic/reductase	0.462	Yes	hydrogenase 2 large subunit	No	Yes	0.69	8.5E-05	No
nuoB	anaerobic/reductase	0.431	Yes	NADH:quinone oxidoreductase subunit B	Yes	Yes	0.93	4.4E-08	No
nirB	Zhang	0.428	Yes	nitrite reductase (NADH) large subunit	Yes	Yes	-1.03	7.2E-05	No
ynfH	anaerobic/reductase	0.412	Yes	putative menaquinol dehydrogenase	Yes	Yes	-2.35	5.0E-13	No
hyaC	anaerobic/reductase	0.411	Yes	hydrogenase 1 cytochrome b subunit	No	Yes	-0.36	8.7E-02	No
nuoH	anaerobic/reductase	0.404	Yes	NADH:quinone oxidoreductase subunit H	Yes	Yes	1.86	1.2E-11	No
glpA	anaerobic/reductase	0.399	Yes	anaerobic glycerol-3-phosphate dehydrogenase subunit A	Yes	No	ND	ND	No
frdA	anaerobic/reductase	0.395	Yes	fumarate reductase flavoprotein subunit	Yes	Yes	-3.79	3.1E-16	No
soxS	Zhang	0.382	Yes	DNA-binding transcriptional dual regulator	Yes	Yes	0.32	3.9E-01	No
glpC	anaerobic/reductase	0.372	Yes	anaerobic glycerol-3-phosphate dehydrogenase subunit C	Yes	Yes	-0.64	2.5E-03	No
hybA	anaerobic/reductase	0.357	No	hydrogenase 2 iron-sulfur protein	No	Yes	0.64	8.1E-05	No

Gene KO	Reason for screen-ing	Proportion Dye Remaining (avg)	Genes above the 95% CI?	Gene annotation	FNR regulon?	Detected by RNA-seq?	log2 fold-change Δ fnr vs. wt	adjusted p-value	Altered by FD&C Red No. 40***
cydB	anaerobic/reductase	0.349	No	cytochrome bd-I subunit 2	Yes	Yes	2.33	2.5E-14	No
nuoE	anaerobic/reductase	0.345	No	NADH:quinone oxidoreductase subunit E	Yes	Yes	0.89	7.8E-06	Yes
narI	Zhang	0.342	No	nitrate reductase A subunit gamma	Yes	Yes	-5.83	4.6E-18	No
fadH	anaerobic/reductase	0.340	No	2,4-dienoyl-CoA reductase	No	No	ND	ND	No
ucpA	anaerobic/reductase	0.334	No	oxidoreductase UcpA	No	Yes	1.02	6.2E-07	No
guaC	anaerobic/reductase	0.329	No	GMP reductase	No	Yes	-0.52	2.0E-02	No
azoR	control, Zhang	0.324	No	FMN dependent NADH:quinone oxidoreductase	No	Yes	-0.91	5.3E-04	No
nirD	Zhang	0.323	No	nitrite reductase (NADH) small subunit	Yes	No	ND	ND	No
grxA	Zhang	0.322	No	reduced glutaredoxin 1	No	Yes	-5.31	1.5E-15	No
hybB	anaerobic/reductase	0.321	No	hydrogenase 2 membrane subunit	No	Yes	0.85	6.6E-05	No
nuoA	anaerobic/reductase	0.317	No	NADH:quinone oxidoreductase subunit A	Yes	Yes	0.45	2.8E-03	No
nuoL	anaerobic/reductase	0.312	No	NADH:quinone oxidoreductase subunit L	Yes	Yes	2.07	5.1E-11	No
nuoG	anaerobic/reductase	0.310	No	NADH:quinone oxidoreductase subunit G	Yes	Yes	1.58	2.1E-08	No
nuoF	anaerobic/reductase	0.309	No	NADH:quinone oxidoreductase subunit F	Yes	Yes	1.41	2.6E-08	No
nuoN	anaerobic/reductase	0.304	No	NADH:quinone oxidoreductase subunit N	Yes	Yes	1.86	7.6E-12	No

Gene KO	Reason for screening	Proportion Dye Remaining (avg)	Genes above the 95% CI?	Gene annotation	FNR regulon?	Detected by RNA-seq?	log2 fold-change Δ fnr vs. wt	adjusted p-value	Altered by FD&C Red No. 40***
napA	anaerobic/reductase	0.293	No	periplasmic nitrate reductase subunit NapA	Yes	Yes	-0.29	3.3E-01	No
nuoK	anaerobic/reductase	0.293	No	NADH:quinone oxidoreductase subunit K	Yes	Yes	1.44	1.8E-07	No
nrfD	Zhang	0.293	No	putative menaquinol-cytochrome c reductase subunit NrfD	Yes	No	ND	ND	No
pyrD	anaerobic/reductase	0.291	No	dihydroorotate dehydrogenase, type 2	No	Yes	-0.79	1.2E-03	No
glpD	anaerobic/reductase	0.289	No	aerobic glycerol 3-phosphate dehydrogenase	No	Yes	2.49	3.3E-14	No
hcaD	anaerobic/reductase	0.286	No	putative 3-phenylpropionate/cinnamate dioxygenase ferredoxin reductase subunit	via <i>fnrS</i>	Yes	-0.03	9.2E-01	No
yfeH	anaerobic/reductase	0.284	No	putative solute:Na ⁺ symporter	No	Yes	-0.77	2.7E-04	No
kefF	anaerobic/reductase	0.279	No	regulator of KefC-mediated potassium transport and quinone oxidoreductase	No	Yes	0.64	1.3E-02	No
nrfC	Zhang	0.278	No	putative menaquinol-cytochrome c reductase 4Fe-4S subunit	Yes	No	ND	ND	No
uxuB	anaerobic/reductase	0.277	No	D-mannonate oxidoreductase	No	Yes	1.44	3.4E-08	No
mltD	anaerobic/reductase	0.277	No	membrane-bound lytic murein transglycosylase D	No	Yes	0.05	8.0E-01	No

Gene KO	Reason for screen-ing	Proportion Dye Remaining (avg)	Genes above the 95% CI?	Gene annotation	FNR regulon?	Detected by RNA-seq?	log2 fold-change Δ fnr vs. wt	adjusted p-value	Altered by FD&C Red No. 40***
frdB	Zhang	0.274	No	fumarate reductase iron-sulfur protein	Yes	Yes	-3.81	1.0E-15	No
proC	anaerobic/reductase	0.271	No	pyrroline-5-carboxylate reductase	No	Yes	0.20	1.9E-01	No
hokD	anaerobic/reductase	0.271	No	qin prophage; toxic protein	No	Yes	-1.23	4.6E-06	No
fxsA	anaerobic/reductase	0.269	No	F exclusion of bacteriophage T7	No	Yes	-0.58	1.3E-02	No
nfsA	anaerobic/reductase	0.267	No	NADPH-dependent nitroreductase / NADPH-dependent quinone reductase	No	Yes	-1.09	1.5E-08	No
frmA	anaerobic/reductase	0.266	No	S-(hydroxymethyl)glutathione dehydrogenase	No	Yes	0.78	2.9E-06	No
nuoM	anaerobic/reductase	0.264	No	NADH:quinone oxidoreductase subunit M	Yes	Yes	1.98	7.9E-12	No
narH	Zhang	0.264	No	nitrate reductase A subunit beta	Yes	Yes	-6.02	6.9E-24	No
glpB	anaerobic/reductase	0.263	No	anaerobic glycerol-3-phosphate dehydrogenase subunit B	Yes	No	ND	ND	No
hyaB	anaerobic/reductase	0.262	No	hydrogenase 1 large subunit	No	Yes	-0.50	2.6E-02	No
yciK	anaerobic/reductase	0.261	No	putative oxidoreductase	No	Yes	-0.72	7.1E-05	No
ybdK	anaerobic/reductase	0.260	No	putative glutamate -- cysteine ligase 2	No	Yes	0.32	5.5E-02	Yes
yhaH	anaerobic/reductase	0.258	No	putative inner membrane protein	No	Yes	-0.24	1.9E-01	Yes

Gene KO	Reason for screen-ing	Proportion Dye Remaining (avg)	Genes above the 95% CI?	Gene annotation	FNR regulon?	Detected by RNA-seq?	log2 fold-change Δ fnr vs. wt	adjusted p-value	Altered by FD&C Red No. 40***
ndh	anaerobic/reductase	0.258	No	NADH:quinone oxidoreductase II	Yes	Yes	2.21	6.6E-13	No
narP	Zhang	0.258	No	DNA-binding transcriptional dual regulator// nitrate/nitrite response regulator	No	Yes	-0.21	8.9E-02	No
napF	anaerobic/reductase	0.258	No	ferredoxin-type protein	Yes	No	ND	ND	No
cybC	anaerobic/reductase	0.256	No	pseudogene	No	Yes	-0.25	7.9E-02	No
guaC	anaerobic/reductase	0.253	No	GMP reductase	No	Yes	-0.52	2.0E-02	No
fdnI	anaerobic/reductase	0.252	No	formate dehydrogenase N subunit gamma	Yes	Yes	-1.51	1.3E-04	No
fixC	anaerobic/reductase	0.252	No	putative oxidoreductase FixC	Yes	No	ND	ND	No
ytfG	anaerobic/reductase	0.251	No	NAD(P)H:quinone oxidoreductase	No	Yes	0.56	7.3E-02	No
yhdH	anaerobic/reductase	0.251	No	acrylyl-CoA reductase	No	Yes	-0.25	1.5E-01	No
narV	anaerobic/reductase	0.250	No	nitrate reductase Z subunit gamma	No	No	ND	ND	No
aldA	anaerobic/reductase	0.250	No	aldehyde dehydrogenase A	Yes	Yes	2.76	1.6E-13	Yes
torY	anaerobic/reductase	0.249	No	cytochrome c quinol dehydrogenase TorY	No	No	ND	ND	No
napC	anaerobic/reductase	0.247	No	periplasmic nitrate reductase cytochrome c protein	Yes	Yes	-0.77	5.5E-03	No
torC	anaerobic/reductase	0.247	No	cytochrome c menaquinol dehydrogenase TorC	No	Yes	-0.52	1.9E-03	No

Gene KO	Reason for screen-ing	Proportion Dye Remaining (avg)	Genes above the 95% CI?	Gene annotation	FNR regulon?	Detected by RNA-seq?	log2 fold-change Δ fnr vs. wt	adjusted p-value	Altered by FD&C Red No. 40***
caiA	anaerobic/reductase	0.244	No	crotonobetainyl-CoA reductase	Yes	Yes	-0.15	4.4E-01	No
hyaA	anaerobic/reductase	0.243	No	hydrogenase 1 small subunit	No	Yes	-0.51	1.6E-02	No
hmp	anaerobic/reductase	0.240	No	nitric oxide dioxygenase	Yes	Yes	6.79	7.4E-21	No
ccmH	anaerobic/reductase	0.239	No	holocytochrome c synthase CcmH component	Yes	Yes	0.11	6.4E-01	No
torZ	anaerobic/reductase	0.237	No	trimethylamine N-oxide reductase 2	No	Yes	-0.29	2.4E-01	No
nrfG	anaerobic/reductase	0.236	No	putative formate-dependent nitrite reductase complex subunit NrfG	Yes	No	ND	ND	No
ydfG	anaerobic/reductase	0.236	No	3-hydroxy acid dehydrogenase	No	Yes	-0.52	2.9E-03	No
narW	anaerobic/reductase	0.234	No	NarW, putative private chaperone for NarZ nitrate reductase subunit	No	No	ND	ND	No
mhpF	anaerobic/reductase	0.232	No	acetaldehyde dehydrogenase (acetylating) MhpF	No	Yes	-0.12	4.5E-01	No
dmsC	Brige & Zhang	0.231	No	dimethyl sulfoxide reductase subunit C	Yes	Yes	-1.54	1.5E-09	No
trxB	Zhang	0.231	No	thioredoxin reductase	No	Yes	-0.18	1.7E-01	No
ynfE	anaerobic/reductase	0.231	No	putative selenate reductase YnfE	Yes	Yes	-3.82	6.2E-18	No
xdhD	anaerobic/reductase	0.230	No	fused putative xanthine/hypoxanthine oxidase: molybdopterin-binding subunit and Fe-S binding subunit	No	Yes	5.45	7.9E-18	No

Gene KO	Reason for screen-ing	Proportion Dye Remaining (avg)	Genes above the 95% CI?	Gene annotation	FNR regulon?	Detected by RNA-seq?	log2 fold-change Δ fnr vs. wt	adjusted p-value	Altered by FD&C Red No. 40***
ccp	anaerobic/reductase	0.230	No	cytochrome c peroxidase	Yes	Yes	-3.85	3.3E-17	No
yqhD	Zhang	0.229	No	NADPH-dependent aldehyde reductase YqhD	No	Yes	-2.40	2.5E-11	No
fdnG	anaerobic/reductase	0.228	No	formate dehydrogenase N subunit alpha	Yes	Yes	-1.89	5.0E-08	No
wrbA	Zhang	0.226	No	NAD(P)H:quinone oxidoreductase	No	Yes	-1.61	2.2E-10	Yes
nrfB	Zhang	0.226	No	periplasmic nitrite reductase penta-heme c-type cytochrome	Yes	No	ND	ND	No
dmsA	Brige & Zhang	0.226	No	dimethyl sulfoxide reductase subunit A	Yes	Yes	-3.73	1.3E-17	No
yccW	anaerobic/reductase	0.225	No	(rImI) 23S rRNA m5C1962 methyltransferase	No	Yes	-0.33	4.9E-02	No
narG	Zhang	0.224	No	nitrate reductase A subunit α	Yes	Yes	-5.97	2.4E-24	No
nrfA	Zhang	0.224	No	cytochrome c 552 nitrite reductase	Yes	No	ND	ND	No
torA	anaerobic/reductase	0.220	No	trimethylamine N-oxide reductase 1	No	Yes	-0.53	3.1E-02	No
ftnA	anaerobic/reductase	0.217	No	ferritin iron-storage complex	No	Yes	2.32	2.5E-10	No
ybdH	anaerobic/reductase	0.213	No	hydroxycarboxylate dehydrogenase A	No	Yes	-0.44	2.4E-02	No
dmsB	Brige & Zhang	0.212	No	dimethyl sulfoxide reductase subunit B	Yes	No	ND	ND	No
nadB	anaerobic/reductase	0.212	No	L-aspartate oxidase	No	Yes	0.16	4.6E-01	No
ynfE	anaerobic/reductase	0.209	No	putative selenate reductase YnfE	Yes	Yes	-3.82	6.2E-18	No

Gene KO	Reason for screen-ing	Proportion Dye Remaining (avg)	Genes above the 95% CI?	Gene annotation	FNR regulon?	Detected by RNA-seq?	log2 fold-change Δ fnr vs. wt	adjusted p-value	Altered by FD&C Red No. 40***
narZ	anaerobic/reductase	0.207	No	nitrate reductase Z subunit alpha	No	Yes	2.21	9.6E-14	No
nfsB	anaerobic/reductase	0.206	No	NAD(P)H-dependent nitroreductase	No	Yes	-1.55	6.5E-12	No
wild-type	control	0.202	No	-	-	-	-	-	-
narY	anaerobic/reductase	0.202	No	nitrate reductase Z subunit beta	No	Yes	2.53	9.7E-13	No
sdhA	anaerobic/reductase	0.201	No	succinate:quinone oxidoreductase, FAD binding protein	Yes	Yes	4.68	2.9E-19	No
fdnH	anaerobic/reductase	0.201	No	formate dehydrogenase N subunit beta	Yes	No	ND	ND	No
yceM	anaerobic/reductase	0.198	No	putative oxidoreductase YceM	No	Yes	0.21	2.1E-01	No
nuoJ	anaerobic/reductase	0.195	No	NADH:quinone oxidoreductase subunit J	Yes	Yes	1.54	9.1E-10	No
panE	anaerobic/reductase	0.174	No	2-dehydropantoate 2-reductase	No	Yes	0.12	4.2E-01	No
srID	anaerobic/reductase	0.173	No	sorbitol-6-phosphate 2-dehydrogenase	No	Yes	0.33	5.9E-02	No

Supplemental Table S3.2. RNA sequencing metadata and sequencing statistics.

Sample identifier	<i>E. coli</i> strain	Growth condition	Timepoint	High-quality sequencing reads
S1	WT	R40	mid-exponential	9.02E+06
S2	WT	R40	mid-exponential	1.13E+07
S3	WT	R40	mid-exponential	1.14E+07
S4	WT	Vehicle	mid-exponential	1.16E+07
S5	WT	Vehicle	mid-exponential	8.21E+06
S6	WT	Vehicle	mid-exponential	1.16E+07
S7	Δ fnr	R40	mid-exponential	7.67E+06
S8	Δ fnr	R40	mid-exponential	1.03E+07
S9	Δ fnr	R40	mid-exponential	3.89E+07
S10	Δ fnr	Vehicle	mid-exponential	1.18E+07
S11	Δ fnr	Vehicle	mid-exponential	1.28E+07
S12	Δ fnr	Vehicle	mid-exponential	1.28E+07
S13	WT	R40	stationary phase	1.26E+07
S14	WT	R40	stationary phase	1.49E+07
S15	WT	Vehicle	stationary phase	1.11E+07
S16	WT	Vehicle	stationary phase	1.05E+07
S17	WT	Vehicle	stationary phase	1.27E+07
S18	Δ fnr	R40	stationary phase	1.73E+07
S19	Δ fnr	R40	stationary phase	8.21E+06
S20	Δ fnr	R40	stationary phase	1.02E+07
S21	Δ fnr	Vehicle	stationary phase	9.88E+06
S22	Δ fnr	Vehicle	stationary phase	1.07E+07
S23	Δ fnr	Vehicle	stationary phase	1.01E+07

Supplemental Table S3.2. Peptide metadata and peptide counts.

Sample identifier	<i>E. coli</i> strain	Growth condition	Timepoint	Unique Peptides Count
S1	WT	R40	mid-exponential	909
S2	WT	R40	mid-exponential	905
S3	WT	R40	mid-exponential	921
S4	WT	Vehicle	mid-exponential	913
S5	WT	Vehicle	mid-exponential	902
S6	WT	Vehicle	mid-exponential	904
S7	fnr	R40	mid-exponential	870
S8	fnr	R40	mid-exponential	873
S9	fnr	R40	mid-exponential	882
S10	fnr	Vehicle	mid-exponential	856
S11	fnr	Vehicle	mid-exponential	893
S12	fnr	Vehicle	mid-exponential	904
S13	WT	R40	stationary phase	882
S14	WT	R40	stationary phase	880
S15	WT	Vehicle	stationary phase	880
S16	WT	Vehicle	stationary phase	866
S17	WT	Vehicle	stationary phase	857
S18	fnr	R40	stationary phase	879
S19	fnr	R40	stationary phase	852
S20	fnr	R40	stationary phase	868
S21	fnr	Vehicle	stationary phase	895
S22	fnr	Vehicle	stationary phase	834
S23	fnr	Vehicle	stationary phase	843

Table S3.4. Multivariate analysis of differentially abundant E. coli proteins and transcripts.

+Number of features significantly enriched or depleted by each component of the interaction model ~Strain*Growth Phase*Condition for Proteomics. Significance is based on log2 fold-change $\geq |1|$ and padj < 0.1 .

*Number of features significantly enriched or depleted by each component of the interaction model ~Strain*Growth Phase*Condition for Transcriptomics. Significance is based on log2 fold-change $\geq |1|$ and padj < 0.1 .

		Single Component Interaction Terms			Multiple Component Interaction Terms			
	Fold-change	Δfnr vs. wt	Mid Exponential: vs. Stat	Red 40 vs. Vehicle	Δfnr : Mid Exponential vs. wt: Stationary	Δfnr : Vehicle vs. wt: Red 40	Mid Exponential: Vehicle vs. Stat: Red 40	Δfnr : Mid Exponential: Vehicle vs. wt: Stat: Red40
Proteins+	Enriched	92	65	0	33	0	1	1
	Depleted	124	44	3	4	0	32	1
	No change	801	908	1014	980	1017	984	1015
Transcripts*	Enriched	315	221	1	523	73	0	0
	Depleted	273	273	35	584	88	0	0
	No change	2552	2646	3104	2033	2979	3140	3140

Table S3.5. Genes with consistent changes in both RNA-seq and proteomics data.

Gene	log2 fold- Δ protein	Adj.p-value protein	log2 fold- Δ RNA	Adj. p-value RNA	Gene identifier	Functional category	Gene annotation
pta	1.91	2.19E-06	1.61	2.67E-10	BW25113_2297	acetyltransferase	phosphate acetyltransferase
gadc	2.31	1.10E-03	1.48	1.79E-08	BW25113_1492	antiporter	L-glutamate:4-aminobutyrate antiporter
aspg2	10.92	3.10E-08	3.76	7.27E-18	BW25113_2957	asparaginase	L-asparaginase 2
ag43	-3.25	8.40E-03	-3.54	1.85E-11	BW25113_2000	autotransporter	CP4-44 prophage; self recognizing antigen 43 (Ag43) autotransporter
taba	1.72	3.21E-03	1.10	6.78E-09	BW25113_4252	biofilm protein	DUF386 domain-containing toxin-antitoxin biofilm protein TabA
osmy	-3.68	6.60E-05	-2.80	2.80E-15	BW25113_4376	chaperone	periplasmic chaperone OsmY
ivy	1.16	1.69E-02	1.51	2.53E-09	BW25113_0220	chaperone	periplasmic chaperone, inhibitor of vertebrate C-type lysozyme
dlst	7.40	2.03E-05	4.36	3.90E-20	BW25113_3110	cysteine detoxification	cysteine detoxification protein CyuP
naga	1.20	7.23E-04	1.63	4.59E-11	BW25113_0677	deacetylase	N-acetylglucosamine-6-phosphate deacetylase

Gene	log2 fold- Δ protein	Adj. p-value protein	log2 fold- Δ RNA	Adj. p-value RNA	Gene identifier	Functional category	Gene annotation
nagb	1.32	4.54E-05	2.12	5.73E-15	BW25113_0678	deaminase	glucosamine-6-phosphate deaminase
adia	3.40	2.62E-08	4.93	8.45E-21	BW25113_4117	decarboxylase	arginine decarboxylase, degradative
rne	1.10	9.09E-05	-6.79	7.37E-21	BW25113_2552	degradosome	degradosome
tdcb	5.21	3.62E-03	-1.99	3.37E-05	BW25113_3117	dehydratase	catabolic threonine dehydratase
acnb	-1.76	1.18E-05	-2.28	2.13E-14	BW25113_0118	dehydratase	bifunctional aconitate hydratase B and 2-methylisocitrate dehydratase
fdhe	-6.78	1.01E-02	-1.91	5.44E-11	BW25113_3891	dehydrogenase	formate dehydrogenase formation protein
ykgf	-5.64	3.42E-06	-3.98	2.01E-16	BW25113_0307	dehydrogenase	putative amino acid dehydrogenase with NAD(P)-binding domain and ferrodoxin-like domain
xdhd	-4.94	7.87E-05	-5.45	7.85E-18	BW25113_2881	dehydrogenase	predicted xanthine dehydrogenase
fdog	-4.77	7.47E-02	-2.83	6.56E-14	BW25113_3894	dehydrogenase	formate dehydrogenase O
alda	-4.08	1.29E-03	-2.76	1.55E-13	BW25113_1415	dehydrogenase	aldehyde dehydrogenase A, NAD-linked

Gene	log2 fold- Δ protein	Adj. p-value protein	log2 fold- Δ RNA	Adj. p-value RNA	Gene identifier	Functional category	Gene annotation
fdoh	-3.58	2.12E-02	-3.37	3.19E-16	BW25113_3893	dehydrogenase	formate dehydrogenase O
glda	2.91	5.63E-05	1.69	1.34E-11	BW25113_3945	dehydrogenase	L-1,2-propanediol dehydrogenase / glycerol dehydrogenase
dada	-2.77	8.78E-02	-3.17	2.65E-15	BW25113_1189	dehydrogenase	D-amino acid dehydrogenase
tdh	-2.41	2.19E-06	-1.32	3.86E-09	BW25113_3616	dehydrogenase	threonine dehydrogenase
mdh	-2.05	4.54E-09	-2.77	1.02E-16	BW25113_3236	dehydrogenase	malate dehydrogenase
glpd	-1.69	1.58E-02	-2.49	3.28E-14	BW25113_3426	dehydrogenase	aerobic glycerol 3-phosphate dehydrogenase
yham	6.41	1.43E-06	3.80	3.79E-20	BW25113_4470	desulfidase	putative L-cysteine desulfidase CyuA
iscs	1.10	1.04E-05	3.42	1.11E-19	BW25113_2530	desulfidase	cysteine desulfurase IscS
hmp	-11.24	3.10E-08	-6.79	7.37E-21	BW25113_2552	dioxygenase	nitric oxide dioxygenase
flic	-10.07	4.75E-08	-5.93	2.44E-24	BW25113_1923	flagellum	flagellum
fumb	8.88	5.44E-04	2.27	6.50E-10	BW25113_4122	fumarase	fumarase B
fuma	1.51	2.30E-03	-2.83	7.05E-15	BW25113_1612	fumarase	fumarase A
ftna	-2.39	1.19E-05	-2.32	2.47E-10	BW25113_1905	iron storage	ferritin iron-storage complex
uxac	-1.07	6.60E-05	-1.24	3.22E-08	BW25113_3092	isomerase	D-glucuronate/D-

Gene	log2 fold-Δ protein	Adj. p-value protein	log2 fold-Δ RNA	Adj. p-value RNA	Gene identifier	Functional category	Gene annotation
							galacturonate isomerase
glpk	-1.76	2.98E-07	-3.41	1.17E-09	BW25113_3926	kinase	glycerol kinase
acka	1.48	2.27E-06	1.83	2.43E-12	BW25113_2296	kinase	acetate kinase
aspa	2.83	2.82E-10	2.16	1.20E-11	BW25113_4139	lyase	aspartate ammonia-lyase
grca	2.15	1.67E-08	3.73	1.20E-15	BW25113_2579	lyase	stress-induced alternate pyruvate formate-lyase subunit
pflb	2.15	2.31E-09	2.13	2.64E-11	BW25113_0903	lyase	pyruvate formate-lyase (inactive)
ompx	2.92	4.79E-02	1.16	4.90E-10	BW25113_0814	membrane protein	outer membrane protein X
loip	-3.46	4.13E-02	-1.27	3.59E-09	BW25113_2936	metalloprotease	metalloprotease LoIP
ygfk	-10.82	1.90E-02	-5.54	2.83E-18	BW25113_2878	oxidoreductase	putative oxidoreductase, Fe-S subunit
sdha	-7.13	1.19E-04	-4.68	2.95E-19	BW25113_0723	oxidoreductase	succinate:quinone oxidoreductase subcomplex
ykge	-4.73	9.71E-02	-5.17	3.25E-16	BW25113_0306	oxidoreductase	putative lactate utilization oxidoreductase YkgE
sdhb	-4.46	2.09E-02	-4.71	8.26E-19	BW25113_0724	oxidoreductase	succinate:quinone oxidoreductase subcomplex

Gene	log2 fold- Δ protein	Adj. p-value protein	log2 fold- Δ RNA	Adj. p-value RNA	Gene identifier	Functional category	Gene annotation
ydja	-1.35	2.91E-03	-1.05	2.92E-07	BW25113_1765	oxidoreductase	putative oxidoreductase YdjA
nuoc	-4.18	4.40E-02	-1.23	1.00E-09	BW25113_2286	oxidoreductase	NADH:quinone oxidoreductase I
ygey	-5.97	3.72E-04	-5.53	5.92E-18	BW25113_2872	peptidase	putative peptidase YgeY
yfex	1.86	8.42E-06	1.30	3.33E-11	BW25113_2431	peroxidase	porphyrinogen peroxidase
tpx	-2.30	7.07E-06	2.70	5.08E-16	BW25113_0605	peroxidase	lipid hydroperoxide peroxidase
glpq	1.07	1.37E-03	-1.05	4.69E-07	BW25113_2239	phosphodiesterase	glycerophosphoryl diester phosphodiesterase, periplasmic
yqjh	10.60	2.07E-06	7.80	8.10E-18	BW25113_3070	reductase	NADPH-dependent ferric-chelate reductase
frda	7.77	2.65E-03	3.79	3.10E-16	BW25113_4154	reductase	fumarate reductase flavoprotein subunit // fumarate reductase iron-sulfur protein // fumarate reductase membrane protein FrdC // fumarate reductase membrane protein FrdD
narg	7.14	6.25E-03	5.97	2.44E-24	BW25113_1224	reductase	nitrate reductase A
narh	5.81	6.78E-04	6.02	6.90E-24	BW25113_1225	reductase	nitrate reductase A

Gene	log2 fold-Δ protein	Adj.p-value protein	log2 fold-Δ RNA	Adj. p-value RNA	Gene identifier	Functional category	Gene annotation
ynff	5.70	5.13E-02	4.38	1.36E-15	BW25113_1588	reductase	selenate reductase
dmsa	5.14	3.97E-03	3.73	1.27E-17	BW25113_0894	reductase	dimethyl sulfoxide reductase subunit A // dimethyl sulfoxide reductase subunit B // dimethyl sulfoxide reductase subunit C
frdb	5.02	2.03E-05	3.81	1.00E-15	BW25113_4153	reductase	fumarate reductase flavoprotein subunit // fumarate reductase iron-sulfur protein // fumarate reductase membrane protein FrdC // fumarate reductase membrane protein FrdD
narj	4.68	5.53E-04	6.43	2.16E-21	BW25113_1226	reductase	nitrate reductase 1 molybdenum cofactor assembly chaperone
nirb	4.43	6.36E-02	1.03	7.21E-05	BW25113_3365	reductase	nitrite reductase (NADH) small subunit // nitrite reductase (NADH) large subunit

Gene	log2 fold- Δ protein	Adj. p-value protein	log2 fold- Δ RNA	Adj. p-value RNA	Gene identifier	Functional category	Gene annotation
nrdd	4.43	2.18E-02	3.22	8.88E-18	BW25113_4238	reductase	anaerobic nucleoside-triphosphate reductase activating system
dkgb	-3.54	4.70E-02	1.36	3.92E-08	BW25113_0207	reductase	methylglyoxal reductase DkgB
yqhd	1.14	2.12E-04	2.40	2.45E-11	BW25113_3011	reductase	NADPH-dependent aldehyde reductase YqhD
ilvc	-3.83	9.68E-03	-1.16	2.07E-08	BW25113_3774	reductoisomerase	ketol-acid reductoisomerase (NADP+)
ytfe	-9.67	1.29E-05	-5.28	6.55E-17	BW25113_4209	repair protein	iron-sulfur cluster repair protein YtfE
yqji	6.57	6.34E-06	8.56	4.25E-26	BW25113_3071	repressor	DNA-binding transcriptional repressor
uspe	2.41	1.82E-04	1.38	4.21E-10	BW25113_1333	stress protein	universal stress protein E
uspf	1.69	1.22E-04	2.05	1.30E-09	BW25113_1376	stress protein	universal stress protein F
cfa	8.13	4.88E-04	1.43	3.52E-10	BW25113_1661	synthase	cyclopropane fatty acyl phospholipid synthase
succ	-5.96	5.13E-02	-4.38	7.85E-18	BW25113_0728	synthetase	succinyl-CoA synthetase subunit alpha // succinyl-CoA synthetase subunit beta
sucd	-2.41	1.93E-02	-4.13	2.29E-17	BW25113_0729	synthetase	succinyl-CoA synthetase subunit alpha // succinyl-

Gene	log2 fold- Δ protein	Adj.p-value protein	log2 fold- Δ RNA	Adj. p-value RNA	Gene identifier	Functional category	Gene annotation
							CoA synthetase subunit beta
moeb	-8.90	3.31E-02	-1.45	4.07E-10	BW25113_0826	transferase	molybdopterin-synthase adenylyltransferase
ynje	7.90	3.35E-04	3.87	1.81E-17	BW25113_1757	transferase	molybdopterin synthase sulfurtransferase
gabt	-5.23	4.74E-03	-1.14	4.21E-05	BW25113_2662	transferase	4-aminobutyrate aminotransferase GabT
mete	-7.65	2.01E-04	-1.67	1.12E-11	BW25113_3829	transmethylase	cobalamin-independent homocysteine transmethylase
ynjb	5.09	7.59E-03	1.91	1.63E-09	BW25113_1754	transport complex	putative transport complex, ABC superfamily
glti	-6.43	9.09E-05	-1.93	1.77E-13	BW25113_0655	transporter	glutamate / aspartate ABC transporter
hisj	-5.97	1.31E-03	-1.54	6.88E-09	BW25113_2309	transporter	histidine ABC transporter
dppa	-4.10	1.54E-03	-4.74	1.02E-20	BW25113_3544	transporter	dipeptide ABC transporter
yjia	4.33	5.59E-03	2.57	6.67E-16	BW25113_4352	triphosphatase	P-loop guanosine triphosphatase YjiA
tnaa	-1.03	5.80E-06	-4.78	2.98E-16	BW25113_3708	tryptophanase	tryptophanase
yjii	5.37	3.28E-05	1.20	1.24E-07	BW25113_4380	unknown	DUF3029 domain-

Gene	log2 fold- Δ protein	Adj. p-value protein	log2 fold- Δ RNA	Adj. p-value RNA	Gene identifier	Functional category	Gene annotation
							containing protein Yjjl
yccj	1.30	2.37E-02	1.47	1.95E-10	BW25113_1003	unknown	PF13993 family protein YccJ
yfcz	1.15	2.09E-02	1.40	3.92E-07	BW25113_2343	unknown	DUF406 domain-containing protein YfcZ

Table S3.6. Genes that were consistently affected at the transcript and protein level and significant in our knockout screen.

Gene	<i>frdA</i>	<i>narJ</i>	<i>nirB</i>	<i>nuoC</i>
Mean proportion dye remaining in Keio screen	0.39	0.52	0.43	0.55
log2 fold-change (Δfnr vs. <i>wt</i>) (RNA)	-3.79	-6.43	-1.03	1.23
adjusted <i>p</i> -value (RNA)	3.10E-16	2.16E-21	7.21E-05	1.00E-09
log2 fold-change (Δfnr vs. <i>wt</i>) (protein)	-7.77	-4.68	-4.43	4.18
adjusted <i>p</i> -value (protein)	2.65E-03	5.53E-04	6.36E-02	4.40E-02
Annotation	fumarate reductase flavoprotein subunit // fumarate reductase iron-sulfur protein // fumarate reductase membrane protein FrdC // fumarate reductase membrane protein FrdD	nitrate reductase 1 molybdenum cofactor assembly chaperone	nitrite reductase (NADH) small subunit // nitrite reductase (NADH) large subunit	NADH:quinone oxidoreductase I

Table S3.7. Strains used in Chapter 3.

Strain	Source	Notes
<i>E. coli</i> BW25113 wt + pCA24N (no gene)	ASKA Collection	
<i>E. coli</i> BW25113 wt	Keio Collection	
<i>E. coli</i> BW25113 Δ azoR::KanR	Keio Collection	
<i>E. coli</i> BW25113 Δ fnr::KanR	Keio Collection	
<i>E. coli</i> BW25113 Δ rstA::KanR	Keio Collection	
<i>E. coli</i> BW25113 Δ dcuR::KanR	Keio Collection	
<i>E. coli</i> BW25113 Δ arcA::KanR	Keio Collection	
<i>E. coli</i> BW25113 Δ nfeR::KanR	Keio Collection	
<i>E. coli</i> BW25113 Δ ihfA::KanR	Keio Collection	
<i>E. coli</i> BW25113 Δ narL::KanR	Keio Collection	
<i>E. coli</i> BW25113 Δ fisA::KanR	Keio Collection	
<i>E. coli</i> BW25113 Δ ihfB::KanR	Keio Collection	
<i>E. coli</i> BW25113 Δ cyuR::KanR	Keio Collection	in Keio collection as "ybaO"
<i>E. coli</i> BW25113 Δ crp::KanR	Keio Collection	
<i>E. coli</i> BW25113 Δ cyuA::KanR	Keio Collection	in Keio collection as "yhaM"
<i>E. coli</i> BW25113 Δ cyuP::KanR	Keio Collection	in Keio collection as "yhaO"
<i>E. coli</i> BW25113 Δ azoR	This paper	Created using pSIJ8 to remove KanR
<i>E. coli</i> BW25113 Δ fnr	This paper	Created using pSIJ8 to remove KanR
<i>E. coli</i> BW25113 Δ cyuR Δ azoR::KanR	This paper	Created using pSIJ8 to remove KanR from cyuR background, azoR::KanR introduced via P1 transduction
<i>E. coli</i> BW25113 Δ fnr + pCA24N-fnr	This paper	plasmid from ASKA collection
<i>E. coli</i> BW25113 Δ fnr + pCA24N-azoR	This paper	plasmid from ASKA collection
<i>E. coli</i> BW25113 Δ fnr + pCA24N (no gene)	This paper	plasmid from ASKA collection

Strain	Source	Notes
<i>E. coli</i> BW25113 $\Delta marA::kanR$	This paper	plasmid from ASKA collection
<i>E. coli</i> BW25113 $\Delta soxS::kanR$	This paper	plasmid from ASKA collection
<i>E. coli</i> BW25113 $\Delta fnr + pCA24N-marA$	This paper	plasmid from ASKA collection
<i>E. coli</i> BW25113 $\Delta fnr + pCA24N-soxS$	This paper	plasmid from ASKA collection
<i>E. coli</i> BW25113 $\Delta fnr + pCA24N-cyuR$	This paper	plasmid from ASKA collection
<i>E. coli</i> MG1655 wt		
<i>E. coli</i> MG1655 $\Delta fnr::kanR$	This paper	Via P1 transduction, using Keio $\Delta fnr::kan$ strain as template
<i>E. coli</i> MG1655 $\Delta fnrS::kanR$	Storz lab	Durand and Storz (2010)
<i>E. coli</i> MG1655 + <i>pBD-pLac</i>	Storz lab	
<i>E. coli</i> MG1655 + <i>pBD-pLac-fnrS</i>	Storz lab	
<i>E. coli</i> MG1655 $\Delta fnr::kanR + pBD-pLac$	This paper	
<i>E. coli</i> MG1655 $\Delta fnr::kanR + pBD-pLac-fnrS$	This paper	
<i>E. coli</i> MG1655 $\Delta fnrS::kanR + pBD-pLac$	This paper	
<i>E. coli</i> MG1655 $\Delta fnrS::kanR + pBD-pLac-fnrS$	This paper	

Table S3.8. Primers used in Chapter 3.

Primer name	Primer sequence (5' --> 3')	Notes
arcA_down_R	CCGGCATGATGTTTGTGACC	
arcA_up_F	GGCAGGTCAGGGACTTTTGT	
azoR_downstream_F	ATTGCGCTATCTCTGTGGCA	
azoR_internal_F_2203	AGGAAGCTCTGGCACTTTCC	
azoR_internal_R_2203	CTGTGCTTTCGCTGCCATTT	
azoR_upstream_R	GAGCACAAAGCGATGCTGAG	
crp_down_R	CATAGCACCAGCGTTTGTGCG	
crp_up_F	CAGACCCGACTCTCGAATGG	
cyuA_1_F_RTqPCR	AGGTTTGTAGCGGTAAGTGGG	
cyuA_1_R_RTqPCR	GATCTTTCGCCAGCAAACCG	
cyuP_1_F_RTqPCR	GCCGCACAGTTTATTAGCGG	
cyuP_1_R_RTqPCR	TACGATCCCTTGCGTTGCTT	
cyuR_1_F_RTqPCR	CGACGGTATCCTTATCGGCA	
cyuR_1_R_RTqPCR	CGTGACAAAGCGGCAATACC	
dcuR_down_R	ATGTAACTTGTGCCAGGGT	
dcuR_up	GCGTCGGTTTAGCACTTGTC	
decR_down	ACATCAGGATCTGCCCGGTA	
decR_up	GACGCACTGGCTGGACTATC	
fis_down_R	TAAGTTCGGCAAGCGCATCA	
fis_down_R	TAAGTTCGGCAAGCGCATCA	
fis_up_F	ATTGAGGATGCCAGCGAACA	
fis_up_F	ATTGAGGATGCCAGCGAACA	
fnr_2_F_RTqPCR	TGGCGGTTGTGCTATCCATT	
fnr_2_R_RTqPCR	TCACCAGCCTTAAACAGCGT	
fnr_downstream	TGTGCCAGCTTGTTCACT	
fnr_upstream	GGTTGGTCGTCCTGGTTAGG	
fnrS_1_F_RTqPCR	CGTTGCGCTCCATATTGTCT	
fnrS_1_R_RTqPCR	TCGGCGTCGTACGAATCAA	
ihfA_down_R	GCTGAAGTGTGTCATGGCGTTG	
ihfA_up_F	CCAGCCGTACTACTCGAAGAA	
ihfB_down_R	CCCGACAGGTGCTTTTCTCT	
ihfB_up_F	GCCGCCCTTAATCAATGCAG	
k1	CAGTCATAGCCGAATAGCCT	Reference: Baba <i>et al.</i> (2006).
k2	CGGTGCCCTGAATGAACTGC	
narL_down_R	CAGCGTGTCTCCTCCAATGA	
narL_up_F	GTGCCTGAAAATGCCATCCG	
nfeR_down_R	GGTTGGAGCACAAATGCCTGA	

Primer name	Primer sequence (5' --> 3')	Notes
nfeR_up_F	CCTTATCCGGCCTACGTGTG	
rst_down_R	CCAACCAGCAGAGACATCACA	
rst_up_R	ACCAGCCCCATCTTTTACCT	

3.6 Materials and Methods

Bacterial strains, media, and chemicals. Bacterial strains used in this study are listed in **Table S7**. *E. coli* was grown in Lysogeny broth (LB) supplemented with 0.05% L-Cysteine under anaerobic conditions at 37°C unless otherwise noted. L-Cysteine was purchased from Sigma-Aldrich (St. Louis, MO).

Assaying whole cell *E. coli* depletion of azo compounds. Cultures were grown in LB medium supplemented with 250 µM of azo dye and 0.05% L-Cysteine. Azo dye stocks were prepared in DMSO at a concentration of 25 mM and used at a final concentration of 1% (v/v). Cultures were incubated at 37°C. All incubations were performed in a COY Laboratory Products Inc anaerobic chamber (Grass Lake, MI) with the following atmosphere: 5% H₂, 20% CO₂, and 75% N₂. All reagents were equilibrated in the anaerobic chamber for at least 24 hours before use. After 24 hours of growth, samples were removed, spun down to remove cells, and analyzed for residual dye concentration by measuring absorbance at appropriate wavelengths. Wavelengths used: 450 nm (FD&C Red No. 40, FD&C Yellow No. 6, D&C Orange No. 4), 360 nm (Balsalazide, Sulfasalazine, Olsalazine, Phenazopyridine), 420 nm (D&C Brown No. 1), 500 nm (D&C Red No. 6, FD&C Red No. 4), 535 nm (D&C Red No. 33). Residual dye concentration was determined by creating a standard curve of dye concentrations in LB and measuring absorbance at appropriate wavelengths. Concentration was calculated using the equation of the fitted standard curve line.

Clean deletion creation from Keio collection strains. Competent cells were created from cultures of the Keio collection knockouts for *azoR::Kan^R* and *fnr::Kan^R* knockouts. Cells were electroporated with vector pSIJ8 (88) and transformants were selected for on

LB supplemented with 25 µg/ml ampicillin at 30°C. Removal of the kanamycin cassette at FRT sites was done on liquid culture of transformant using 50 mM L-rhamnose for 4 hours, then patch streaked to identify loss of kanamycin resistance. Transformants were streak-purified two times, then PCR was performed for confirmation of kanamycin cassette removal. Strains were cured of plasmid at 37°C and then patch-streaked to confirm loss of ampicillin resistance.

Purified enzyme preparation. The ASKA collection (38) strain expressing *azoR* was streaked from glycerol stock on LB with 30 µg/ml chloramphenicol. A single colony was used to inoculate an overnight culture of 25 ml of LB with 30 µg/ml chloramphenicol aerobically at 37°C. The overnight was used to inoculate 1 L of LB with 30 µg/ml chloramphenicol at a 1:10 dilution factor. Sample was grown with shaking aerobically to an optical density of 0.5 at 600 nm. Expression of *azoR* was induced with isopropyl β-d-1-thiogalactopyranoside (IPTG) at 1 mM concentration for 3 hours. Cells were then cooled on ice water for 10 mins and centrifuged at 3,500 x g for 15 mins, washed with ice cold PBS, spun again, then cell pellets were frozen overnight at -20°C. Pellets were thawed for on ice, then resuspended in Buffer A (50 mM HEPES, 300 mM NaCl, 10 mM imidazole, pH 7.5) and sonicated with the following program: 5 mins, 4 sec on, 4 sec off, at 30% amplitude. Protease tablets (Sigma Aldrich) were crushed in buffer A and added to the sonicated lysate. Lysate was centrifuged for 20 mins at 16,000 RPM. Supernatant was applied to a 5 ml nickel-nitrilotriacetic acid column (Qiagen, Valencia, CA) that had been equilibrated with buffer A at a rate of 1 ml/min. The column was washed with 5 column volumes of buffer A, then 5 column volumes of buffer B (50 mM HEPES, 300 mM NaCl, 20 mM imidazole, pH 7.5). Protein was eluted with Buffer C (50 mM Hepes, 300 mM NaCl,

300 mM imidazole) and collected in 5 ml portions. Fractions were analyzed for protein content using a sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gel. Fractions containing proteins corresponding to AzoR molecular weight were pooled and dialyzed with 3 L of buffer D (50 mM Hepes, 300 mM NaCl, 15% glycerol by weight, in cold water, pH 7.5) overnight at 4°C. Protein concentration was determined using Nanodrop (Thermo Scientific).

Purified Enzyme Experiments. AzoR reactions with azo dyes were carried out in 20 mM sodium phosphate buffer, 100 µM azo dye (unless otherwise noted), 2000 µM NADH, and 20 µM FMN with varying concentrations of enzyme as appropriate. Reactions were activated by the addition of NADH. Dye concentration was monitored at absorbance 450 nm every 30 seconds after a linear shake of 15 seconds. All reactions were performed aerobically at 37°C.

Keio collection screening for alternative azo reducers. Homologs to genes implicated in azoreduction by *Shewanella oneidensis* MR-1 (40) were identified by first gene name and second gene function (tested or predicted) in *E. coli*. Other strains of interest were identified based on gene name and function. All strains were cross referenced on EcoCyc (41). Strains were revived from glycerol stocks on LB agar with 25 µg/ml kanamycin. Single isolates were picked from agar plates and grown overnight in 1 ml of LB supplemented with 0.05% L-Cysteine in a 96-deep well plate under anaerobic conditions (see above). For azoreduction assays, previously anaerobically-equilibrated liquid media was supplemented with 250 µM FD&C Red No. 40 and 0.05% L-Cysteine. 1 ml of media was aliquoted to 2 ml deep 96 well plates and strains were inoculated at 1:100 from overnight cultures in triplicate. Plates were sealed with TempPlate Sealing Foil (USA

Scientific, Cat #2923-0110) and allowed to grow for 24 hours. After 24 hours, cells were spun down and 100 μ l of supernatant was used to measure dye absorbance at 450 nm. Proportion dye remaining was calculated by dividing the absorbance of each well by the average absorbance of uninoculated control wells. Where standard curves were used, a 1:2 dilution series of FD&C Red No. 40 was made in LB with 0.05% L-Cysteine in triplicate and dye absorbance measured at 450 nm. Concentrations were calculated using the equation of the standard curve line in Excel.

Validation of Keio collection knockouts. To ensure the knockouts used from the Keio collection were indeed the annotated gene, primers were designed up and downstream of the gene of interest. PCR and Sanger sequencing of the kanamycin insert and surrounding regions were performed to ensure the expected band size and gene alignment. Primers are listed in **Table S8**.

Complement strain construction and assay. Complement strains were made by preparing competent cells for the background strain of interest. Plasmids were prepared by growing up overnight cultures from the ASKA strain of interest, then using Qiagen Plasmid Mini Kit (Qiagen; Cat #12125) kit to extract the pCA24N plasmid with the gene of interest. The plasmid preparation was desalted using a membrane filter on water, then electroporated to the background strain of interest. Correct insertion was ensured by growth of the transformant on chloramphenicol. A mini prep was also made from the transformant and PCR and Sanger sequencing done to ensure the expected gene insertion. The same procedure was used to introduce the pBD-pLac and pBD-pLac-fnrS into strains of interest.

RNA-Sequencing and proteomics sample growth and exposure. 100 ml of overnight cultures of both *wt* and Δ *fnr* *E. coli* were grown in LB supplemented with 0.05% L-Cysteine under anaerobic conditions. Overnight culture OD_{600 nm} was measured for both cultures and normalized to a starting OD_{600 nm} of 0.05 in 100ml of LB with 0.05% L-Cysteine in 12 replicates for each strain. OD_{600 nm} was monitored until mid-exponential phase for each strain (previously determined). At mid-exponential 3 replicates for each strain were treated with a final concentration of 250 μ M FD&C Red No. 40 in DMSO, or 1% DMSO. After 45 mins of exposure, 15 ml of the sample was removed for RNA extraction, while the remaining volume was used for proteomics. Both volumes samples were spun down to pellet cells, supernatant removed, and cell pellets were flash frozen in liquid nitrogen then stored at -80°C. The remaining samples were allowed to reach stationary phase, then they were treated with a final concentration of 250 μ M FD&C Red No. 40 in DMSO, or 1% DMSO. Samples were processed the same way as for mid-exponential.

RNA-sequencing sample preparation. 1 μ l of TRI Reagent (Sigma Aldrich catalog number T9424) was added to bacterial pellets and incubated for 10 mins, samples were then transferred to 2 ml Lysing Matrix E tubes (MP Biomeicals, catalog number 116914050). Cells were lysed for 5 mins in the bead beater at room temperature. 200 μ l of chloroform was added. Samples were vortexed for 15 sec and incubated at room temperature for 10 mins. Samples were then centrifuged at 16,000 x g for 15 mins at 4°C. 500 μ l of the upper aqueous phase was transferred to a new tube, 500 μ l of 100% ethanol added, and vortexed to mix. Mixture was transferred to a spin column (PureLink RNA Mini Kit; Life Technologies; catalog number 12183025) and centrifuged at \geq 12,000 x g for 30 sec, discarding flow-through, until all of the material had been added to the column.

To the spin column, 350 μ l Wash Buffer I (PureLink RNA Mini Kit; Life Technologies; catalog number 12183025) was added, then the column was centrifuged at $\geq 12,000 \times g$ for 30 sec, discarding flow-through. 80 μ l of PureLink DNase mix was added to the column and incubated at room temperature for 15 mins. 350 μ l of Wash Buffer I (Purelink RNA mini kit) was added and the column spun at $>12000 \times g$ for 30 sec. The column was transferred to a new collection tube, and 500 μ l Wash Buffer II was added, followed by centrifugation at $\geq 12,000 \times g$ for 30 sec, discarding flow-through. The column was centrifuged at $\geq 12,000 \times g$ for 60 sec and dried, then moved to a collection tube. 50 μ l RNase-free water was added, and the column was incubated at room temperature for 1 min. Finally, the column was centrifuged for 1 min at $\geq 12,000 \times g$, retaining the flow-through, which contained total RNA.

Samples were DNase treated again using TURBO-DNase (Ambion; ThermoFisher catalog number AM2238) and incubated at 37°C for 30 mins. RNA Ampure XP Beads were used to clean up the reactions. 1.8 volumes of the RNA Ampure XP beads were added to 1 volume of RNA sample and allowed to sit for 5 mins at room temperature. Tubes were placed on a magnetic stand until the liquid cleared. Then liquid was removed and beads washed with 200 μ l of 100% Ethanol. Samples were incubated for 30 sec, then removed. This was repeated twice more. Samples were allowed to dry for 5 mins, removed from the magnetic stand, and 30 μ l of RNase-free water was added. This was incubated for 2 min at room temperature then placed back on the magnetic rack where liquid was collected after turning clear.

rRNA was removed from total RNA using Ribominus Transcriptome Isolation Kit for Bacteria and Yeast (Invitrogen; catalog number K155004, LOT: 2116711), following

manufacturer's protocol. RNA fragmentation, cDNA synthesis, and library preparation were performed using the NEBNext Ultra RNA library Prep Kit for Illumina and NEBNext Multiplex Oligos for Illumina (Dual Index Primers) (Ipswich, MA). Samples were dual-end sequenced (2 x 75 bp) using the NextSeq Mid Output platform (**Table S2**). Reads were mapped to the *E. coli* K12 BW25113 genome sequence (NCBI Reference Sequence: GCA_000750555.1) using Bowtie2 (89) and HTSeq (90) was used to count the number of reads to *E. coli* genes. Differential gene expression was analyzed using limma (91). Differentially expressed genes were defined as transcripts exhibiting an absolute log₂ fold-change ≥1 and a FDR<0.1. Detection of expression was based on limma's default filtering parameters in the "filter genes by expression level" function.

Proteomics sample preparation. Cell pellets were washed three times with 5 ml of cold PBS and resuspended in 1 ml, transferred to 2 ml Lysing Matrix E tubes (MP Biomeicals, catalog number 116914050) and lysed via bead beating for 1 min, then resting tubes on ice for 2 mins. This was repeated twice. Samples were centrifuged for 15 min at 4°C at 16,000 x g. A BCA assay was performed to quantify protein concentration in the lysate supernatant. Urea was added to the samples to a target concentration of 8 M followed by an appropriate volume of Dithiothreitol to obtain a 5 mM concentration. Samples were then incubated at 60°C for 30 mins. While samples were incubating, the trypsin was pre-activated for 10 mins at 37°C. Samples were diluted 10-fold with 100 mM NH₄HCO₃. 1 M CaCl₂ was added at an appropriate volume to create a final sample concentration of 1 mM CaCl₂. Samples were then digested for 3 hours with Trypsin at 37°C at a concentration of 1 µg trypsin/ 50 µg protein. Samples were snap frozen.

Samples were cleaned with a C₁₈ column on a vacuum manifold. The column was conditioned with 3 ml of Methanol, then rinsed with 2 ml of 0.1% TFA acidified water. Sample was run through column, then column was rinsed with 4 ml of 95:5 H₂O:acetonitrile (ACN), 0.1% trifluoroacetic acid (TFA). The column was allowed to go to dryness, then eluted slowly to dryness with 1 ml of 80:20 ACN:H₂O, 0.1% TFA into a collection tube. Samples were concentrated in the speed-vac to a volume of approximately 50-100 µl. Protein quantification was done using a BCA test, then samples were stored at -80°C until analysis.

Proteomics. A Waters nano-Acquity M-Class dual pumping UPLC system (Milford, MA) was configured for on-line trapping of a 5 µL injection at 5 µL/min with reverse-flow elution into the analytical column at 300 nL/min. The trapping column was packed in-house (PNNL) using 360 µm o.d. fused silica (Polymicro Technologies Inc., Phoenix, AZ) with 5 mm Kasil frits for media retention and contained Jupiter C₁₈ media (Phenomenex, Torrance, CA) in 5 µm particle size for the trapping column (150 µm i.d. × 4 cm long) and 3 µm particle size for the analytical column (75 µm i.d. × 70 cm long). Mobile phases consisted of (A) water with 0.1% formic acid, and (B) acetonitrile with 0.1% formic acid. The following gradient profile was performed (min, %B): 0, 1; 2, 8; 25, 12; 85, 35; 105, 55; 110, 95; 115, 95; 117, 50; 119, 95; 121, 95; 123, 1.

MS analysis was performed using a Velos Orbitrap Elite mass spectrometer (Thermo Scientific, San Jose, CA) outfitted with an in-house made nano-electrospray ionization interface. Electrospray emitters were prepared using 150 µm o.d. × 20 µm i.d. chemically etched fused silica (92). The ion transfer tube temperature and spray voltage were 325 °C and 2.2 kV, respectively. Data were collected for 100 min following a 20 min

delay from sample injection. FT-MS spectra were acquired from 400 to 2000 m/z at a resolution of 35k (AGC target $3e6$) and while the top 12 FT-HCD-MS/MS spectra were acquired in data dependent mode with an isolation window of 2.0 m/z and at a resolution of 17.5k (AGC target $1e5$) using a normalized collision energy of 30 sec and a 30 sec exclusion time. Extracted MS data were run through the MSGF+ database and filtered by an FDR of >0.01 using the MSGF+ generated decoy database. Differential peptide abundance was analyzed using DEP (93) and limma (91) using a cutoff of an absolute \log_2 fold-change ≥ 1 and a $FDR < 0.1$.

Features found significantly differentially expressed (\log_2 fold change $> |1|$, $p_{adj} < 0.1$) in both the proteomics and transcriptomics data sets were imported to EcoCyc (41) for enrichment analysis. The gene list was analyzed using the "Genes enriched for transcriptional/translational regulators (direct only)". The following settings were used: p -value $p_{adj} < 0.1$, Fisher Exact Statistic, Benjamini-Hochberg Correction). Proportion of regulon represented was determined using the list of genes "directly regulated by gene" feature of EcoCyc for total regulon size. Then, the number of genes regulated by each regulator was divided by number of total genes regulated by that particular regulator.

Hydrogen sulfide quantification. H_2S was measured using the Cline reaction (94) with modifications. Samples were put into a zinc acetate buffer (16.7 mM) on ice at a ratio of 1:3 sample:buffer. Then 180 μl of zinc acetate solution was transferred and mixed with 20 μl of cline reagent (2 g of N,N-dimethyl-p-phenylenediamine sulfate with 3 g of $FeCl_3$ in 500 ml of cold 50% v/v HCl). The reaction was allowed to proceed for 20-30 minutes at room temperature in the dark. Then absorbances were measured at 670 nm. A

standard curve of sodium sulfide was created in zinc acetate and reacted with cline reagents for sulfide concentration calculations.

RT-qPCR assays. Samples were grown under anaerobic conditions with 250 μ M Red 40. At mid-exponential phase, 4.2 mM L-Cysteine was added to cultures for 15 min. Then samples were removed from anaerobic conditions, centrifuged for 5 min at $\geq 12,000 \times g$, supernatant removed, and pellets were frozen at -80°C . Samples were thawed and RNA extracted as previously described for RNA-sequencing. RT-qPCR was done using RioRad's 1-step Quantitative Reverse Transcription PCR kit (Bio-Rad, Hercules, CA, catalog number 1725150) according to kit protocol.

To analyze the differences between *wt* and Δ *fnr* samples, fold-change of sample ΔC_t minus the maximum ΔC_t for each gene across all samples in *wt* or Δ *fnr* *E. coli* was calculated. Within each replicate the C_t of the housekeeping gene was subtracted from the C_t of the gene of interest (ΔC_t). The maximum ΔC_t across samples from both strains was subtracted from all samples to give $\Delta\Delta C_t$. Fold-change was calculated as $2^{(-\Delta\Delta C_t)}$. To analyze the impact of L-Cysteine on gene expression, fold-change was calculated using expression of the sample with L-Cysteine compared to the vehicle. Within each replicate the C_t of the housekeeping gene was subtracted from the C_t of the gene of interest (ΔC_t). The ΔC_t from the paired replicate control (vehicle exposure) samples was subtracted from the exposed sample ΔC_t to give $\Delta\Delta C_t$. Fold-change was calculated as $2^{(-\Delta\Delta C_t)}$.

Double gene knockout creation. P1 lysates were generated of each strain of interest carrying the kanamycin resistance cassette (Keio collection) adapting methods from previously described techniques (95). Briefly, 150 μ l of overnight culture in LB

supplemented with 12.5 µg/ml kanamycin was mixed with 1 to 25 µl of P1 phage (previously propagated from ATCC on MG1655). This mixture was incubated for 10 min at 37°C to aid absorption, added to 3 ml of 0.7% agar, and overlaid on pre-warmed LB agar supplemented with 25 µg/ml kanamycin and 10 mM MgSO₄. Plates were incubated overnight at 37°C and phage were harvested by adding 5 ml of SM buffer, incubating at room temperature for 10 mins, and breaking the top agar for phage harvest. The mixture was briefly centrifuged to pellet agar, then supernatant was passed through a 100 µm cell straining, then 0.45 µm syringe filter. Lysates were stored at 4°C.

A clean deletion of the recipient strain of interest was created using the above described method with pSIJ8 (88). To transduce the clean recipient strain, 1 ml of an overnight culture of recipient strain was pelleted and resuspended in $\frac{1}{3}$ volume of LB 10 mM MgSO₄, 5 mM CaCl₂. 100 µl of cells were mixed with 1 µl to 10 µl of P1 lysate and incubated for 60 min at 37°C. 200 µl of 1M sodium citrate was added with 1 ml of LB to minimize secondary infection. This mixture was incubated at 37°C for 2 hrs, then plated on LB supplemented with 10 mM sodium citrate and 25 µg/ml kanamycin to select for transductants. Transduction was confirmed using PCR and Sanger sequencing. Azoreduction and sulfide activity were tested as described above.

Statistical analysis. For multiple comparisons of data, an ordinary one-way analysis of variance (ANOVA) test with Dunnet's correction was chosen, unless otherwise specified. For the targeted knockout screen, a 95% confidence interval of averages from two experiments was calculated using GraphPad Prism 9. Michaelis-menten curves were created in Prism 9. For proteomics and transcriptomics analysis, R package limma was used to determine the fold-change and adjusted *p*-values. Feature count differences and

Mann-Whitney significance testing were done in R. For the regulator enrichment analysis, EcoCyc's enrichment analysis for "Genes enriched for transcriptional/translational regulators (direct only)" was used with a Fisher's exact test with Benjamini-Hochberg Correction.

Data and material availability. RNA-seq data are available through the NCBI Gene Expression Omnibus (GEO) online data repository under accession number GSE235465. Proteomics data is available through ProteomeXchange.org, accession number TBD. All other data are provided in the supplemental tables or available upon request.

3.7 References

1. Spanogiannopoulos P, Bess EN, Carmody RN, Turnbaugh PJ. 2016. The microbial pharmacists within us: a metagenomic view of xenobiotic metabolism. *Nat Rev Microbiol* 14:273–287.
2. Collins SL, Patterson AD. 2020. The gut microbiome: an orchestrator of xenobiotic metabolism. *Acta Pharm Sin B* 10:19–32.
3. Clarke G, Sandhu KV, Griffin BT, Dinan TG, Cryan JF, Hyland NP. 2019. Gut Reactions: Breaking Down Xenobiotic-Microbiome Interactions. *Pharmacol Rev* 71:198–224.
4. Yee Yip L, Chun Yong Chan E. 2015. Special Section on Drug Metabolism and the Microbiome—Minireview Investigation of Host–Gut Microbiota Modulation of Therapeutic Outcome. *Drug Metab Dispos* 43:1619–1631.
5. Koppel N, Maini Rekdal V, Balskus EP. 2017. Chemical transformation of xenobiotics by the human gut microbiota. *Science* 356:eaag2770.
6. Nayak RR, Turnbaugh PJ. 2016. Mirror, mirror on the wall: which microbiomes will help heal them all? *BMC Med* 14:72.
7. Peppercorn MA. 1984. Sulfasalazine: Pharmacology, Clinical Use, Toxicity, and Related New Drug Development. *Ann Intern Med* 101:377.
8. Center for Food Safety, Nutrition A. Color Additives History. US Food and Drug Administration. FDA. <https://www.fda.gov/industry/color-additives/color-additives-history>. Retrieved 24 June 2022.
9. Misal SA, Gawai KR. 2018. Azoreductase: a key player of xenobiotic metabolism. *Bioresources and Bioprocessing* <https://doi.org/10.1186/s40643-018-0206-8>.

10. Chung KT, Fulk GE, Egan M. 1978. Reduction of azo dyes by intestinal anaerobes. *Appl Environ Microbiol* 35:558–562.
11. Chung KT, Stevens SE Jr, Cerniglia CE. 1992. The reduction of azo dyes by the intestinal microflora. *Crit Rev Microbiol* 18:175–190.
12. Wang R-F, Chen H, Paine DD, Cerniglia CE. 2004. Microarray method to monitor 40 intestinal bacterial species in the study of azo dye reduction. *Biosens Bioelectron* 20:699–705.
13. Zahran SA, Ali-Tammam M, Hashem AM, Aziz RK, Ali AE. 2019. Azoreductase activity of dye-decolorizing bacteria isolated from the human gut microbiota. *Sci Rep* 9:5508.
14. Ryan A. 2017. Azoreductases in drug metabolism. *Br J Pharmacol* 174:2161–2173.
15. Gingell R, Bridges JW, Williams RT. 1971. The role of the gut flora in the metabolism of prontosil and neoprontosil in the rat. *Xenobiotica* 1:143–156.
16. Sousa T, Yadav V, Zann V, Borde A, Abrahamsson B, Basit AW. 2014. On the Colonic Bacterial Metabolism of Azo-Bonded Prodrug of 5-Aminosalicylic Acid. *J Pharm Sci* 103:3171–3175.
17. Stevens LJ, Kuczek T, Burgess JR, Stochelski MA, Arnold LE, Galland L. 2013. Mechanisms of behavioral, atopic, and other reactions to artificial food colors in children. *Nutr Rev* 71:268–281.
18. Zou L, Spanogiannopoulos P, Pieper LM, Chien H-C, Cai W, Khuri N, Pottel J, Vora B, Ni Z, Tsakalozou E, Zhang W, Shoichet BK, Giacomini KM, Turnbaugh PJ. 2020. Bacterial metabolism rescues the inhibition of intestinal drug absorption by food and drug additives. *Proc Natl Acad Sci U S A* 117:16009–16018.

19. Feng J, Cerniglia CE, Chen H. 2012. Toxicological significance of azo dye metabolism by human intestinal microbiota. *Front Biosci* 4:568–586.
20. Yahagi T, Degawa M, Seino Y, Matsushima T, Nagao M, Sugimura T, Hashimoto Y. 1975. Mutagenicity of carcinogenic azo dyes and their derivatives. *Cancer Lett* 1:91–96.
21. Chung K-T. 1983. The significance of azo-reduction in the mutagenesis and carcinogenesis of azo dyes. *Mutation Research/Reviews in Genetic Toxicology* 114:269–281.
22. He Z, Chen L, Catalan-Dibene J, Bongers G, Faith JJ, Suebsuwong C, DeVita RJ, Shen Z, Fox JG, Lafaille JJ, Furtado GC, Lira SA. 2021. Food colorants metabolized by commensal bacteria promote colitis in mice with dysregulated expression of interleukin-23. *Cell Metab* 33:1358–1371.e5.
23. Sandhya S. 2010. Biodegradation of Azo Dyes Under Anaerobic Condition: Role of Azoreductase, p. 39–57. *In* Atacag Erkurt, H (ed.), *Biodegradation of Azo Dyes*. Springer Berlin Heidelberg, Berlin, Heidelberg.
24. Nakanishi M, Yatome C, Ishida N, Kitade Y. 2001. Putative ACP phosphodiesterase gene (*acpD*) encodes an azoreductase. *J Biol Chem* 276:46394–46399.
25. Nachiyar CV, Rajakumar GS. 2005. Purification and characterization of an oxygen insensitive azoreductase from *Pseudomonas aeruginosa*. *Enzyme Microb Technol* 36:503–509.
26. Wang C-J, Hagemeyer C, Rahman N, Lowe E, Noble M, Coughtrie M, Sim E, Westwood I. 2007. Molecular Cloning, Characterisation and Ligand-bound Structure of an Azoreductase from *Pseudomonas aeruginosa*. *J Mol Biol* 373:1213–1228.

27. Ryan A, Kaplan E, Nebel J-C, Polycarpou E, Crescente V, Lowe E, Preston GM, Sim E. 2014. Identification of NAD(P)H Quinone Oxidoreductase Activity in Azoreductases from *P. aeruginosa*: Azoreductases and NAD(P)H Quinone Oxidoreductases Belong to the Same FMN-Dependent Superfamily of Enzymes. *PLoS One* 9:e98551.
28. Crescente V, Holland SM, Kashyap S, Polycarpou E, Sim E, Ryan A. 2016. Identification of novel members of the bacterial azoreductase family in *Pseudomonas aeruginosa*. *Biochem J* 473:549–558.
29. Chen H, Wang R-F, Cerniglia CE. 2004. Molecular cloning, overexpression, purification, and characterization of an aerobic FMN-dependent azoreductase from *Enterococcus faecalis*. *Protein Expr Purif* 34:302–310.
30. Liu Z-J, Chen H, Shaw N, Hopper SL, Chen L, Chen S, Cerniglia CE, Wang B-C. 2007. Crystal structure of an aerobic FMN-dependent azoreductase (AzoA) from *Enterococcus faecalis*. *Arch Biochem Biophys* 463:68–77.
31. Chalansonnet V, Mercier C, Orenga S, Gilbert C. 2017. Identification of *Enterococcus faecalis* enzymes with azoreductases and/or nitroreductase activity. *BMC Microbiol* 17:126.
32. Morrison JM, John GH. 2013. The non-enzymatic reduction of azo dyes by flavin and nicotinamide cofactors under varying conditions. *Anaerobe* 23:87–96.
33. Dai Q, Zhang S, Liu H, Huang J, Li L. 2020. Sulfide-mediated azo dye degradation and microbial community analysis in a single-chamber air cathode microbial fuel cell. *Bioelectrochemistry* 131:107349.

34. Wolfson SJ, Hitchings R, Peregrina K, Cohen Z, Khan S, Yilmaz T, Malena M, Goluch ED, Augenlicht L, Kelly L. 2022. Bacterial hydrogen sulfide drives cryptic redox chemistry in gut microbial communities. *Nat Metab* 4:1260–1270.
35. Haiser HJ, Turnbaugh PJ. 2013. Developing a metagenomic view of xenobiotic metabolism. *Pharmacol Res* 69:21–31.
36. Zimmermann M, Zimmermann-Kogadeeva M, Wegmann R, Goodman AL. 2019. Mapping human microbiome drug metabolism by gut bacteria and their genes. *Nature* <https://doi.org/10.1038/s41586-019-1291-3>.
37. Bisanz JE, Soto-Perez P, Noecker C, Aksenov AA, Lam KN, Kenney GE, Bess EN, Haiser HJ, Kyaw TS, Yu FB, Rekdal VM, Ha CWY, Devkota S, Balskus EP, Dorrestein PC, Allen-Vercoe E, Turnbaugh PJ. 2020. A Genomic Toolkit for the Mechanistic Dissection of Intractable Human Gut Bacteria. *Cell Host Microbe* 27:1001–1013.e9.
38. Kitagawa M, Ara T, Arifuzzaman M, Ioka-Nakamichi T, Inamoto E, Toyonaga H, Mori H. 2005. Complete set of ORF clones of Escherichia coli ASKA library (A complete set of E. coli K-12 ORF archive): unique resources for biological research. *DNA Res* 12:291–299.
39. Zhang H-K, Lu H, Wang J, Liu G-F, Zhou J-T, Xu M-Y. 2013. Global transcriptome analysis of Escherichia coli exposed to immobilized anthraquinone-2-sulfonate and azo dye under anaerobic conditions. *Appl Microbiol Biotechnol* 97:6895–6905.
40. Brigé A, Motte B, Borloo J, Buyschaert G, Devreese B, Van Beeumen JJ. 2008. Bacterial decolorization of textile dyes is an extracellular process requiring a multicomponent electron transfer pathway. *Microb Biotechnol* 1:40–52.

41. Karp PD, Ong WK, Paley S, Billington R, Caspi R, Fulcher C, Kothari A, Krummenacker M, Latendresse M, Midford PE, Subhraveti P, Gama-Castro S, Muñiz-Rascado L, Bonavides-Martinez C, Santos-Zavaleta A, Mackie A, Collado-Vides J, Keseler IM, Paulsen I. 2018. The EcoCyc Database. *EcoSal Plus* 8.
42. Uden G, Trageser M. 1991. Oxygen regulated gene expression in *Escherichia coli*: Control of anaerobic respiration by the FNR protein. *Antonie Van Leeuwenhoek* 59:65–76.
43. Spiro S, Guest JR. 1990. FNR and its role in oxygen-regulated gene expression in *Escherichia coli*. *FEMS Microbiol Lett* 75:399–428.
44. Myers KS, Yan H, Ong IM, Chung D, Liang K, Tran F, Keleş S, Landick R, Kiley PJ. 2013. Genome-scale analysis of *Escherichia Coli* FNR reveals complex features of transcription factor binding. *PLoS Genet* 9:e1003565.
45. Srinivasan S, Venkatesh KV. 2014. Steady state analysis of the genetic regulatory network incorporating underlying molecular mechanisms for anaerobic metabolism in *Escherichia coli*. *Mol Biosyst* 10:562–575.
46. Kang Y, Weber KD, Qiu Y, Kiley PJ, Blattner FR. 2005. Genome-wide expression analysis indicates that FNR of *Escherichia coli* K-12 regulates a large number of genes of unknown function. *J Bacteriol* 187:1135–1160.
47. Gunsalus RP, Park SJ. 1994. Aerobic-anaerobic gene regulation in *Escherichia coli*: control by the ArcAB and Fnr regulons. *Res Microbiol* 145:437–450.
48. Ito K, Nakanishi M, Lee W-C, Sasaki H, Zenno S, Saigo K, Kitade Y, Tanokura M. 2006. Three-dimensional structure of AzoR from *Escherichia coli*. An oxidoreductase conserved in microorganisms. *J Biol Chem* 281:20567–20576.

49. Blake T, Barnard A, Busby SJW, Green J. 2002. Transcription activation by FNR: evidence for a functional activating region 2. *J Bacteriol* 184:5855–5861.
50. Bilous PT, Weiner JH. 1985. Dimethyl sulfoxide reductase activity by anaerobically grown *Escherichia coli* HB101. *J Bacteriol* 162:1151–1155.
51. Uden G, Bongaerts J. 1997. Alternative respiratory pathways of *Escherichia coli*: energetics and transcriptional regulation in response to electron acceptors. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* 1320:217–234.
52. Kim S, Pevzner PA. 2014. MS-GF+ makes progress towards a universal database search tool for proteomics. *Nat Commun* 5:1–10.
53. Caglar MU, Houser JR, Barnhart CS, Boutz DR, Carroll SM, Dasgupta A, Lenoir WF, Smith BL, Sridhara V, Sydykova DK, Vander Wood D, Marx CJ, Marcotte EM, Barrick JE, Wilke CO. 2017. The *E. coli* molecular phenotype under different growth conditions. *Sci Rep* 7:45303.
54. Makinoshima H, Aizawa S-I, Hayashi H, Miki T, Nishimura A, Ishihama A. 2003. Growth phase-coupled alterations in cell structure and function of *Escherichia coli*. *J Bacteriol* 185:1338–1345.
55. Kobayashi A, Hirakawa H, Hirata T, Nishino K, Yamaguchi A. 2006. Growth phase-dependent expression of drug exporters in *Escherichia coli* and its contribution to drug tolerance. *J Bacteriol* 188:5693–5703.
56. Keseler IM, Collado-Vides J, Santos-Zavaleta A, Peralta-Gil M, Gama-Castro S, Muñiz-Rascado L, Bonavides-Martinez C, Paley S, Krummenacker M, Altman T, Kaipa P, Spaulding A, Pacheco J, Latendresse M, Fulcher C, Sarker M, Shearer AG,

- Mackie A, Paulsen I, Gunsalus RP, Karp PD. 2011. EcoCyc: a comprehensive database of *Escherichia coli* biology. *Nucleic Acids Res* 39:D583–90.
57. Keseler IM, Gama-Castro S, Mackie A, Billington R, Bonavides-Martínez C, Caspi R, Kothari A, Krummenacker M, Midford PE, Muñiz-Rascado L, Ong WK, Paley S, Santos-Zavaleta A, Subhraveti P, Tierrafría VH, Wolfe AJ, Collado-Vides J, Paulsen IT, Karp PD. 2021. The EcoCyc Database in 2021. *Front Microbiol* 12:711077.
58. Inada T, Takahashi H, Mizuno T, Aiba H. 1996. Down regulation of cAMP production by cAMP receptor protein in *Escherichia coli*: an assessment of the contributions of transcriptional and posttranscriptional control of adenylate cyclase. *Mol Gen Genet* 253:198–204.
59. Zheng D, Constantinidou C, Hobman JL, Minchin SD. 2004. Identification of the CRP regulon using in vitro and in vivo transcriptional profiling. *Nucleic Acids Res* 32:5874–5893.
60. Shimada T, Fujita N, Yamamoto K, Ishihama A. 2011. Novel roles of cAMP receptor protein (CRP) in regulation of transport and metabolism of carbon sources. *PLoS One* 6:e20081.
61. Kargeti M, Venkatesh KV. 2017. The effect of global transcriptional regulators on the anaerobic fermentative metabolism of *Escherichia coli*. *Mol Biosyst* 13:1388–1398.
62. Loddeke M, Schneider B, Oguri T, Mehta I, Xuan Z, Reitzer L. 2017. Anaerobic Cysteine Degradation and Potential Metabolic Coordination in *Salmonella enterica* and *Escherichia coli*. *J Bacteriol* 199.
63. Zhou Y, Imlay JA. 2022. *Escherichia coli* Uses a Dedicated Importer and Desulfidase To Ferment Cysteine. *MBio* e0296521.

64. Shimada T, Tanaka K, Ishihama A. 2016. Transcription factor DecR (YbaO) controls detoxification of L-cysteine in *Escherichia coli*. *Microbiology* 162:1698–1707.
65. Zhao H-Q, Huang S-Q, Xu W-Q, Wang Y-R, Wang Y-X, He C-S, Mu Y. 2019. Undiscovered Mechanism for Pyrogenic Carbonaceous Matter-Mediated Abiotic Transformation of Azo Dyes by Sulfide. *Environ Sci Technol* 53:4397–4405.
66. Li K, Xin Y, Xuan G, Zhao R, Liu H, Xia Y, Xun L. 2019. *Escherichia coli* Uses Separate Enzymes to Produce H₂S and Reactive Sulfane Sulfur From L-cysteine. *Front Microbiol* 10:298.
67. Hungate RE. 1950. The anaerobic mesophilic cellulolytic bacteria. *Bacteriol Rev* 14:1–49.
68. Fukushima RS, Weimer PJ, Kunz DA. 2003. Use of photocatalytic reduction to hasten preparation of culture media for saccharolytic *Clostridium* species. *Braz J Microbiol* 34:22–26.
69. Breznak JA, Costilow RN. 2014. Physicochemical Factors in Growth, p. 309–329. *In* *Methods for General and Molecular Microbiology*. ASM Press, Washington, DC, USA.
70. Seo SW, Kim D, Szubin R, Palsson BO. 2015. Genome-wide Reconstruction of OxyR and SoxRS Transcriptional Regulatory Networks under Oxidative Stress in *Escherichia coli* K-12 MG1655. *Cell Rep* 12:1289–1299.
71. Durand S, Storz G. 2010. Reprogramming of anaerobic metabolism by the FnrS small RNA. *Mol Microbiol* 75:1215–1231.
72. Boysen A, Møller-Jensen J, Kallipolitis B, Valentin-Hansen P, Overgaard M. 2010. Translational regulation of gene expression by an anaerobically induced small non-coding RNA in *Escherichia coli*. *J Biol Chem* 285:10690–10702.

73. Ito K, Nakanishi M, Lee W-C, Sasaki H, Zenno S, Saigo K, Kitade Y, Tanokura M. 2005. Crystallization and preliminary X-ray analysis of AzoR (azoreductase) from *Escherichia coli*. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 61:399–402.
74. Punj S, John GH. 2009. Purification and identification of an FMN-dependent NAD(P)H azoreductase from *Enterococcus faecalis*. *Curr Issues Mol Biol* 11:59–65.
75. Morrison JM, Wright CM, John GH. 2012. Identification, Isolation and characterization of a novel azoreductase from *Clostridium perfringens*. *Anaerobe* 18:229–234.
76. Bardi L, Marzona M. 2010. Factors Affecting the Complete Mineralization of Azo Dyes. *The Handbook of Environmental Chemistry* https://doi.org/10.1007/698_2009_50.
77. Mani A, Hameed SAS. 2016. Accelerated Production of Oxygen-Insensitive Azoreductase from Mutant *Pseudomonas* Species for Degradation Azo Dyes under Aerobic Condition. *Asian Journal of Chemistry* <https://doi.org/10.14233/ajchem.2016.20113>.
78. Clark DP. 1989. The fermentation pathways of *Escherichia coli*. *FEMS Microbiol Rev* 5:223–234.
79. Jones SA, Chowdhury FZ, Fabich AJ, Anderson A, Schreiner DM, House AL, Autieri SM, Leatham MP, Lins JJ, Jorgensen M, Cohen PS, Conway T. 2007. Respiration of *Escherichia coli* in the mouse intestine. *Infect Immun* 75:4891–4899.
80. Jones SA, Gibson T, Maltby RC, Chowdhury FZ, Stewart V, Cohen PS, Conway T. 2011. Anaerobic respiration of *Escherichia coli* in the mouse intestine. *Infect Immun* 79:4218–4226.
81. Zhou Y, Imlay JA. 2020. *Escherichia coli* K-12 Lacks a High-Affinity Assimilatory Cysteine Importer. *MBio* 11.

82. Braccia DJ, Jiang X, Pop M, Hall AB. 2021. The Capacity to Produce Hydrogen Sulfide (H₂S) via Cysteine Degradation Is Ubiquitous in the Human Gut Microbiome. *Front Microbiol* 12:705583.
83. Loubinoux J, Bronowicki J-P, Pereira IAC, Mougénel J-L, Faou AE. 2002. Sulfate-reducing bacteria in human feces and their association with inflammatory bowel diseases. *FEMS Microbiol Ecol* 40:107–112.
84. Rey FE, Gonzalez MD, Cheng J, Wu M, Ahern PP, Gordon JI. 2013. Metabolic niche of a prominent sulfate-reducing human gut bacterium. *Proc Natl Acad Sci U S A* 110:13582–13587.
85. Pitcher MC, Beatty ER, Cummings JH. 2000. The contribution of sulphate reducing bacteria and 5-aminosalicylic acid to faecal sulphide in patients with ulcerative colitis. *Gut* 46:64–72.
86. Barton LL, Ritz NL, Fauque GD, Lin HC. 2017. Sulfur Cycling and the Intestinal Microbiome. *Dig Dis Sci* 62:2241–2257.
87. Arnold LE, Lofthouse N, Hurt E. 2012. Artificial food colors and attention-deficit/hyperactivity symptoms: conclusions to dye for. *Neurotherapeutics* 9:599–609.
88. Jensen SI, Lennen RM, Herrgård MJ, Nielsen AT. 2016. Seven gene deletions in seven days: Fast generation of *Escherichia coli* strains tolerant to acetate and osmotic stress. *Sci Rep* 5:17874.
89. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359.
90. Anders S, Pyl PT, Huber W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31:166–169.

91. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43:e47.
92. Kelly RT, Page JS, Luo Q, Moore RJ, Orton DJ, Tang K, Smith RD. 2006. Chemically etched open tubular and monolithic emitters for nanoelectrospray ionization mass spectrometry. *Anal Chem* 78:7796–7801.
93. Zhang X, Smits AH, van Tilburg GB, Ovaa H, Huber W, Vermeulen M. 2018. Proteome-wide identification of ubiquitin interactions using UbiA-MS. *Nat Protoc* 13:530–550.
94. Cline JD. 1969. Spectrophotometric determination of hydrogen sulfide in natural waters. *Limnol Oceanogr* 14:454–458.
95. Thomason LC, Costantino N, Court DL. 2007. *E. coli* Genome Manipulation by P1 Transduction, p. 1.17.1–1.17.8. *In* *Current Protocols in Molecular Biology*. John Wiley & Sons, Inc., Hoboken, NJ, USA.

Chapter 4: Gut Microbiome Environment Impacts Sulfasalazine pharmacokinetics and Sulfide Levels

4.1 Introduction

While research into the interaction of the microbiome and many drugs is in its infancy, other pharmaceuticals, such as sulfasalazine (SSZ) have been rationally designed to utilize microbial activity (1–3). SSZ is an anti-inflammatory prodrug used for inflammatory bowel disease (IBD) and rheumatoid arthritis (RA) (1). It is thought that poor absorption of SSZ allows for delivery to the colon, the site of microbial activators, resulting in site-specific release of the active moiety, 5-ASA, and the inactive sulfapyridine (SP) (1, 2) (**Figure 4.1a**). Azoreductase enzymes are considered the microbial method of activation (2, 4, 5). However, there is a lack of direct evidence of azoreductase causality *in vivo*.

Furthermore, recent research from our lab and others shows that microbial azoreductase enzymes are not the only source of azo compound reduction. Microbially produced hydrogen sulfide (H₂S) is sufficient to deplete azo compounds (6, 7). I previously described how this process is regulated in the model gut microbe, *Escherichia coli* by the fumarate and nitrate reductase (*fnr*) regulator and that L-Cysteine (L-Cys) are required for azo depletion *in vitro* (7). These findings leave open the question of how sulfide and environmental conditions impact the azoreduction *in vivo*? Moreover, while the microbiome is considered a key player in SSZ activation, no studies have looked at the impact of microbiota or diet, a key modulator of microbiome environment (8), on SSZ pharmacokinetics.

In this chapter, I focus on SSZ as a model azo drug to better understand how microbial azo depletion occurs *in vivo*. I explore the impact of microbial community

composition, intestinal sulfide concentrations, and diet as factors that could impact azoreduction. Additionally, I hypothesize sulfide produced from microbial L-Cys metabolism impacts SSZ PK. Overall, we found that the presence of a microbiota expedites SSZ conversion to SP as well as the removal of SP from plasma and stool samples. Additionally, the presence of a microbiota leads to greater concentrations of fecal and cecal sulfide. Similarly, antibiotics are sufficient to elicit a change in SSZ PK and sulfide levels. Lastly, we found gastrointestinal sulfide levels were robust to dietary perturbation in mice with various microbiota. Overall, these findings highlight the need for a better understanding of the *in vivo* role of the microbiome on SSZ PK. Our findings serve as a basis for future studies to better understand the impact of microbial community environment on inter-individual variability and the efficacy of SSZ.

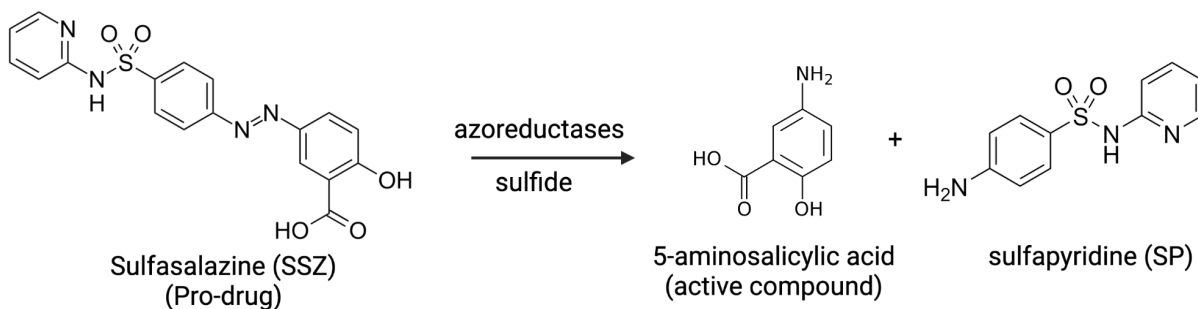


Figure 4.1. Sulfasalazine can be activated by enzymatic and non-enzymatic mechanisms. Sulfasalazine is reduced to 5-aminosalicylic acid and sulfapyridine.

4.2 Results

4.2.1 The gut microbiota impacts the pharmacokinetics of sulfasalazine.

While the microbiota is touted as the activation agent of SSZ (1, 2, 5), no PK experiments have been published to further evaluate microbial impact on activation in germ-free (GF) and conventionalized mice (CONV-D). To evaluate the impact of colonization on SSZ activation and PK of both the parent (SSZ) and metabolite (SP), GF mice were either colonized with the cecal contents of a conventionally raised specific pathogen free mouse (CONV-R), or left GF (n=6 mice/group). After two weeks of colonization, mice were dosed with 100 mg/kg SSZ and both plasma and fecal samples were collected for 22 hrs for LC-MS/MS analysis (**Figure 4.2a**). First, we assessed the PK dynamics of the prodrug, SSZ (**Figure 4.2b**). CONV-D mice demonstrated a faster decrease in SSZ GF mice in both plasma (**Figure 4.2c**) and fecal (**Figure 4.2g**) samples. In plasma samples, the maximum concentration (C_{max}) of SSZ was significantly higher in GF mice than CONV-D mice (**Figure 4.2d**). Plasma total area under the curve (AUC) for each individual mouse also trended higher in GF mice compared to CONV-D (**Figure 4.2e**). In feces, significantly higher SSZ AUC and C_{max} were observed (**Figure 4.2h-i**). However, the time at which the SSZ concentration was highest (T_{max}) was not significantly different between groups in either plasma or feces.

Next, we measured the same PK parameters for the metabolite SP (**Figure 4.3a**). Here, we saw inverse trends as with the parent compound. By LC-MS/MS, both plasma and fecal levels of SP were higher in CONV-D mice than GF mice (**Figure 4.3b, f**). The CONV-D group has significantly higher SP AUC and C_{max} than the GF group in both

plasma and fecal samples (**Figure 4.3c-d, g-h**). Additionally, The GF group had significantly later T_{\max} than the CONV-D group in both sample types (**Figure 4.3e, i**).

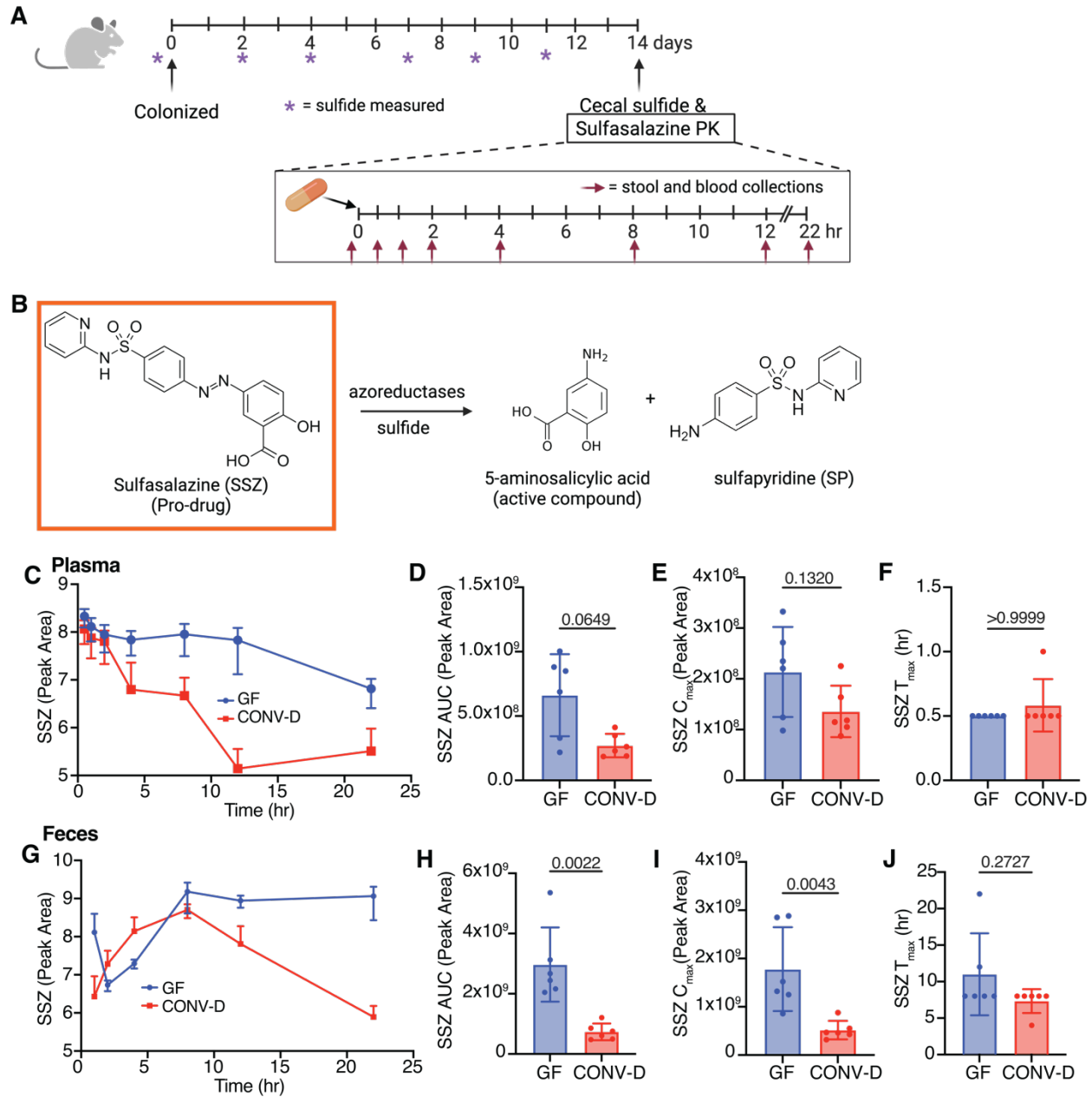


Figure 4.2. Gut microbiota presence impacts Sulfasalazine activation and pharmacokinetics. (a) Experimental schematic. Mice are conventionalized or remain GF for two weeks. Fecal sulfide was measured intermittently. After two weeks mice were dosed with 100 mg/kg SSZ and blood and feces were collected for pharmacokinetics. (b) SSZ activation reaction. The parent compound, measured in the rest of the figure, is highlighted in orange. (c, g) LC-MS/MS peak area of plasma (c) and fecal (g) SSZ, plotted on a log scale. (e-f; h-j) PK parameters for plasma or fecal SSZ, respectively (d, h) area under the curve (AUC), (e, i) peak concentration (C_{max}), and (f, j) the time at which concentrations are highest (t_{max}). ($n=6$ mice/group, values and bars represent mean \pm standard deviation, statistics done with Mann-Whitney t-test, fecal and plasma LC-MS/MS peak areas are normalized to background levels pre-SSZ dose).

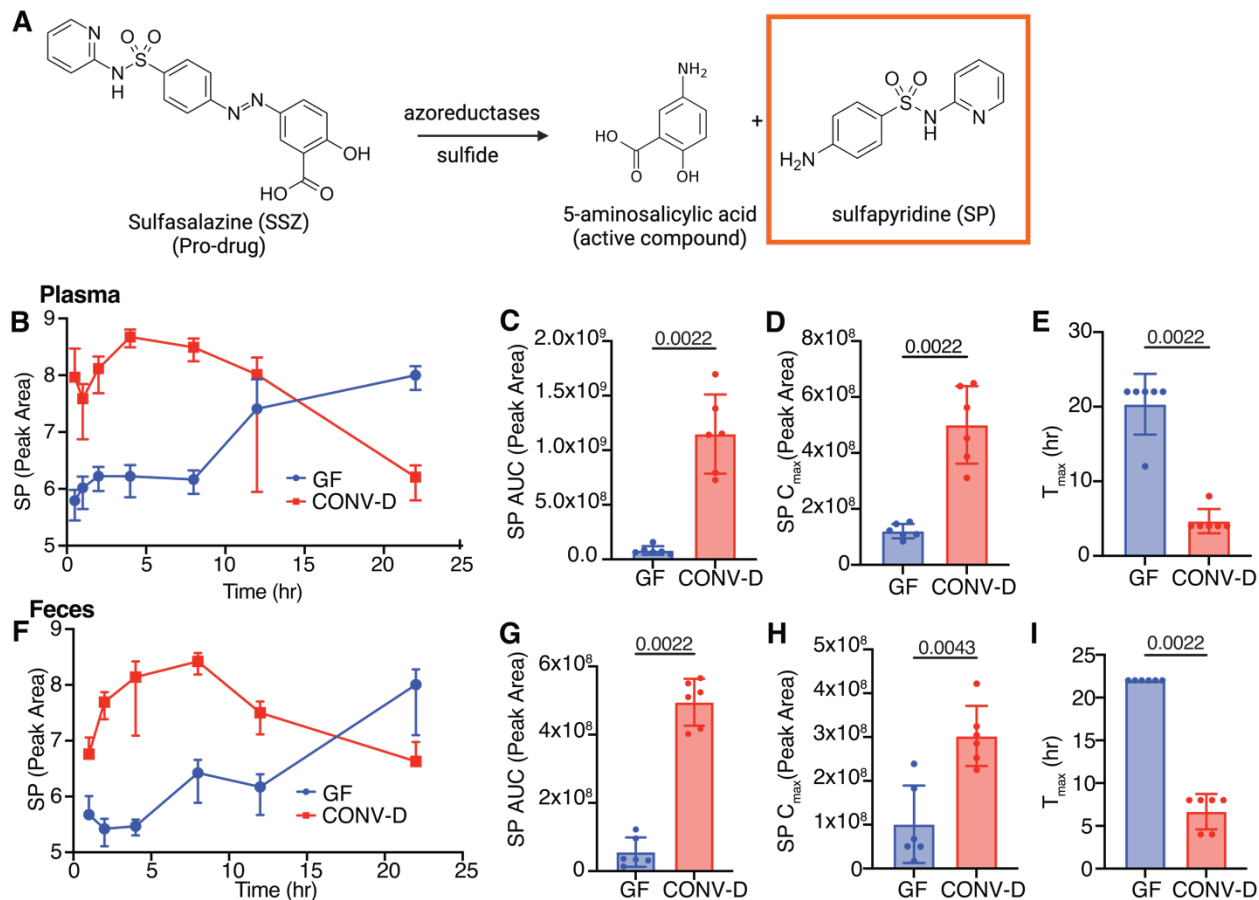


Figure 4.3. Gut microbiota presence impacts sulfasalazine metabolite, sulfapyridine pharmacokinetics. (b) SSZ activation reaction. The metabolite compound sulfapyridine, measured in the rest of the figure, is highlighted in orange. (b, f) LC-MS/MS peak area of plasma (c) and fecal (g) SP, plotted on a log scale. (c-e; g-i) PK parameters for plasma or fecal SP, respectively (c, g) area under the curve (AUC), (d, h) peak concentration (C_{max}), and (e, i) the time at which concentrations are highest (t_{max}). (n= 6 mice/group, values and bars represent mean +/- standard deviation, statistics done with Mann-Whitney t-test, fecal and plasma LC-MS/MS peak areas are normalized to background levels pre-SSZ dose).

To test the hypothesis that microbial sulfide production plays a role in SSZ activation, I also measured fecal sulfide over the course of this experiment to assess the dynamics of complex community introduction (**Figure 4.2a**). Upon colonization, there was an increase in fecal sulfide levels compared to pre-colonization and the GF group, with CONV-D levels stabilizing around day 10 (**Figure 4.4a**). The GF group consistently had low levels of fecal sulfide, often below the limit of detection (**Figure 4.4a**) using the Cline assay (9). After 14 days of colonization and SSZ PK, cecal sulfide was significantly higher in the CONV-D group compared to GF (**Figure 4.4b**).

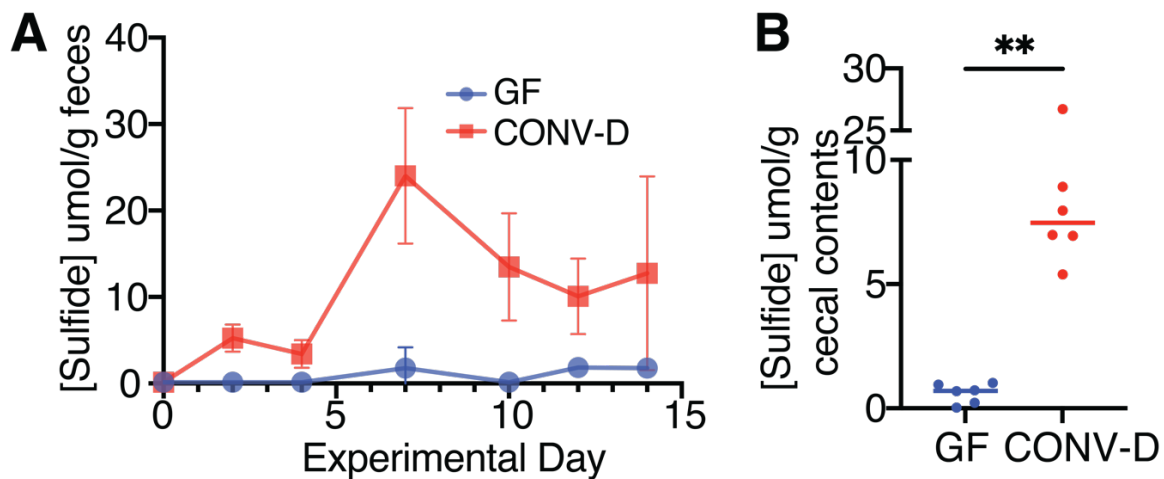


Figure 4.4. Conventionalization leads to increased sulfide levels. (a) Sulfide measured over the experiment using the Cline assay. (b) Cecal sulfide measured upon sacrifice after SSZ PK. (n= 6 mice/group, bars represent mean +/- standard deviation, statistics done with Mann-Whitney t-test, * $p < 0.05$).

Next, I asked if altering the gut microbial community with antibiotics would similarly impair SSZ activation. CONV-R mice were given streptomycin (strep) water or remained on tap-water. After one week of strep water treatment, mice were dosed with 200 mg/kg of SSZ for PK using fecal samples (n = 4 mice/group) (**Figure 4.5a**). I observed a faster depletion of SSZ from CONV-R stool by LC-MS/MS measurements (**Figure 4.5b**). This trend was also mirrored in the AUC and C_{max} parameters where the CONV-R had significantly lower AUC and C_{max} parameters (**Figure 4.5c-d**). The metabolite SP levels were less clear. In CONV-R mice there was a trend towards more SP earlier in the time course, but SP levels matched those of the strep treated group by 6 hrs (**Figure 4.5f**). Similarly, none of the PK parameters were significantly different between groups (**Figure 4.5 g-i**).

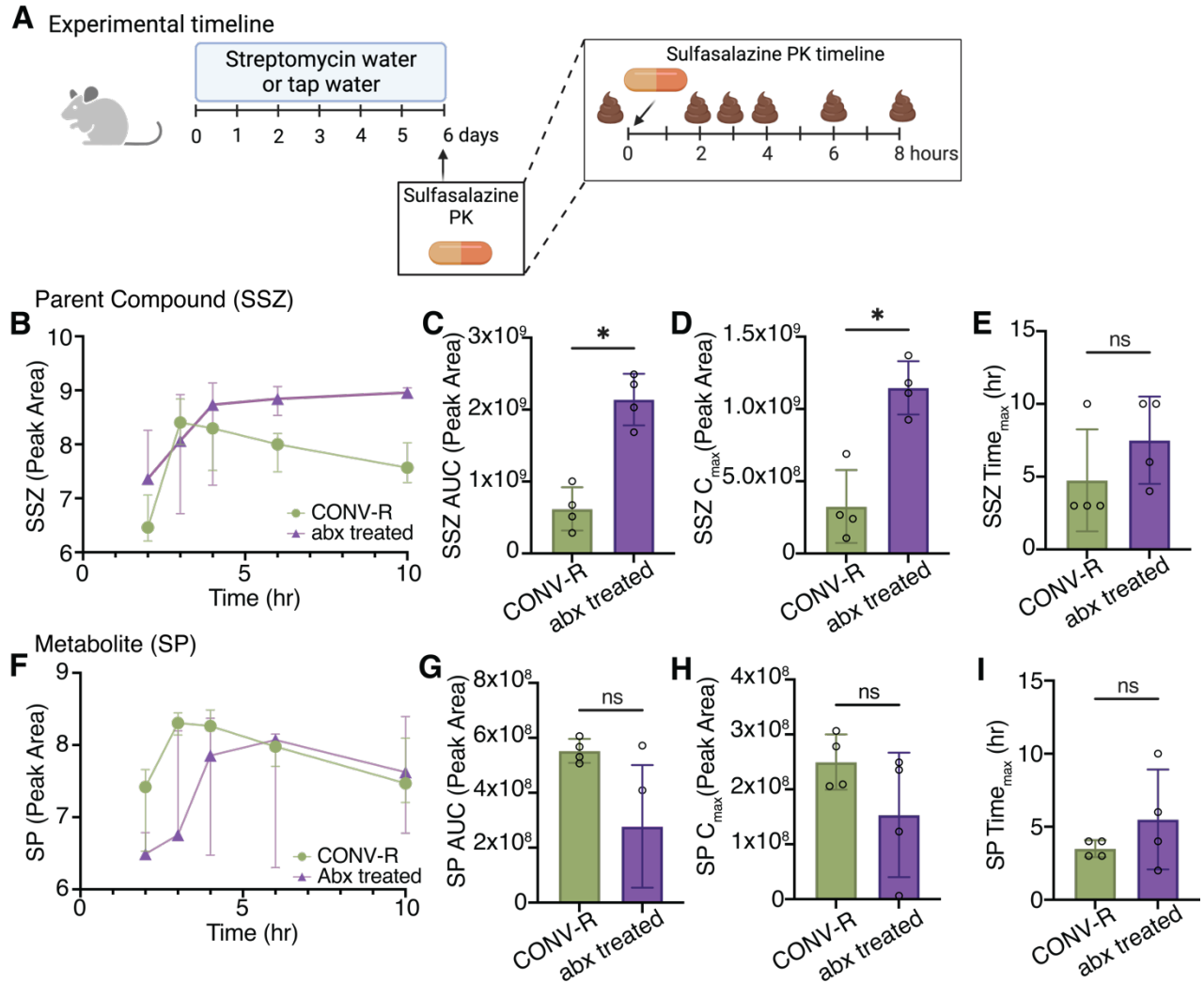


Figure 4.5. Streptomycin interferes with gut bacterial azo reduction. (a) Experimental schematic. Mice are given streptomycin water or remain on normal tap water. After 6 days, mice were dosed with 200 mg/kg SSZ and feces were collected for pharmacokinetics. (b, f) LC-MS/MS peak area of fecal SSZ (b) and SP (f) plotted on a log scale. (c-e, g-h) PK parameters for fecal SSZ or SP, respectively (c, g) area under the curve (AUC), (d, h) peak fecal concentration (C_{max}), and (e, i) the time at which concentrations are highest (t_{max}). (n= 4 mice/group, values and bars represent mean +/- standard deviation, statistics done with Mann-Whitney t-test, *p<0.05, fecal LC-MS/MS peak areas are normalized to background levels pre-SSZ dose).

4.2.2 Sulfasalazine pharmacokinetics are not altered by the addition of *E. coli* or dietary interventions.

Based on our previous work describing the depletion of azo compounds by microbially produced sulfide as a by-product of *in vitro*, I hypothesized that manipulating dietary L-Cysteine (L-Cys) may influence SSZ PK. In the same vein, my previous research found that knocking out the *fnr* regulator in *E. coli* led to no detectable sulfide *in vitro*. Thus, I also asked how adding *wt* or Δ *fnr* engraftment alters SSZ PK dynamics. I first validated that these *E. coli* findings held true for sulfasalazine *in vitro* using our LC-MS/MS method (Figure 4.6a-b). Then, I utilized the previously described strep model allowing for high biomass *E. coli* engraftment in strep-treated CONV-R mice with a minimal background microbiota (10). Mice engrafted with *wt E. coli* did not demonstrate different levels of fecal SSZ compared to strep-treated mice when dosed with 200 mg/kg SSZ (Figure 4.7a). We also did not see any differences in metabolite SP concentrations between groups (Figure 4.8a) (n = 4 mice/group).

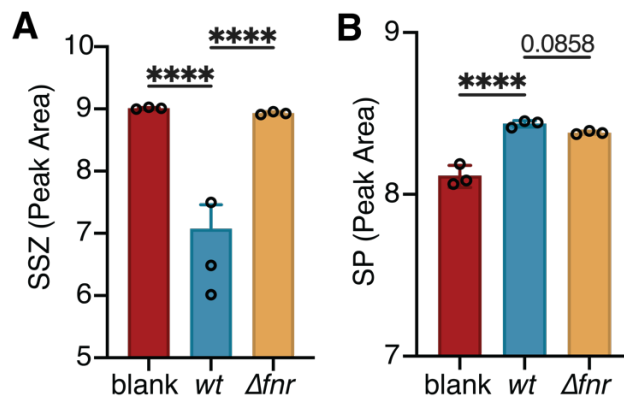


Figure 4.6. Sulfasalazine is significantly depleted in *wt E. coli* compared to Δ *fnr E. coli*. (a) Parent compound SSZ and (b) metabolite SP LC-MS/MS peak area of spend media supernatant after 24 hour incubation of 250 μ M Sulfasalazine with *wt* or Δ *fnr E. coli*. Values and bars represent mean +/- standard deviation, one-way ANOVA (n = 3 biological replicates)

While there were no observable differences, we hypothesized that changing mice to a high sulfur amino acid (SAA) diet would boost sulfide production and potentially SSZ activation. To control for sulfide production in engrafted *E. coli*, we also included a Δfnr *E. coli* group. However, we still did not see any differences in either plasma (**Figure 4.7b**) or fecal (**Figure 4.7c**) SSZ levels after dosing mice (n = 8 mice/group, SSZ dose 100 mg/kg). Similarly, we saw no significant differences in standard PK parameters (AUC of individuals, C_{max} , and T_{max}) (data not shown). Additionally, SP PK demonstrated no differences between groups in either plasma (**Figure 4.8b**) or fecal (**Figure 4.8c**) samples. Plasma and fecal PK Measurements of SSZ and SP showed no differences in measured levels over time by LC-MS/MS (**Figure 4.9d, h, l, p**). Similarly, none of the classic pharmacokinetic metrics demonstrated any significant differences (**Figure 4.9e-g, i-k, m-o, q-s**).

To further understand the impact of defined communities and known sulfide producer *E. coli* engraftment, we utilized the Altered Schaedler Flora (ASF) community (11). Others have used this 8-member augmented by *E. coli* genetically modified strains to address sulfide-related quotations (12). We hypothesized that a background microbiota would help drive competition for resources, likely pushing *E. coli* to utilize L-Cys metabolism. We colonized GF mice with the ASF community, then engrafted *wt* or Δfnr *E. coli* MG1655 after one week followed by a 100 mg/kg SSZ dose for SSZ and SP PK (n = 6-8 mice/group). I observed no differences between plasma SSZ levels between the ASF only, ASF + *wt E. coli*, or ASF + $\Delta fnr E. coli$ groups (**Figure 4.7d**). The same was true of the SP metabolite (**Figure 4.8d**).

Again, I reasoned that perhaps standard chow did not provide enough sulfur amino acids to drive a significant difference, so I repeated with the high SAA diet. Additionally, those that previously observed differences used a high SAA diet (12). Mice were colonized with ASF, then engrafted with *wt* or Δ *fnr* *E. coli* MG1655 and later dosed with 100 mg/kg SSZ for PK (n = 10-12 mice/group). There were no differences in SSZ concentrations between the ASF + *wt* and ASF + Δ *fnr* *E. coli* MG1655 groups for either plasma (**Figure 4.7e**) or fecal (**Figure 4.7f**) samples. Similarly, no differences were seen in SP levels between the groups in plasma (**Figure 4.8e**) or fecal (**Figure 4.8f**) samples.

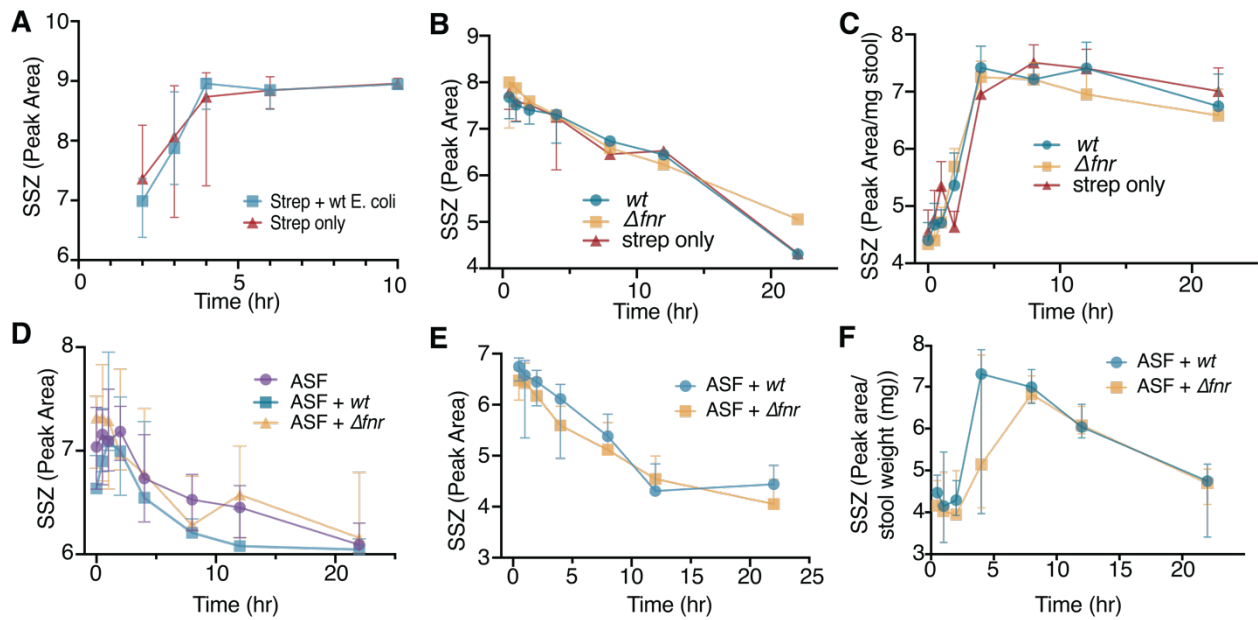


Figure 4.7. Combinations of *E. coli* engraftment and/or diet does not impact Sulfasalazine pharmacokinetics. (a) Parent compound SSZ LC-MS/MS peak area of plasma. Mice were put on streptomycin water then engrafted with *wt strep^R E. coli* MG1655. Mice were dosed with 200 mg/kg of SSZ. (n = 4 mice/group). (b-c) Mice were put on streptomycin water then engrafted with *wt strep^R* or *Δfnr::Kan^R strep^R E. coli* MG1655. Mice were dosed with 100 mg/kg of SSZ. SSZ LC-MS/MS peak area of plasma (b) and fecal (c) plotted on a log scale. (n = 8 mice/group). (d) SSZ LC-MS/MS peak area of plasma. GF mice were colonized with the ASF community, then engrafted with *wt* or *Δfnr::Kan^R E. coli* MG1655, or no *E. coli*. (n = 6-8 mice/group) (e-f) GF mice were colonized with the ASF community, then engrafted with *wt* or *Δfnr::Kan^R E. coli* MG1655. Mice were dosed with 100 mg/kg SSZ for PK. SSZ LC-MS/MS peak area of plasma (d) and feces (e) plotted on a log scale. (n = 10-12 mice/group)

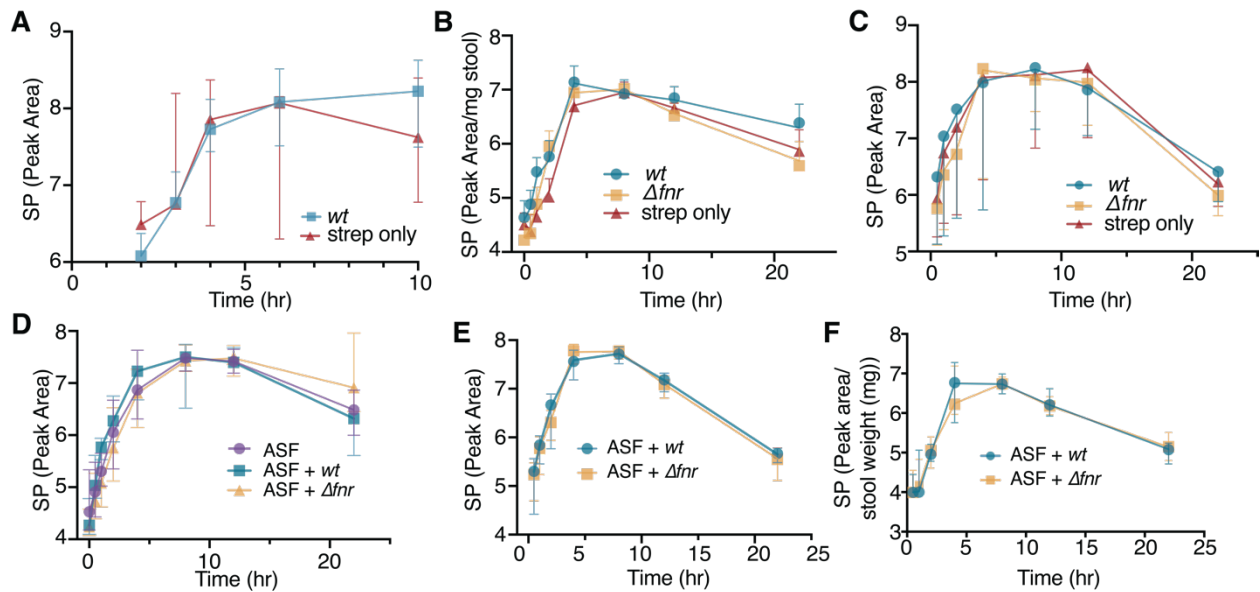


Figure 4.8. Combinations of *E. coli* engraftment and/or diet does not impact sulfapyridine pharmacokinetics. (a) Metabolite SP LC-MS/MS peak area of plasma. Mice were put on streptomycin water then engrafted with *wt strep^R* *E. coli* MG1655. Mice were dosed with 200 mg/kg of SSZ. (n = 4 mice/group). (b-c) Mice were put on streptomycin water then engrafted with *wt strep^R* or Δ *fnr::Kan^R strep^R* *E. coli* MG1655. Mice were dosed with 100 mg/kg of SSZ. SP LC-MS/MS peak area of plasma (b) and fecal (c) plotted on a log scale. (n = 8 mice/group). (d) Metabolite SP LC-MS/MS peak area of plasma. GF mice were colonized with the ASF community, then engrafted with *wt* or Δ *fnr::Kan^R* *E. coli* MG1655, or no *E. coli*. (n = 6-8 mice/group) (e-f) GF mice were colonized with the ASF community, then engrafted with *wt* or Δ *fnr::Kan^R* *E. coli* MG1655. Mice were dosed with 100 mg/kg SSZ for PK. SSZ LC-MS/MS peak area of plasma (e) and feces (f) plotted on a log scale. (n = 10-12 mice/group)

4.2.3 Gastrointestinal sulfide levels are robust to diet and *E. coli* strain engraftment

status. In addition to SSZ and SP PK dynamics, I was interested in how *E. coli* engraftment or diet impacts GI sulfide levels. First, I asked how L-Cys modulation changed sulfide levels. Mice were placed on defined, matched diets with high (0.35% L-Cys) or no L-Cys (0% L-Cys), L-methionine matched diets, similar to those used in previously published sulfide impact studies (6). There were no measurable differences between the two groups overtime (**Figure 4.9a**) or in cecal contents at the experimental endpoint (**Figure 4.9b**) as measured by the Cline reaction (9) (n = 7-8 mice/group). Importantly, L-Methionine (L-Met) can be converted to L-Cys, so I hypothesized that changing both amino acids might have an impact, as observed by other groups (12). Animals were placed on either a high (0.8% L-Cys and 1.5% L-Met) or low sulfur amino acid diet (0.03% L-Cys and 0.3% L-Met) for three weeks (n = 6 mice/group). There were no measurable differences between the two groups overtime (**Figure 4.9c**) or in cecal contents at the experimental endpoint as measured by the cline reaction (**Figure 4.9d**). Interestingly, there was a difference between fecal sulfide levels between standard chow and the defined diets in both experiments (pre-day 0 vs post-defined diet) (**Figure 4.9a, c**).

As there were observed stark differences between GF and CONV-D GI levels of sulfide, I asked how different colonization states impact sulfide levels. With treatment of streptomycin there was an immediate decrease in sulfide production that only recovered slightly after 8-9 days; this impact was consistent on both regular mouse chow (**Figure 4.9e**) and on high SAA diet (**Figure 4.9g**). Additionally, there were no differences between mice engrafted with *wt*, Δfnr *E. coli* MG1655, or no *E. coli*. This was observed in both

fecal pellets over time (**Figure 4.9e, g**) and cecal contents after 9 (**Figure 4.9f**) or 21 days (**Figure 4.9h-i**). Similar trends were observed in ASF + *E. coli* mice (**Figure 4.9j-l**).

Mice colonized with ASF alone, or ASF + *wt* or Δ *fnr* *E. coli* on standard chow demonstrated no differences between groups in cecal contents (**Figure 4.9j**). Similarly, ASF mice engrafted with either *wt* or Δ *fnr* *E. coli* MG1655 on high SAA diet had significant differences in cecal sulfide levels by the cline reaction (**Figure 4.9k**). While not significant, the sulfide production potential by the lead acetate method (13) did trend higher in the ASF + *wt* group (**Figure 4.9l**).

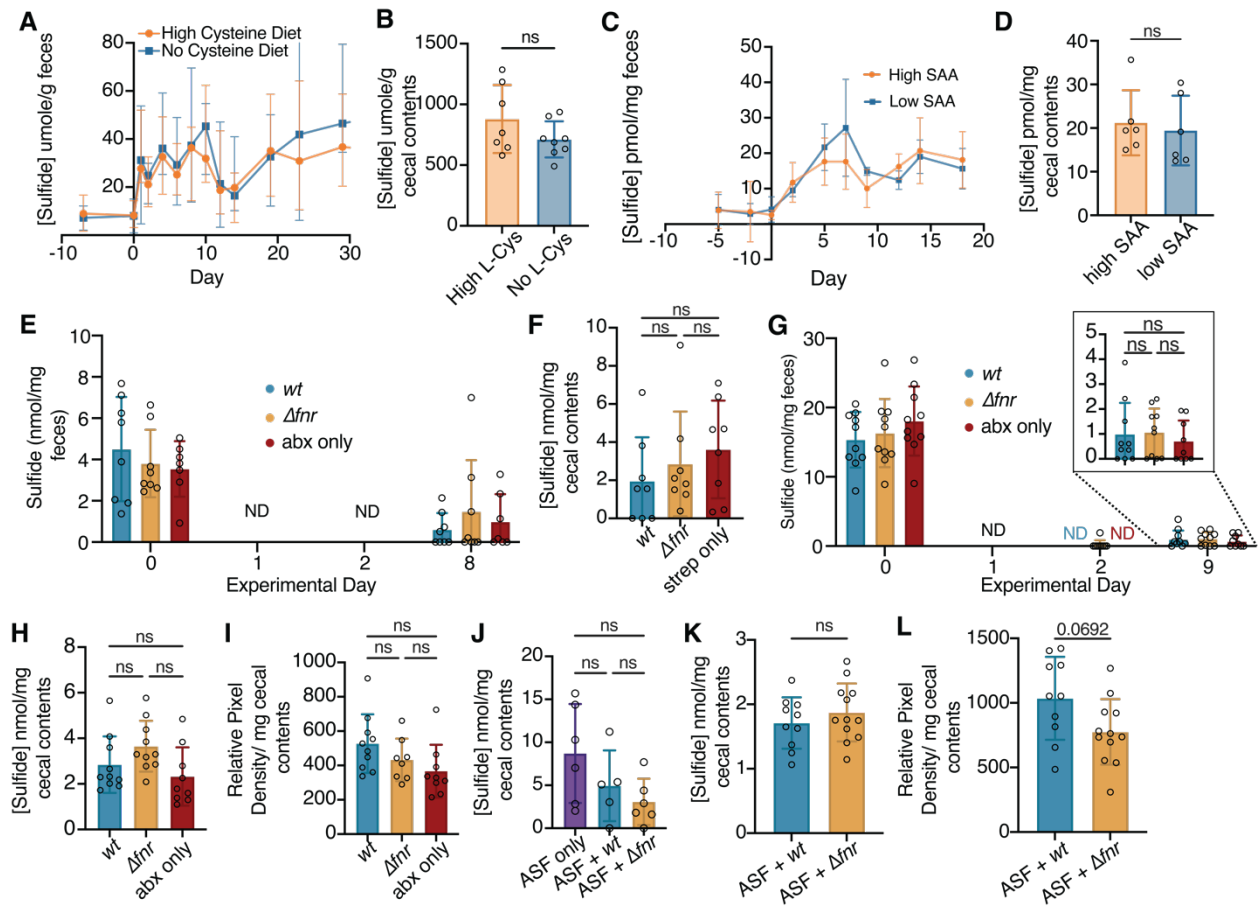


Figure 4.9. Murine GI sulfide levels are robust to dietary intervention. (a-b) Mice were given a high L-cysteine (0.35% L-Cys) or no cysteine diet (0% L-Cys) for 55 days. (a) Fecal sulfide over time measured by the Cline reaction (b) Cecal sulfide measured after 55 days of dietary intervention, measured by the Cline reaction. (n= 7-8 mice/group). (c-d) Mice were given a high (0.8% L-Cys and 1.5% L-Met) or low sulfur amino acid (0.03% L-Cys and 0.3% L-Met) diet for 21 days. (c) Fecal sulfide was measured before and after dietary intervention by the Cline reaction. (d) Cecal sulfide, after 21 days of dietary intervention, measured by the Cline reaction. (n= 6 mice/group). (e-f) All mice treated with strep water and given a high sulfur amino acid diet (0.8% L-Cys and 1.5% L-Met). Mice were engrafted with strep-resistant *wt E. coli*, $\Delta fnr::Kan^R E. coli$ MG1655, or left on antibiotics with no engraftment. (e) Fecal sulfide was measured throughout the engraftment timeline. Fecal sulfide measured from mice before and throughout antibiotic treatment and engraftment using the Cline reaction. (f) Sulfide levels from the cecal contents at the time of sacrifice measured by the Cline reaction. (n = 8 mice/group). (g-i) Mice are given streptomycin water and engrafted with strep-resistant *wt* or $\Delta fnr::Kan^R E. coli$ MG1655, or left on antibiotics with no engraftment. (g) Fecal sulfide was measured throughout the engraftment timeline using the Cline reaction. (h) Sulfide levels from the cecal contents at the time of sacrifice measured by the Cline reaction. (i) Sulfide production potential of the cecal contents. (Figure Caption continued on the next page)

(Figure caption continued from the previous page) Evolved sulfide from microbial reaction with L-Cysteine was trapped on lead acetate paper and sulfide was measured using pixel densitometry. (n = 10/group) (j) All mice were engrafted with the ASF microbiota and given a high sulfur amino acid diet (0.8% L-Cys and 1.5% L-Met). After 7 days mice were engrafted with *wt* or Δ *fnr* *E. coli*, or remained engrafted with only ASF. Cecal sulfide measured by the Cline reaction. (k-l) All mice were engrafted with the ASF microbiota and remained on standard chow. After 7 days mice were engrafted with *wt* or Δ *fnr* *E. coli*. (k) Cecal sulfide measured by the Cline reaction. (l) Cecal sulfide production potential. Evolved sulfide from microbial reaction with L-Cysteine was trapped on lead acetate paper and sulfide was measured using pixel densitometry. (n= 10-12 mice/group). (a-j) values and bars represent mean +/- standard deviation, statistics done with Mann-Whitney t-test or one-way ANOVA, Cline reaction limit of detection = 0.97uM).

4.3 Discussion

Overall, our results provide insight to the dynamics of SSZ activation and PK under a variety of conditions. Additionally, they provide a framework to build up for understanding the importance of microbiota composition and dietary impact on SSZ in addition to sulfide production.

The presence of a microbiome impacts SSZ activation and PK dynamics. While microbial activation of SSZ has long been accepted, here we provide the first comprehensive look at parent compound SSZ and metabolite SP PK over time in both CONV-D and GF mice. These results provide a base model to build upon for asking how other microbial communities impact the activation and efficacy of SSZ. As SSZ is frequently prescribed to patients with conditions where we know the microbiome can be different from healthy controls, or moreover, play role in disease (inflammatory bowel disease (1, 14, 15) and rheumatoid arthritis (16)), having a gnotobiotic model where we can directly test the mechanist impacts of these components will prove invaluable. Similarly, treatment with streptomycin impacted SSZ activation and PK metrics. Not only do these data provide complementary evidence that the microbiome plays a key role in SSZ activation and PK, they also highlight the importance of understanding how microbial interventions such as antibiotics can impact drug interactions without a clear obvious drug-drug interaction in the traditional sense.

Additionally, GI sulfide levels were microbiome status dependent. Both CONV-D and non-antibiotic treated mice produced more sulfide than their respective control groups. Based on previous research, sulfide is an alternative activator of azo-bonded compounds. Microbes are the main source of sulfide in the gut (17–20) and this

production is more wide-spread across microbial taxa than originally thought (21). Our experiments provide correlative data for microbial colonization status, sulfide levels, and SSZ activation. While of course this does not provide causative evidence, our results provide a baseline understanding of how these three components might interact and provide a model for future mechanistic experimentation.

E. coli engraftment status did not impact SSZ PK or sulfide levels. *E. coli* is a known sulfide producer (6, 22, 23) and azo reducing bacterium (6, 24). My previous work demonstrated that the *fnr* regulator controls the majority of *E. coli* sulfide production *in vitro*. However, wt *E. coli* engraftment was insufficient to induce PK changes or sulfide level differences *in vivo*. Additionally, there were no measurable differences between wt and *fnr* *E. coli* engraftment. These results may suggest that addition of one microbial species is insufficient to modulate SSZ dynamics; however, more work with additional *E. coli* strains is necessary to test the generalizability of these findings. These experiments provide important background information for future studies of potential probiotics or other uses to gut microbiome manipulation to improve patient SSZ activation and response.

Lastly, GI sulfide levels were robust to dietary intervention. A number of factors could account for this phenomenon. Specifically, there are likely redundant pathways that lead to L-Cys catabolism and metabolism to hydrogen sulfide. Additionally, for host health reasons, complete removal of dietary sulfur amino acids was unadvisable without causing off target host disease. Lastly, sulfide is volatile and easily interacts with many compounds, making accurate *in vivo* measurements of sulfide levels difficult and variable. Diet induced differences may have been masked by this volatility and simply unmeasured using our current methods.

There were multiple limitations to this study. Firstly, we focused on only one gene within *E. coli*, *fnr*, which is a regulator of many processes (25, 26). Because L-Cys metabolism is a redundant process in *E. coli*, knocking out a major contributor regulator (*cyuR*) or the L-Cys import pathway (*cyuP*) (22, 27) might reveal larger impacts *in vivo*. Similarly, we limited our studies to *E. coli* while there may be other gut species that might contribute more readily and relevantly to sulfide production. Furthermore, both of our low-diversity background models for adding in *E. coli* do not have known sulfide or azoreduction capacities. *In vitro* work with the individual ASF strains, in community and individually, could shed light on these contributions. Lastly, we did not survey the microbial composition and relative abundance and activity of each microbe in these studies. Doing so would help get a better understanding of which microbes are background sulfide levels, alone or in addition to added *E. coli* strains.

Despite these limitations, our results clearly emphasize that the presence of a microbial community impacts both SSZ PK and GI sulfide levels. We've provided a framework for continued investigation of these dynamics both in defined communities and CONV-R mice. While strong evidence exists for the impact of diet on the microbiome and the microbiome on drug metabolism, especially azo-bonded drugs like SSZ, the intersection of diet, microbiome, and drug metabolism is still an emerging field. Our work serves as an excellent scaffolding for continued research on this question and optimization of therapeutic strategies for pharmaceuticals that rely on microbial activation.

4.4 Materials and Methods

Sulfasalazine pharmacokinetics. All animal experiments were approved by the University of California, San Francisco (UCSF) IACUC. Sample sizes were chosen based on pilot experiments to have 95% power at a significance level of 0.01 using an unpaired t-test (G*Power3.1). Mice were randomized based on age, the cage of origin, and weight to experimental groups. Pharmacokinetic studies were not blinded. Mice were housed at 67-74°F, 30-70% humidity, and 12/12-hr light/dark cycle. Unless otherwise noted, mice were dosed USP grade SSZ (Sigma) that was freshly prepared in 0.1M NaOH in PBS. SSZ was delivered by oral gavage (0.1 ml) at a dose of 100 mg/kg. Blood samples (~20 µl per time point) were obtained from the mouse tail vein at 0, 0.5, 1, 2, 4, 8, 12, and 22 hours post-SSZ administration. Blood samples were collected in Fisherbrand heparinized glass microhematocrit capillary tubes (Waltham, MA) and centrifuged at 3500g for 15 min to recover blood plasma, which was stored at -80°C until analysis.

Experiments with GF mouse conventionalization. C57BL/6J mice (females and males, ages 6-10 weeks) were obtained from the University of California, San Francisco (UCSF) Gnotobiotics core facility (gnotobiotics.ucsf.edu) and housed in gnotobiotic isolators for the duration of each experiment (Class Biologically Clean) or were housed in Iso positive cages (Tecniplast). Mice were colonized via oral gavage with the cecal contents of a female C57BL/6J mouse from the Turnbaugh Lab colony.

Experiments with *E. coli* engraftment in CONV-R mice given streptomycin. C57BL/6J mice (females or males, ages 6-10 weeks) were ordered from Jackson Labs. Streptomycin water was prepared by dissolving USP grade streptomycin sulfate (VWR 0382) in autoclaved tap water to a final concentration of 5 g/ml and passing through 0.22

µm filtration units. Mice were provided streptomycin water for 1 day, followed by oral gavage of 0.1 mL streptomycin-resistant *E. coli* (*wt* or Δ *fnr::Kan^R*), or remained only on streptomycin water. Mice were kept on streptomycin water thereafter to maintain colonization. *E. coli* colonization was confirmed using Colony Forming Units (CFU) counting at endpoint.

Experiments with ASF and *E. coli* engraftment in GF mice. C57BL/6J mice (females and males, ages 6-10 weeks) were obtained from the University of California, San Francisco (UCSF) Gnotobiotics core facility (gnotobiotics.ucsf.edu) and housed in gnotobiotic isolators for the duration of each experiment (Class Biologically Clean) or were housed in Iso positive cages (Tecniplast). Mice were colonized via oral gavage with the cecal contents of an Altered Schadler Flora (ASF) colonized mouse from Taconic. After one week, mice were engrafted with *E. coli* MG1655 (*wt* or Δ *fnr::Kan^R*). *E. coli* colonization was confirmed using CFU counting at endpoint.

Diets. Custom diets with 0.35% L-Cys or no L-Cys, L-methionine matched, diets were purchased from Teklad. Diets with high (0.8% L-Cys and 1.5% L-Met) or low sulfur amino acid diet (0.03% L-Cys and 0.3% L-Met) were purchased from Reserach Diets, Inc (Cat No. A15121502N, A15121501N, respectively, formulated with no dye. Otherwise, a chow diet (Lab Diet 5058) was used for SPF mice and an autoclaved chow diet (Lab Diet 5021) was used for the gnotobiotic mice. Diets used for gnotobiotic experiments were either autoclaved or irradiated and vacuum sealed to ensure sterility.

Plasma extraction for LC-MS/MS analysis. 5 µl of plasma was extracted with 50 µl of sterile Mili-Q water and 180 µl of ethyl acetate. This mixture was vortexed 5 mins and rested on ice for 5 min 2 times. Then samples were centrifuged at max speed for 30 mins

at 4°C. The ethyl acetate layer was moved to a new 1.5ml tube and dried using a SpeedVac Vacuum until dry. Samples were then resuspended in 150 µl of 100% MeOH and stored at -20°C until analyzed.

Fecal sample extraction for LC-MS/MS analysis. Fecal samples were weighed (averaging approximately 25 mg) and used for extraction and moved to Precellys Bead Tubes (Cat. No. P000911-LYSK0-4). 150 µl of sterile Mili-Q water and 350 µl of ethyl acetate were added to pellets. Samples were homogenized using the Cryolys Evolution tissue homogenizer (Berlin Technologies) with default hard tissue settings. Then samples were centrifuged at max speed for 30 mins at 4°C. The ethyl acetate layer was moved to a new 1.5ml tube and dried using a SpeedVac Vacuum until dry. Samples were then resuspended in 500 µl of 100% MeOH and stored at -20°C until analyzed.

LC-MS/MS analysis.

Sulfasalazine (SSZ) and sulfapyridine (SP) from conditioned media samples, mouse plasma, and mouse fecal samples were analyzed by liquid chromatography-triple quadrupole mass spectrometry (LC-MS/MS) using a SCIEX Triple Quad 7500 instrument with a linear ion QTRAP (Redwood City, CA). Chromatographic separation was achieved using a Synergi 4 µm Fusion-RP 80 A column (50 x 2 mm) at 45°C. The mobile phase consisted of methanol + 0.1% formic acid for solvent A and HPLC-grade water + 0.1% formic acid as solvent B. A flow rate of 0.4 ml/min was used at the following gradient elution profile: 40% B at 0-2 min, gradient to 90% B from 2-5.5 min, gradient to 40% B at 5.6-9 min. The autosampler was maintained at an internal temperature of 4°C. The eluate from the column was ionized in the LC-MS/MS using an electrospray ionization source

(ESI) in positive polarity. Data analysis was performed using the built-in SCIEX OS Software.

Sulfide measurements from fecal samples.

Cline assay. Fecal pellets were collected and immediately placed in 2% zinc acetate solution on ice in the mouse facility. Cecal contents were immediately added to 2% zinc acetate solution on ice upon sac. After collection was completed, samples were homogenized in the zinc acetate solution using a wooden stick or cut, wide-bore pipette tip. Then 300 μ l of the DPD/FeCl₃ Cline Reaction Reagent was added. A few fecal pellets or cecal contents were reserved in zinc acetate without the addition of the Cline Reaction Reagent (equal percentage of HCl solution was used instead) was reserved as a background control. The reaction was incubated at 37°C for 20min, centrifuged for 4min at max speed and absorbance was measured in OD 670nm.

Lead Acetate. Lead acetate paper was made as previously described (13), briefly, pre-cut Whatman paper was incubated in 20mM lead acetate solution for 20 min at room temperature and then was dried for 20 min at 110°C and kept in the dark until use. Cecal contents were added to PBS solution with 10 mM L-Cysteine (VENDOR) and 1 mM pyridoxal-5-phosphate (Sigma Milipore), then homogenized with a pipette tip to mix. 200 μ l of the slurry was moved to 96-well plates in triplicate. The plate was tightly covered with lead acetate paper and incubated at 37°C in an anaerobic chamber for 2 hrs. The lead acetate papers were then photographed, and densitometry analysis was performed using ImageJ2 (28).

References

1. Peppercorn MA. 1984. Sulfasalazine: Pharmacology, Clinical Use, Toxicity, and Related New Drug Development. *Ann Intern Med* 101:377.
2. Sousa T, Yadav V, Zann V, Borde A, Abrahamsson B, Basit AW. 2014. On the Colonic Bacterial Metabolism of Azo-Bonded Prodrugsof 5-Aminosalicylic Acid. *J Pharm Sci* 103:3171–3175.
3. Spanogiannopoulos P, Bess EN, Carmody RN, Turnbaugh PJ. 2016. The microbial pharmacists within us: a metagenomic view of xenobiotic metabolism. *Nat Rev Microbiol* 14:273–287.
4. Jain A, Gupta Y, Jain SK. 2006. Azo chemistry and its potential for colonic delivery. *Crit Rev Ther Drug Carrier Syst* 23:349–400.
5. Zimmermann M, Zimmermann-Kogadeeva M, Wegmann R, Goodman AL. 2019. Mapping human microbiome drug metabolism by gut bacteria and their genes. *Nature* <https://doi.org/10.1038/s41586-019-1291-3>.
6. Wolfson SJ, Hitchings R, Peregrina K, Cohen Z, Khan S, Yilmaz T, Malena M, Goluch ED, Augenlicht L, Kelly L. 2022. Bacterial hydrogen sulfide drives cryptic redox chemistry in gut microbial communities. *Nat Metab* 4:1260–1270.
7. Pieper LM, Spanogiannopoulos P, Volk RF, Miller CJ, Wright AT, Turnbaugh PJ. 2022. The global anaerobic metabolism regulator *fnr* is necessary for the degradation of food dyes and drugs by *Escherichia coli*. *bioRxiv*.

8. Carmody RN, Gerber GK, Luevano JM, Gatti DM, Somes L, Svenson KL, Turnbaugh PJ. 2015. Diet Dominates Host Genotype in Shaping the Murine Gut Microbiota. *Cell Host Microbe* 17:72–84.
9. Cline JD. 1969. Spectrophotometric determination of hydrogen sulfide in natural waters¹. *Limnol Oceanogr* 14:454–458.
10. Myhal ML, Laux DC, Cohen PS. 1982. Relative colonizing abilities of human fecal and K 12 strains of *Escherichia coli* in the large intestines of streptomycin-treated mice. *Eur J Clin Microbiol* 1:186–192.
11. Orcutt RP, Gianni FJ, Judge RJ. 1987. Development of an “altered Schaedler flora” for NCI gnotobiotic rodents. *Microecol Ther.*
12. Lobel L, Cao YG, Fenn K, Glickman JN, Garrett WS. 2020. Diet posttranslationally modifies the mouse gut microbial proteome to modulate renal function. *Science* 369:1518–1524.
13. Hine C, Mitchell JR. 2017. Endpoint or Kinetic Measurement of Hydrogen Sulfide Production Capacity in Tissue Extracts. *Bio Protoc* 7.
14. Das KM. 1989. Sulfasalazine therapy in inflammatory bowel disease. *Gastroenterol Clin North Am* 18:1–20.
15. Sutherland LR, May GR, Shaffer EA. 1993. Sulfasalazine revisited: a meta-analysis of 5-aminosalicylic acid in the treatment of ulcerative colitis. *Ann Intern Med* 118:540–549.

16. Plosker GL, Croom KF. 2005. Sulfasalazine: a review of its use in the management of rheumatoid arthritis. *Drugs* 65:1825–1849.
17. Blachier F, Beaumont M, Kim E. 2019. Cysteine-derived hydrogen sulfide and gut health: a matter of endogenous or bacterial origin. *Curr Opin Clin Nutr Metab Care* 22:68–75.
18. Barton LL, Ritz NL, Fauque GD, Lin HC. 2017. Sulfur Cycling and the Intestinal Microbiome. *Dig Dis Sci* 62:2241–2257.
19. Hanson BT, Dimitri Kits K, Löffler J, Burrichter AG, Fiedler A, Denger K, Frommeyer B, Herbold CW, Rattei T, Karcher N, Segata N, Schleheck D, Loy A. 2021. Sulfoquinovose is a select nutrient of prominent bacteria and a source of hydrogen sulfide in the human gut. *ISME J* <https://doi.org/10.1038/s41396-021-00968-0>.
20. Flannigan KL, McCoy KD, Wallace JL. 2011. Eukaryotic and prokaryotic contributions to colonic hydrogen sulfide synthesis. *Am J Physiol Gastrointest Liver Physiol* 301:G188–93.
21. Braccia DJ, Jiang X, Pop M, Hall AB. 2021. The Capacity to Produce Hydrogen Sulfide (H₂S) via Cysteine Degradation Is Ubiquitous in the Human Gut Microbiome. *Front Microbiol* 12:705583.
22. Zhou Y, Imlay JA. 2022. *Escherichia coli* Uses a Dedicated Importer and Desulfidase To Ferment Cysteine. *MBio* e0296521.

23. Li K, Xin Y, Xuan G, Zhao R, Liu H, Xia Y, Xun L. 2019. *Escherichia coli* Uses Separate Enzymes to Produce H₂S and Reactive Sulfane Sulfur From L-cysteine. *Front Microbiol* 10:298.
24. Nakanishi M, Yatome C, Ishida N, Kitade Y. 2001. Putative ACP phosphodiesterase gene (*acpD*) encodes an azoreductase. *J Biol Chem* 276:46394–46399.
25. Uden G, Schirawski J. 1997. The oxygen-responsive transcriptional regulator FNR of *Escherichia coli*: the search for signals and reactions. *Mol Microbiol* 25:205–210.
26. Spiro S, Guest JR. 1990. FNR and its role in oxygen-regulated gene expression in *Escherichia coli*. *FEMS Microbiol Lett* 75:399–428.
27. Zhou Y, Imlay JA. 2020. *Escherichia coli* K-12 Lacks a High-Affinity Assimilatory Cysteine Importer. *MBio* 11.
28. Rueden CT, Schindelin J, Hiner MC, DeZonia BE, Walter AE, Arena ET, Eliceiri KW. 2017. ImageJ2: ImageJ for the next generation of scientific image data. *BMC Bioinformatics* 18:529.

Chapter 5: Conclusion

1.1 Conclusion

My thesis highlights the critical role of environmental factors in governing a clinically relevant metabolic activity of the human gut microbiome. The findings presented here are distinctly pertinent given the escalating intake of food dyes in processed foods (1–3) and emerging evidence of their unintended effects on gut epithelial (4) and immune cells (5). Understanding the mechanisms responsible for azo dye metabolism in *E. coli* and other gut bacterial species is crucial for predicting and controlling the location and extent of this metabolic activity.

Importantly, our research emphasizes the need to consider the chemical environment in which drugs are administered and reveals how the human gut microbiome mitigates the impact of chemical exposures. The translational implications of these findings are evident, providing a foundation for rational selection of excipients to minimize their off-target effects on human and microbial cells.

Furthermore, our results demonstrate the profound influence of the gut microbiota on sulfasalazine (SSZ) pharmacokinetics and gastrointestinal sulfide levels. By establishing a framework for continued investigation of these dynamics in defined communities and CONV-R mice, we contribute to the understanding of the complex interplay between diet, the microbiome, and drug metabolism—particularly in the context of azo-bonded drugs like SSZ. As this area of study continues to evolve, our work serves as a solid groundwork for future research on optimizing therapeutic strategies for pharmaceuticals that rely on microbial activation.

In conclusion, my thesis underscores the importance of considering the intricate relationships between the human gut microbiome, xenobiotic intake, and drug metabolism and absorption. The findings presented here not only expand our understanding of the gut microbiome's vital role in human health but also offer promising avenues for further research and intervention in the microbiome field.

5.2 References:

1. Center for Food Safety, Nutrition A. Color Additives History. US Food and Drug Administration. FDA. <https://www.fda.gov/industry/color-additives/color-additives-history>. Retrieved 24 June 2022.
2. Arnold LE, Lofthouse N, Hurt E. 2012. Artificial food colors and attention-deficit/hyperactivity symptoms: conclusions to dye for. *Neurotherapeutics* 9:599–609.
3. Lehmkuhler A, Miller MD, Bradman A, Castorina R, Chen M-A, Xie T, Mitchell AE. 2022. Levels of FD&C certified food dyes in foods commonly consumed by children. *J Food Compost Anal* 112:104649.
4. Zou L, Spanogiannopoulos P, Pieper LM, Chien H-C, Cai W, Khuri N, Pottel J, Vora B, Ni Z, Tsakalozou E, Zhang W, Shoichet BK, Giacomini KM, Turnbaugh PJ. 2020. Bacterial metabolism rescues the inhibition of intestinal drug absorption by food and drug additives. *Proc Natl Acad Sci U S A* 117:16009–16018.
5. He Z, Chen L, Catalan-Dibene J, Bongers G, Faith JJ, Suebsuwong C, DeVita RJ, Shen Z, Fox JG, Lafaille JJ, Furtado GC, Lira SA. 2021. Food colorants metabolized by commensal bacteria promote colitis in mice with dysregulated expression of interleukin-23. *Cell Metab* 33:1358–1371.e5.

Publishing Agreement

It is the policy of the University to encourage open access and broad distribution of all theses, dissertations, and manuscripts. The Graduate Division will facilitate the distribution of UCSF theses, dissertations, and manuscripts to the UCSF Library for open access and distribution. UCSF will make such theses, dissertations, and manuscripts accessible to the public and will take reasonable steps to preserve these works in perpetuity.

I hereby grant the non-exclusive, perpetual right to The Regents of the University of California to reproduce, publicly display, distribute, preserve, and publish copies of my thesis, dissertation, or manuscript in any form or media, now existing or later derived, including access online for teaching, research, and public service purposes.

DocuSigned by:

C0CAA68F8B7D42F... Author Signature

8/28/2023
Date