Characterization of a widely distributed cardiac drug-inactivating enzyme from the human gut bacterium Eggerthella lenta

Characterization of a widely distributed cardiac drug-inactivating enzyme from the human gut bacterium *Eggerthella lenta*

A dissertation presented

by

Nitzan Koppel

to

the Committee on Higher Degrees in Chemical Biology

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Chemical Biology

Harvard University

Cambridge, Massachusetts

January 2018

Dissertation advisor: Emily P. Balskus Nitzan Koppel

Characterization of a widely distributed cardiac drug-inactivating enzyme from the human gut bacterium *Eggerthella lenta*

Abstract

The human body is colonized by trillions of microorganisms that are increasingly implicated in modulating human health and disease. In particular, the human gut microbiota is involved in the metabolism of over fifty pharmaceuticals, yielding metabolites with altered biological properties and toxicities. It has been known for decades that particular isolates of the human gut bacterium *Eggerthella lenta* transform the plant-derived toxin and cardiac drug digoxin into the inactive metabolite (20*R*)-dihydrodigoxin, leading to decreased efficacy in a considerable subset of patients. Recently, the Turnbaugh lab identified the cardiac glycoside reductase (*cgr*) operon, a digoxin-inducible gene cluster that was predicted to be responsible for digoxin metabolism. In this thesis, we sought to expand our mechanistic understanding of this clinically relevant transformation and investigate its broader implications for gut microbial and human health.

Through heterologous expression and *in vitro* biochemical characterization, we discovered that the *E. lenta* enzyme Cgr2 is sufficient for digoxin reduction and inactivation. Having validated the *cgr2* gene as a biomarker for digoxin reduction, we probed the distribution of this metabolism among *E. lenta* strains and in the general human population. Using culturing and sequencing analyses, we identified seven additional digoxin-metabolizing strains of *E. lenta* with remarkable sequence conservation of the *cgr2* gene (>98% sequence identity). Metagenomic and qRT-PCR analyses confirmed the high sequence conservation of the *cgr2* gene and revealed that *cgr2*$^+$ *E. lenta* are widespread, but often low in abundance in the human gut microbiota.

We next probed the biochemical mechanism of digoxin reduction by the prevalent Cgr2 enzyme. Using a combination of biochemical, bioinformatic, and spectroscopic techniques, we determined that Cgr2 is a unique reductase that requires an FAD cofactor, harbors oxygen-sensitive, redox-active [4Fe-4S] clusters, and may contain a divalent metal cation center. Although the presence of [4Fe-4S] clusters in Cgr2 was unexpected, these metalloclusters proved to be essential for Cgr2 stability and likely serve a catalytic, electron transfer role. We further identified six cysteine residues that are important for Cgr2 activity and may influence metallocofactor binding.

We next explored the evolutionary origins and impacts of digoxin metabolism on *E. lenta*. Despite the high sequence conservation of the *cgr* operon, no obvious physiological benefit (e.g. growth advantage) could be linked to this metabolism. We thus investigated whether digoxin is the endogenous substrate of Cgr2 by assessing the activity of this enzyme toward a panel of alternative candidate substrates that may be relevant in the gut. However, Cgr2 metabolism appears to be restricted to digoxin and highly similar cardenolide toxins produced by plants. We thus propose that the *cgr* pathway may have evolved to protect humans from ingested toxins in an analogous manner to host xenobiotic-detoxifying enzymes. By maintaining host health, this pathway could thus provide a habitat for *E. lenta* colonization.

The studies presented in this thesis shed light on a long-standing clinical problem, and provide a roadmap for the discovery and mechanistic characterization of additional gut microbial-xenobiotic interactions of medical importance. Such knowledge may allow for accurate prediction of xenobiotic metabolism in the clinic and enable the development of therapeutic interventions to modulate these activities and ultimately affect human health.

Table of Contents

Abstract ..... **iii**

Table of Contents ..... **v**

Acknowledgements ..... **ix**

List of Figures ..... **xi**

List of Tables ..... **xiv**

List of Abbreviations ..... **xv**

Chapter 1: The human gut microbiota influences drug efficacy **1**

1.1 Gut microbial involvement in human health 2

1.2 Contrasting metabolism of xenobiotics by gut microbial and human enzymes 6

1.2.1 Host enzymes 8

1.2.2 Hydrolytic enzymes 8

1.2.2 Lyases 10

1.2.3 Reductases 11

1.2.4 Transferase enzymes 13

1.2.5 Radical enzymes 13

1.3 Gut microbial metabolism of pharmaceuticals 15

1.3.1 Anti-inflammatory and gastrointestinal agents 15

1.3.2 Cancer chemotherapeutics 18

1.3.3 Biologics and peptide-based drugs 20

1.3.4 Central nervous system (CNS) drugs 20

1.3.5 Anti-viral agents 22

1.3.6 Herbal supplements and traditional medicines 24

1.3.7 Cardiovascular drugs 25

1.4 Gut microbial inactivation of the cardiac medication digoxin 27

1.4.1 Pharmacology of the foxglove plant and digoxin 27
1.4.2 Mechanism of action of digoxin 29
1.4.3 Inactivation of digoxin by the prominent gut bacterium *Eggerthella lenta* 33
1.5 References 37

Chapter 2: Cgr2 is a unique enzyme encoded in the *cgr* operon that is responsible for digoxin metabolism 47

2.1 Bioinformatic analyses and heterologous expression link the *cgr* operon to digoxin reduction 48
2.2 Cgr2 is a divergent flavin dependent reductase 52
2.3 An expanded *cgr* operon is present in digoxin metabolizing *E. lenta* isolates 59
2.4 *E. lenta* and *cgr2* are prevalent in human gut microbiomes 63
2.5 Discussion 66
2.6 Experimental 68
2.7 References 75

Chapter 3: Biochemical characterization of Cgr2, a [4Fe-4S] and FAD-dependent cardiac glycoside reductase. 78

3.1. Heterologous expression and activity of Cgr1 and Cgr2 towards digoxin 79
3.2. Cgr2 requires [Fe-S] cluster reconstitution and FAD for *in vitro* stability and activity 83
3.3. Characterization of [Fe-S] clusters in Cgr2 87
3.3.1 UV-vis and EPR spectroscopy reveal the presence of [4Fe-4S] cluster(s) in Cgr2 87
3.3.2 Determining the metal dependence and [4Fe-4S] cluster stoichiometry of Cgr2 88
3.4. Identification of cysteine residues important for Cgr2 activity 92
3.5. FAD dependence of Cgr2 95
3.6. NMR characterization confirms production of (20*R)*-dihydrodigoxin by Cgr2 97
3.7. Discussion 103

3.8. Experimental 106

3.9. References 116

Chapter 4: Investigations of the substrate specificity and biological role of Cgr2 119

4.1. Probing the physiological implications of digoxin metabolism 120

4.1.1 Dietary protein influences *E. lenta* growth and digoxin metabolism 120

4.1.2 Digoxin reduction does not influence *E. lenta* growth 123

4.2. Kinetics of Cgr2 towards digoxin 126

4.3. Identifying potential substrates of Cgr2 132

4.3.1 Cardiac glycosides 132

4.3.2 Furanones 134

4.3.3 α,β-unsaturated carboxylic acids 135

4.3.4 Ketosteroids 136

4.3.5 Prostaglandins 137

4.3.6 *In vitro* activity of Cgr2 towards the substrate panel 137

4.4. Discussion 141

4.5. Experimental 146

4.6. References 151

Chapter 5: Future directions and concluding remarks 156

5.1. Conclusions 157

5.2. Further mechanistic studies of Cgr proteins 159

5.2.1 Metallocofactor characterization 159

5.2.2 Structural characterization of Cgr2 161

5.2.3 Reconstitution of the full Cgr complex 163

5.3. Additional factors that may influence the clinical efficacy of digoxin 164

5.4. Examining the broader relevance of gut microbes to pharmaceutical metabolism 166

5.4.1 Strategies and considerations for discovering new gut microbial-xenobiotic interactions166

5.4.2 Moving from clinical observation to biomarkers of drug metabolism 169

5.4.3 Incorporating gut microbial metabolism into drug design and development 169

5.5. Experimental 173

5.6. References 174

Acknowledgements

I would first like to thank my advisor Emily Balskus, for welcoming me into her lab and for allowing me to work on this project during such an exciting and pivotal time in microbiome research. I am greatly appreciative of Emily's mentorship throughout these last five years, for her thorough and critical approach to research, and for encouraging me to tackle meaningful and challenging scientific problems.

I also have to acknowledge the many incredible mentors who have trained, inspired, and encouraged me throughout my career. I'm so appreciative of Michael Marletta and Emily Weinert for providing an incredibly supportive and exciting introduction to research. Nick Toriello, Robert Blazej, Charlie Emrich, and Rich Cohen, thank you for the amazing start-up experience, an invaluable perspective on industry science, and your continued friendship and advice. I also want to thank Dan Kahne, David Liu, and Michael Gilmore who have served on my committee for the last five years and provided constructive feedback and helped me think about the broader implications of my work.

I have had many incredible collaborators over the years who have contributed immensely to this project. In particular, I'm very grateful to Jordan Bisanz, Henry Haiser, and Peter Turnbaugh for their technical expertise on *E. lenta* isolation, genomics and bioinformatics, illuminating scientific discussions, and especially to Peter for his infectious enthusiasm for science and for inspiring my PhD project. I am also indebted to Oren Rosenberg, Manuel Ortega, Cathy Drennan, Jack Nicoludis, and Rachelle Gaudet for periodically hosting me in their labs and sharing their vast knowledge of biochemistry and crystallography. Finally, thank you to Maria-Eirini Pandelia for her tremendous generosity and assistance with the EPR experiments, and for serving on my dissertation committee.

I also want to thank Jason Millberg, Dan Kahne, and Suzanne Walker for their dedication and support of the ChemBio program and its students, and for the occasional karaoke duet at the annual retreat.

To all of Balskus lab members, past and present, you have been wonderful colleagues, and truly inspiring scientists. I have learned so much from all of you, and I am so grateful for your willingness to share your time, knowledge, enthusiasm for science, and snacks.

To the incredible friends that I have made in and outside of lab –Jenn Brophy, Sarah Luppino, Alexandra Cantley, Stephen Wallace, Matt Wilson, Erica Schultz, Doug Kenny, Yolanda Huang, Nathaniel Braffman, Ana Martinez del Campo, Carolyn Brotherton, and especially to Abraham Waldman, Jack Nicoludis, Tamar Sella, Pia Faxon, Ayala and Adi Snir, and of course Goni – thank you for the over-the-top meals and food-themed parties, weekend trips, crossword puzzles, for playing/hosting Porchfests, for indulging my ceramics obsessions and hyperboles, for the endless humor, and most importantly for your empathy and patience during the challenging parts of grad school.

And finally, to my family, Yanki, Sol, Ariel, Danielle, and Inbar, an enormous thank you for sharing your compassion and encouragement from afar, for providing an important, grounding perspective, and for (sort of) learning to stop asking when I will be done with school.

List of Figures

Figure 1: Gut microbial-host metabolic interactions influence human health. 5

Figure 2: The human microbiota and its interactions with xenobiotics. 6

Figure 3: Host metabolism of xenobiotics. 8

Figure 4: Gut microbial hydrolase enzymes involved in xenobiotic metabolism. 9

Figure 5: Gut microbial lyases involved in xenobiotic metabolism. 11

Figure 6: Gut microbial reductases involved in xenobiotic metabolism. 12

Figure 7: Gut microbial transferase enzymes involved in xenobiotic metabolism. 13

Figure 8: Gut microbial radical enzymes involved in xenobiotic metabolism. 14

Figure 9: Gut microbial metabolism of anti-inflammatory and gastrointestinal agents. 17

Figure 10: Gut bacterial metabolism of anti-cancer agents. 19

Figure 11: Gut microbial metabolism of central nervous system (CNS) drugs. 21

Figure 12: Gut and vaginal microbial metabolism of anti-viral agents. 23

Figure 13: Gut microbial metabolism of traditional remedies and herbal supplements. 25

Figure 14: Gut microbial metabolism of simvastatin. 26

Figure 15: Gut microbial inactivation of the cardiac medication digoxin. 27

Figure 16: Na^+/K^+ ATPases function and inhibition by digoxin. 31

Figure 17: Binding of digoxin to Na^+/K^+ ATPases. 32

Figure 18: Identification of the cardiac glycoside reductase (*cgr*) operon in a digoxin-metabolizing strain of *E. lenta*. 36

Figure 19: Annotation of Cgr1. 49

Figure 20: Annotation of Cgr2. 50

Figure 21: Heterologous expression confirms that Cgr2 is sufficient for digoxin reduction. 51

Figure 22: Sequence similarity network (SSN) analysis reveals that Cgr2 is a highly distinct member of a large reductase family that is widespread in human associated microbes. 53

Figure 23: Cgr2 is a distinct flavoprotein reductase. 55

Figure 24: Multiple sequence alignments of fumarate reductases. 56

Figure 25: Multiple sequence alignments of urocanate reductases. 57

Figure 26: Multiple sequence alignments of ketosteroid dehydrogenases. 58

Figure 27: Comparative genomics expands the boundaries of the *cgr* operon 60

Figure 28: An expanded operon is present in all digoxin-reducing *E. lenta* strains. 61

Figure 29: A common Y/N333 polymorphism in Cgr2 leads to altered digoxin metabolism. 62

Figure 30: *E. lenta* and *cgr2* are prevalent in the human gut microbiota. 65

Figure 31: Heterologous expression of Cgr-MBP fusion proteins in *E. coli*. 80

Figure 32: Heterologous expression of Cgr proteins in *R. erythropolis*. 82

Figure 33: Purification of Cgr2 constructs from *R. erythropolis*. 83

Figure 34: *In vitro* digoxin reduction assay using methyl viologen as an electron donor. 85

Figure 35: [Fe-S] cluster(s) affect Cgr2 stability and oligomerization. 86

Figure 36: Cgr2 contains redox active, oxygen-sensitive [4Fe-4S] cluster(s). 88

Figure 37: Purification after reconstitution leads to loss of activity and stability of Cgr2. 90

Figure 38: Metal dependence and [4Fe-4S] cluster content of Cgr2. 91

Figure 39: Identification of six cysteine residues important for Cgr2 activity. 93

Figure 40: *In vitro* stability and activity of Cgr2 cysteine to alanine point mutants 94

Figure 41: Thermal stability and *in vitro* activity assays reveal that Cgr2 binds FAD poorly. 96

Figure 42: ^{1}H NMR spectrum of digoxin (recorded in CD_3OD at 600 MHz). 98

Figure 43: ^{1}H NMR spectrum of dihydrodigoxin generated by Cgr2 (recorded in CD_3OD at 600 MHz). 99

Figure 44: Overlay of ^{1}H NMR spectra of digoxin and dihydrodigoxin generated by Cgr2 (recorded in CD_3OD at 600 MHz) highlighting key protons in the lactone moiety. 100

Figure 45: COSY spectra of dihydrodigoxin (recorded in CD_3OD at 600 MHz). 101

Figure 46: ^{1}H NMR spectrum of dihydrodigoxin generated by Cgr2 (recorded in CD_3COCD_3 at 600 MHz). 102

Figure 47: Preliminary model for digoxin metabolism by Cgr1 and Cgr2. 105

Figure 48: Geographic and dietary factors correlate with digoxin metabolism in humans. 121

Figure 49: *E. lenta* generates ATP through the arginine dihydrolase pathway. 122

Figure 50: L-arginine does not directly inhibit Cgr2. 123

Figure 51: Digoxin and related cardenolides do not affect *E. lenta* growth in rich media. 124

Figure 52: Cardenolides do not affect *E. lenta* growth in basal growth media. 125

Figure 53: Initial kinetic characterization of Cgr2 towards digoxin. 126

Figure 54: Methyl viologen is the optimal electron donor for Cgr2 *in vitro*. 128

Figure 55: Cgr2 activity towards digoxin is reduced in the presence of detergent. 129

Figure 56: Optimized kinetics of Cgr2 variants for digoxin reduction. 130

Figure 57: Cardenolide natural products. 133

Figure 58:Bufadienolide natural products. 134

Figure 59: Furanones encountered in the diet. 135

Figure 60: α,β-unsaturated carboxylic acids compounds. 136

Figure 61: Ketosteroids encountered in the human gut. 136

Figure 62: Host-derived prostaglandins involved in inflammatory and GI processes. 137

Figure 63: The substrate scope of Cgr2 is restricted to cardenolides. 138

Figure 64: Confirmation of dihydrocardenolide product formation from *in vitro* reactions with Cgr2. .. 139

Figure 65: Modest MV oxidation is observed in *in vitro* reactions of Cgr2 with 2(5H)-furanone, although minimal product is formed. 140

Figure 66: Preliminary EPR experiments suggest that the [4Fe-4S] cluster(s) of Cgr2 transfer electrons to the FAD cofactor. 160

Figure 67: Approaches to linking the gut microbial community to drug metabolism. 167

Figure 68: Incorporating microbial metabolism into drug discovery and development. 170

Figure 69: Biomarkers of microbial drug metabolism can inform clinical trial design and personalized medicine. 171

List of Tables

Table 1: Primers and constructs for heterologous expression of Cgr1 and Cgr2 in *Rhodococcus erythropolis*. 69

Table 2: *Eggerthella lenta* and *Coriobacteriaceae* strains analyzed. 73

Table 3: qPCR primers and probes. 74

Table 4: Primers for heterologous expression and purification of Cgr1 and Cgr2. 106

Table 5: Primers for site-directed mutagenesis of Cgr2. 110

Table 6: [Fe-S] cluster-binding motifs that are not found in Cgr2. 113

Table 7: Comparison of Cgr2 kinetic parameters with those of related reductases. 131

List of Abbreviations

Ala	alanine
δ-ALA	aminolevulinic acid
BHI	brain heart infusion media
BV	benzyl viologen
° C	degree Celsius
COSY	homonuclear correlation spectroscopy
cgr	cardiac glycoside reductase
Cys	cysteine
DTT	dithiothreitol
DMF	dimethyl formamide
DMSO	dimethyl sulfoxide
EPR	electron paramagnetic resonance
FAD	flavin adenine dinucleotide
[Fe-S]	iron sulfur cluster
FMN	flavin mononucleotide
FPLC	fast protein liquid chromatography
g	gram
GC-MS	gas chromatography mass spectrometry
GF	germ free
GI tract	gastrointestinal tract
h	hour
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
His6x	Hexahistidine affinity tag
HMP	Human Microbiome Project

kb	kilobase
kDa	kilodalton
L	liter
LC-MS/MS	liquid chromatography tandem mass spectrometry
M	molar
MBP	Maltose-Binding Protein
min	minute
MV	methyl viologen
MW	molecular weight
N	Aspargine
NAD(P)H	nicotinamide adenine dinucleotide (phosphate)
NaDT	sodium dithionite
NMR	nuclear magnetic resonance
PS	phenosafranin
qPCR	quantitative polymerase chain reaction
RFU	relative fluorescence unit
RNA-seq	RNA sequencing
RPM	revolutions per minute
s	sec
SAM	*S*-adenosylmethionine
SEM	standard error of the mean
SSN	sequence similarity network
Tm	melting temperature
Tris	2-amino-2-(hydroxymethyl)propane-1,3-diol
UV-vis	ultraviolet visible
Y	Tyrosine

Chapter 1: The human gut microbiota influences drug efficacy

This chapter was adapted in part from:

(1) Koppel, N.; Rekdal, V.M.; Balskus, E.P. Chemical transformations of xenobiotics by the human gut microbiota. *Science*, **356**, eaag2770 (2017). DOI: 10.1126/science.aag2770.

(2) Koppel, N.; Balskus, E.P. Exploring and understanding the biochemical diversity of the human microbiota. *Cell Chem. Biol*. **23**, 18-30 (2016), DOI: http://dx.doi.org/10.1016/j.chembiol.2015.12.008.

1.1 Gut microbial involvement in human health

The human body is colonized by trillions of microorganisms including bacteria, archaea, viruses, and eukaryotes that are collectively referred to as the human microbiota. Increasing evidence suggests that these microbes have co-evolved with their human hosts (*1*) and play a central role in human health. In particular, the microbes that colonize the gastrointestinal (GI) tract influence human biology through a range of processes which include providing colonization resistance to pathogens, modulating the immune system, and synthesizing essential vitamins and nutrients (*2-4*). In addition, gut microbes can transform a wide array of ingested xenobiotics including dietary components, environmental chemicals and pollutants, and pharmaceuticals into metabolites with altered physical and biological properties (*5, 6*). Although such transformations have been appreciated for over 60 years, this metabolism has typically been attributed to the gut microbial community as a whole, and the specific microorganisms and enzymes involved in xenobiotic processing remain largely unknown (*5, 7*).

Previously, mechanistic studies of the human gut microbiota were limited to poorly defined communities (e.g. fecal samples) or microorganisms that could be isolated and grown in the lab. With the exception of pathogenic, highly abundant, or aero-tolerant microorganisms, many commensal gut microbes have not been extensively studied (*8-10*). Over the last two decades, significant technological advances in DNA sequencing, computational analyses, and culturing techniques have enabled comprehensive studies of the organisms present within the gut community (including those that have previously eluded cultivation) (*8, 9, 11, 12*), leading to a resurgence of interest in the impact of the gut microbiota on human health. Many labs and consortia have since cataloged the presence and abundance of gut microbes and genes in human populations (*11, 13, 14*) and linked gut microbial community composition to various host diseases and phenotypic outcomes including infection, metabolic disease, cancer, cardiovascular disease, autism, and drug response (*15-17*). However, a mechanistic understanding of the specific microbial contributions to human biology has lagged behind and represents a significant challenge for the human microbiota field.

Study of the human gut microbiota is further complicated by the complexity of the system as well as the marked variability among even healthy individuals. The GI tract is comprised of distinct regions (stomach; lower intestine: duodenum, jejunum, ileum; large intestine: caecum, colon) that differ with respect to epithelial cell physiology, pH levels, oxygen gradients, and nutrient content. These different regions support the growth of hundreds of distinct microbial species and influence the types of microbial metabolic processes that can occur (*14, 18, 19*). It has been estimated that 10^{13} microbes inhabit the human GI tract, which is approximately equal to the number of human cells in the body (*19*). The gut microbiota is dominated by obligate anaerobes from the Firmicutes and Bacteroidetes phlya, although large variability in both community composition and microbial abundance is observed among individuals (*11, 20*). Strikingly, the genetic content of these communities outnumbers the ~20,000 human genes by approximately 150- fold (*14*). The vast genetic potential of these organisms encodes expanded and variable metabolic and biosynthetic capabilities with the potential to profoundly impact human health. However, recent large-scale metagenomic studies, including data from the Human Microbiome Project (HMP), reveal a limited capacity to accurately identify the biochemical functions encoded by the human gut microbiome. For example, approximately 50% of the genes identified in the 139 assembled HMP stool metagenomes could not be annotated, and >86% could not be assigned to a known metabolic pathway (*11*). Even less is known about the metabolites associated with the human gut microbiota. For example, a study comparing plasma metabolomes of germ-free (GF) and conventionally reared mice found that over 10% of metabolites were significantly altered in the two groups, with many unique features detected only in the presence of microbes (*21*). Of these altered metabolites, only 2% could be assigned a chemical structure, underscoring a profound deficiency in our understanding of gut microbial chemistry.

Despite these gaps in knowledge, it is clear that the expanded chemical capabilities of gut microbes enable host-microbe interactions and influence human health. Gut bacteria can biosynthesize unique molecules, modify host metabolites, and metabolize dietary components and xenobiotics in distinct ways from the host to yield compounds that impart an array of physiological outcomes (Figure 1).

For example, microbial associated molecular patterns (MAMPs) are unique microbial cellular components or metabolites that interact with the host innate immune system. Perhaps the most extensively characterized MAMP is lipopolysaccharide (LPS), an immune-stimulatory component of the outer membrane of Gram-negative bacteria (Figure 1A) (*22*). Gram-negative bacterial clearance during infection is dependent on binding of LPS to the host receptor TLR4 (*23*), and certain gut microbes chemically modify LPS to evade detection by the immune system (*22*). An additional class of MAMPs is the zwitterionic bacterial capsular polysaccharides (ZPS), which are made by various gut isolates including *Bacteroides fragilis*. These depolymerized molecules bind to host MHCII proteins to elicit downstream immunological effects (*24*). For example, polysaccharide A from *B. fragilis* activates CD4+ T cells, induces secretion of anti-inflammatory cytokines, and maintains a balance in T helper cell levels (*25*).

In addition to synthesizing unique metabolites, gut microbes can also interface with host functions to produce co-metabolites with important biological activities. Gut bacteria can synthesize and metabolize essential amino acids, generating molecules that serve as substrates for host pathways. The co-metabolite *p*-cresyl sulfate arises from gut bacterial metabolism of L-tyrosine to *p*-cresol, followed by sulfation by host enzymes (Figure 1B) (*26*). This molecule is thought to contribute to uremic toxicity and accumulates in the serum of patients with end-stage renal and chronic kidney diseases (*27*). This co-metabolic pathway also indirectly influences host metabolism of other xenobiotics, including the drug acetaminophen, as *p*-cresol competes with these molecules for sulfation by host enzymes (*28*). Another important co-metabolic process is bile acid (BA) metabolism. Primary BAs are steroids synthesized from cholesterol by host liver enzymes and are metabolized by gut microbes into more than 50 different secondary BAs (Figure 1B) (*29*). Primary and secondary BAs have a variety of important physiological roles including solubilization of dietary compounds, anti-microbial activity, and diverse signaling and regulatory functions (*29, 30*). Recently, secondary BA formation by the gut commensal *Clostridium scindens* via the 7α/β-dehydroxylation pathway was shown to enhance resistance to *Clostridium difficile* infections in mice by directly inhibiting growth of this pathogen (*2*).

Finally, the gut microbiota metabolizes ingested compounds, including dietary components and pharmaceuticals. The most well studied of these interactions is gut bacterial fermentation of indigestible dietary polysaccharides to produce short chain fatty acids (SCFAs), including acetate, propionate, and butyrate (Figure 1C). These metabolites serve as energy sources for distinct cell types and tissues that account for up to 10% of daily caloric value, as precursors and regulators of important host molecules such as cholesterol, fatty acids, and glucose, and as inhibitors of autophagy and of histone-deacetylase activity (*31-34*). SCFAs can also activate G-protein coupled receptors and lead to downstream effects including production of the hormone leptin (*35*), anti-inflammatory effects (*36*), and tumor suppression (*37*). Human gut bacteria also contribute to drug metabolism, as discussed further in Section **1.3**.

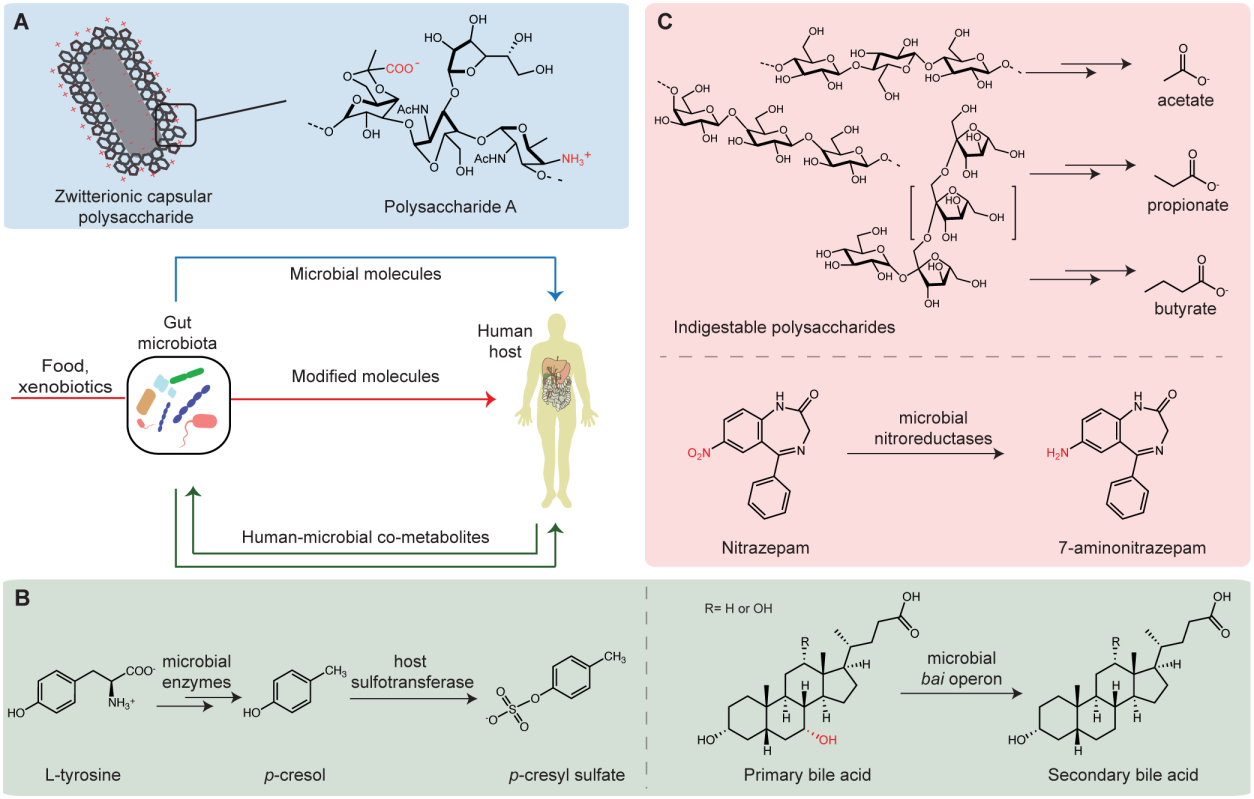


Figure 1: Gut microbial-host metabolic interactions influence human health.
(A) Gut bacterial molecules that interact with the host immune system. (B) Co-metabolic pathways. (C) Gut bacterial conversion of dietary components and other xenobiotics.

1.2 Contrasting metabolism of xenobiotics by gut microbial and human enzymes

The metabolism and circulation of ingested substrates is a highly complex process that occurs across multiple organ systems, and involves both human and microbial enzymes (Figure 2A). Orally ingested compounds pass through the upper GI tract to the small intestine, where they can be modified by secreted, digestive host enzymes and absorbed by host tissues (*38*). Readily absorbed xenobiotics pass between or through intestinal epithelial cells, where they are subject to processing by host enzymes. These molecules can also be transported via the portal vein to the liver, where they can be further modified by a vast collection of host xenobiotic-processing enzymes (Section **1.2.1**). These xenobiotics and their resultant metabolites can enter systemic circulation and distribute throughout host tissues, including distal organs, to impart their pharmacological properties. Compounds administered via other routes, for example intravenously administered compounds, bypass "first-pass" metabolism and are immediately introduced into systemic circulation. Xenobiotics and their metabolites may also undergo biliary excretion, where they are secreted from the liver to the bile and back into the lumen of the GI tract (Figure 2B). Depending on their chemical properties, xenobiotics and their metabolites may be reabsorbed by intestinal cells (enterohepatic circulation), excreted from the feces, or filtered through the kidneys and eliminated into the urine (Figure 2C).

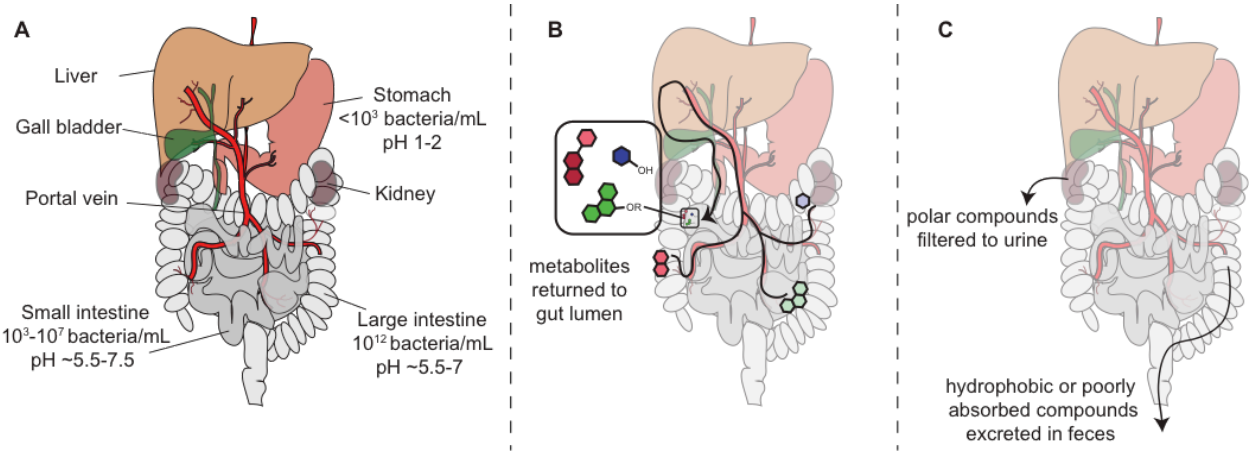


Figure 2: The human microbiota and its interactions with xenobiotics.
(A) Key sites of microbial colonization and xenobiotic metabolism. (B) Biliary excretion returns xenobiotics and metabolites to the gut lumen. (C) Routes of xenobiotic and metabolite excretion from the body.

At multiple points along this metabolic axis, xenobiotics may encounter gut microorganisms and their metabolic enzymes. For example, compounds that undergo biliary excretion or efflux by drug transporters (e.g. P-glycoprotein) are pumped into the gut lumen (*39*). Additionally, poorly absorbed xenobiotics continue through the small intestine into large intestine. Within the densely populated large intestine, these compounds may be subject to extensive microbial metabolism, leading to products with altered bioactivity, bioavailability, and toxicity (*5, 6, 40*). Additionally, these molecules can interfere with the activities of human xenobiotic-metabolizing enzymes, affecting the fates of other ingested molecules. The aggregate metabolism of host and microbe significantly alters the bioactivities and lifetimes of ingested compounds within the human body.

Several important features distinguish gut microbial xenobiotic metabolism from analogous host-mediated transformations. The human gut microbiome encodes a broad diversity of enzymes, many of which are exclusively microbial, and which often perform distinct chemistry from the host (*41-45*). Gut microbes typically carry out hydrolytic and reductive reactions to metabolize xenobiotics (*5, 6*), as opposed to the typically oxidative and conjugative chemistry of human xenobiotic-processing enzymes. These differences are due in part to physiological context, such as oxygen concentrations, but may also reflect distinct evolutionary pressures. For example, the abundance and diversity of host cytochrome P450s may reflect a need to detoxify and excrete diverse compounds to which humans were exposed throughout evolution (*46*), whereas microbial breakdown of compounds could enable niche colonization in the dense, highly competitive environment of the gut (e.g. through use of alternative nutrients or energy sources). Many host and microbial xenobiotic-metabolizing enzymes rely on complex cofactors that must be synthesized or acquired from the diet or the gut environment (*47*), an energetically costly process that further supports the physiological importance of xenobiotic metabolism. It is also possible that certain gut microbial enzymes were not evolutionarily selected to process specific xenobiotics and that metabolism may instead arise from relaxed substrate specificity. Many enzyme classes that are associated with xenobiotic metabolism including hydrolases, lyases, oxidoreductases, and transferases, are widely distributed in sequenced gut microbes (*43-45, 48, 49*) and represent some of the most abundant protein

families in the human gut ecosystem (*50, 51*). Together, the human host and the gut microbiota significantly alter the bioactivities and lifetimes of xenobiotics using a diverse array of enzymes.

1.2.1 Host enzymes

Human xenobiotic metabolism occurs predominantly in the liver and intestinal cells, and modifies nonpolar compounds into hydrophilic, higher molecular weight metabolites (*38, 52*). These metabolites are more easily excreted from the body via the kidney, which filters out hydrophilic compounds. This process occurs in two phases: the installation or exposure of polar functional groups ('phase I') and the conjugation of these groups to more polar metabolites ('phase II') (Figure 3). Phase I enzymes perform oxidative, reductive, or hydrolytic reactions to yield various polar substituents, including hydroxyl groups, epoxides, thiols, and amines. The largest and best-characterized class of enzymes participating in phase I metabolism is the cytochrome P450s, which have 57 isoforms in humans (*53*). Additional phase I enzymes that are important in xenobiotic processing include carboxylesterases and flavin monooxygenases (*38*). Phase II enzymes are predominantly transferases that append glucuronyl, methyl, acetyl, sulfonyl, and glutathionyl groups onto xenobiotics or their phase I metabolites (*52*). Polymorphisms in the genes encoding these enzymes greatly influence how individuals respond to dietary and pharmaceutical interventions (*53*).

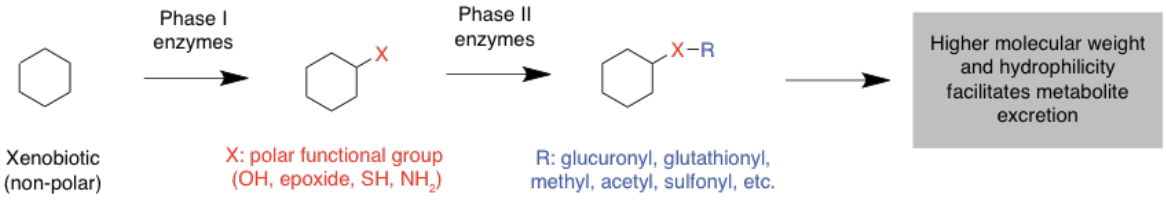


Figure 3: Host metabolism of xenobiotics.

1.2.2 Hydrolytic enzymes

Both the host and gut microbiota use hydrolytic chemistry to break down large luminal substrates into various nutrients and building blocks. Hydrolase enzymes catalyze the addition of a water molecule

to a substrate, followed by bond cleavage. The most abundant and relevant hydrolases in the GI tract are proteases, glycosidases, and sulfatases, and the microbiota contributes a broader range of activities than host enzymes (Figure 4). Proteases cleave the peptide bonds linking the amino acids in polypeptide chains. While the small intestine is dominated by pancreatic serine proteases, colonic microbes express many cysteine- and metalloproteases (*54*) with different substrate specificities and potentially different clinical consequences (*55*). Anaerobic microbes can also reduce disulfide bonds in proteins, enhancing their susceptibility for degradation (*54*). Glycosidases hydrolyze glycosidic bonds using a dyad of carboxylic acid residues and a water molecule, releasing free sugars (*56*) that can be used a carbon source. These enzymes process a huge diversity of glycoconjugates and oligosaccharides, and are widespread in gut microbes; for example, 43% of bacterial genomes in the HMP database contained β-glucuronidase homologs (*43, 57*). Sulfatases are also widely distributed among gut microbes and hydrolyze sulfate esters generated by phase II host metabolism using the unusual amino acid formylglycine (*58*). The hydrate form of this residue is thought to undergo transesterification with a sulfate ester substrate to generate a tetrahedral intermediate that breaks down to release sulfate and reform the aldehyde (*59*).

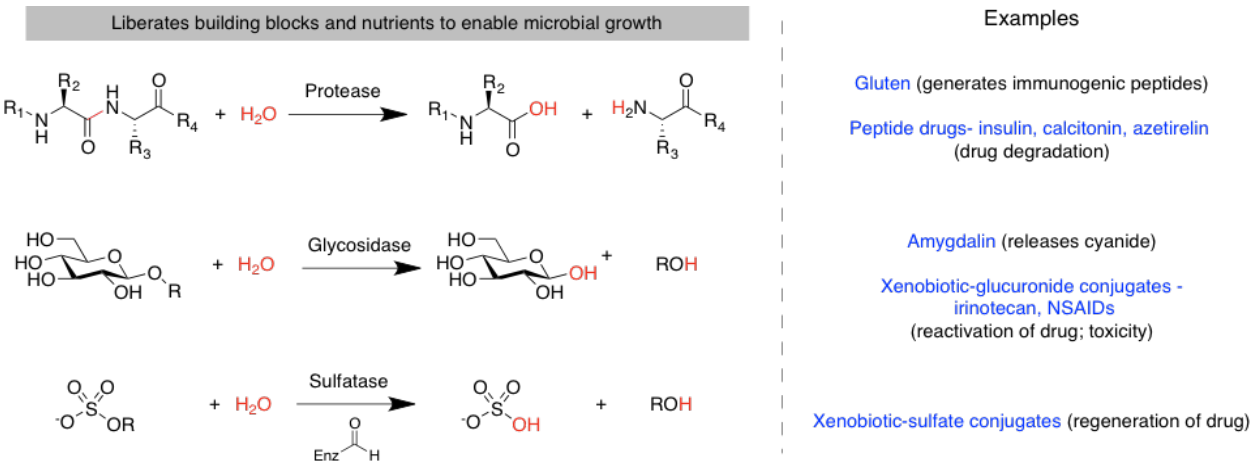


Figure 4: Gut microbial hydrolase enzymes involved in xenobiotic metabolism. Proteases, glycosidases, and sulfatases are the most common gut microbial hydrolase enzymes. Known substrates of these enzymes are shown in blue, and the consequences of metabolism for the human host are shown in parentheses.

Hydrolytic reactions alter both the physical properties and activities of xenobiotics and their metabolites. For example, removal of a glucuronide in the gut lumen is generally accompanied by a decrease in polarity that can allow reabsorption by host cells and thereby extend the lifetime of a molecule within the body, as seen with glucuronide conjugates of nonsteroidal anti-inflammatory drugs (NSAIDs) and the cancer therapy irinotecan (*60, 61*). Hydrolysis can also alter the biological activity or toxicity of xenobiotics, as observed for plant-derived glycosides like amygdalin and the artificial sweetener cyclamate (*62, 63*). Additionally, hydrolysis is often a prerequisite for further transformations, such as the fermentation of sugars released from indigestible polysaccharides (*31*) and the products of hydrolytic reactions (sugars, amino acids, sulfate) often support microbial growth and survival in the gut.

1.2.2 Lyases

Lyases are enzymes that break C–C or C–X bonds (where X = O, N, S, P or halides) without relying on oxidation or addition of water. Within the gut microbiota, the predominant classes of lyase enzymes are polysaccharide lyases (PL) and C-S β-lyases (Figure 5). A single representative microbiome encoded >5000 PLs (*43*), demonstrating an enormous diversity of possible carbohydrate substrates which could support microbial growth. Microbial PLs modify polysaccharides that contain a glycosidic bond β- to a carboxylic acid (e.g. alginate, pectin, chondroitin, and heparan). The presence of the carboxylate enables removal of the α-proton and subsequent β-elimination to yield an α,β-unsaturated sugar and a hemiacetal (*48*).

Microbial C–S β-lyases cleave C–S bonds found in both dietary compounds and cysteine-*S*-conjugates of xenobiotics, which are formed by liver enzymes in the mercapturic acid pathway. These enzymes generate an aldimine linkage between pyridoxal 5-phosphate (PLP) and the α-amino group of the cysteine-derived substituent, acidifying the adjacent proton. β-elimination releases a thiol-containing metabolite and aminoacrylate, the latter of which spontaneously breaks down to form ammonia and pyruvate (*64*). Microbes can further metabolize these thiols, altering their physical properties and localization within the body. For example, gut bacterial C–S β-lyases cleave cysteine-*S*-conjugates of

polychlorinated biphenyls to yield toxic thiol metabolites that are further methylated and accumulate in lipophilic host tissues (*65*). Microbial C–S lyases have also been implicated in the metabolism of conjugates of the herbicide propachlor (*66*) and the motor oil additive bromobenezne (*67*), although the specific enzymes have yet to be discovered. The consequences of C–S lyase chemistry for the gut microbiota are not well understood, and the activity may derive from promiscuous PLP-dependent enzymes involved in "housekeeping" functions (*64*). However, ammonia generated by C–S β-lyases could serve as the sole nitrogen source for the industrial bacterium *Corynbacterium glutamicum* (*68*), pointing to a potential role in nutrient acquisition.

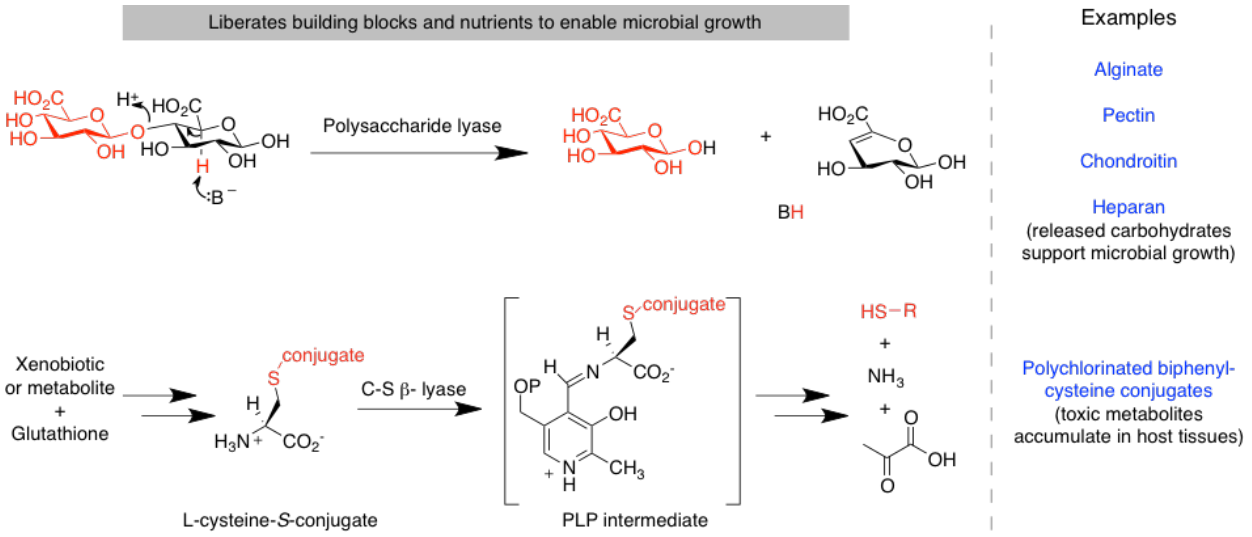


Figure 5: Gut microbial lyases involved in xenobiotic metabolism.
Polysaccharide lyases and C-S β-lyases are the predominant lysases in gut microbes. Known substrates of these enzymes are shown in blue, and the consequences of metabolism for gut microbes and the human host are shown in parentheses.

1.2.3 Reductases

Gut microbes can reduce a wide range of functional groups, including alkenes and α,β-unsaturated carboxylic acid derivatives, nitro-, *N*-oxide, azo-, and sulfoxide groups (Figure 6). These multi-electron reductions cannot be performed by standard amino acid residues and instead rely on a variety of complex cofactors, including NAD(P)H, flavin, [Fe-S] clusters, (siro)heme, molybdenum-cofactor and other metallocofactors to mediate the transfer of electrons or hydride equivalents (H^+, $2e^-$) to

substrates, followed by proton delivery by charged amino acid residues (*69-71*). Biochemical and structural characterization of gut microbial reductases has uncovered individual enzymes that display broad substrate scope and can reduce a variety of functional groups including quinone, azo-, and nitro-groups (*72, 73*), and consequently their endogenous substrates and *in vivo* relevance are often unclear. Other studies have identified multiple enzymes within a single bacterium that all mediate the same reaction (e.g sulfoxide reduction) (*70*), highlighting the breadth and importance of reductive metabolism for particular gut microbes.

Reduction typically decreases the polarity of compounds and can alter charge, hybridization, and electrophilicity, which can affect the lifetimes and activities of metabolites in the body (*74-76*). In some cases, substrates serve as alternative terminal electron acceptors in anaerobic respiration, which confers a growth advantage to metabolizing microbes (*69*). While reductase enzymes are found in humans, many reductive transformations are uniquely microbial (*7, 72, 74*) and many have not yet been linked to known enzymes or organisms.

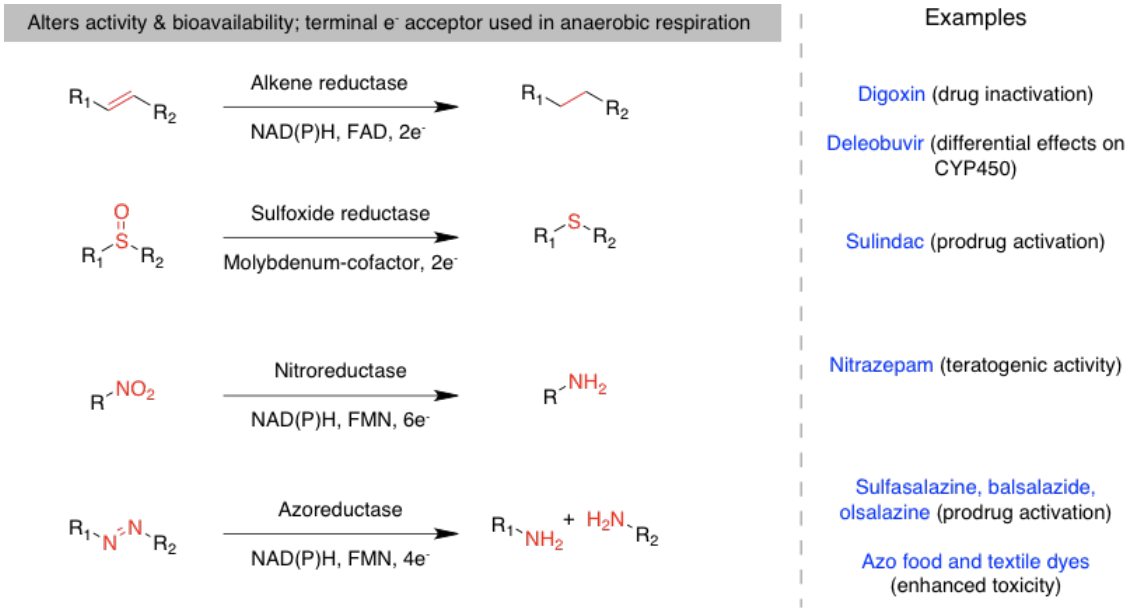


Figure 6: Gut microbial reductases involved in xenobiotic metabolism.
Common gut microbial reductases include alkene, sulfoxide, nitro-, and azo- reductases. Known substrates of these enzymes are shown in blue, and the consequences of metabolism for the human host are shown in parentheses.

1.2.4 Transferase enzymes

These enzymes transfer functional groups between a xenobiotic and a cofactor or activated co-substrate using nucleophilic chemistry (Figure 7). Microbial methyl transferases utilize *S*-adenosylmethionine (SAM) or methylcobalamin as methyl donors (*77*), while demethylation requires nucleophilic cofactors such as corrinoids (e.g. cobalamin) or tetrahydrofolate to remove methyl groups that can be used as a carbon source (*78-80*). Gut microbial enzymes can also acylate xenobiotics; for example using acetyl-coenzyme A (acetyl-CoA) as a donor, microbes *N*-acetylate the anti-inflammatory agent 5-aminosalicylic acid to generate an inactive metabolite (*81*). In addition to modulating the bioactivities of xenobiotics, appendage or removal of functional groups can alter the polarity and bioavailability of these compounds (*65, 82*).

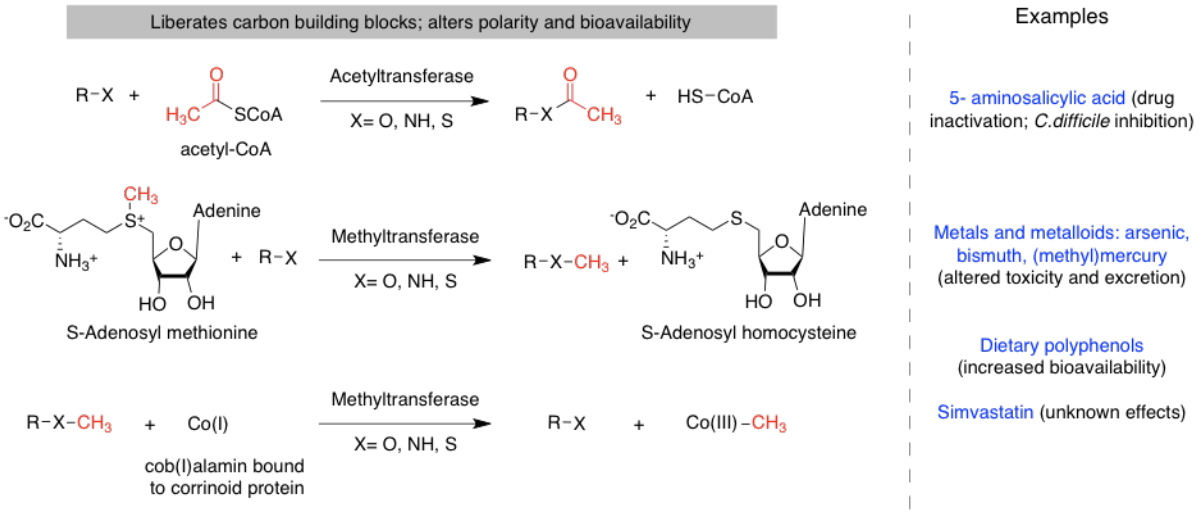


Figure 7: Gut microbial transferase enzymes involved in xenobiotic metabolism.
Common gut microbial transferases include acetyltransferase and methyltransferases. Known substrates of these enzymes are shown in blue, and the consequences of metabolism for the human host are shown in parentheses.

1.2.5 Radical enzymes

Gut microbes employ radical enzymes to carry out difficult chemical reactions such as bond-formation (C–C), bond-cleavage (C–C, C–O, C–N), and skeletal rearrangements using high-energy intermediates containing an unpaired electron (*83*). These enzymes are often inactivated by oxygen and

allow anaerobic organisms in the anoxic environment of the gut to perform challenging reactions. In fact, some of the most abundant protein families in the human gut microbiome are radical enzymes (*50, 84*). The first step in radical catalysis involves generation of a protein- or cofactor- based radical species, either through homolytic bond cleavage or through electron transfer that is mediated by cofactors such as SAM, cobalamin, and [Fe-S] clusters. Following formation of the radical species on the enzyme or cofactor, the radical is transferred onto the substrate. The substrate-based radical then undergoes transformation to a product-based radical. Formation of the final product may be accompanied by regeneration of the initial enzyme- or cofactor- centered radical (*83*) (Figure 8). Radical enzymes are common in primary metabolism, but can also influence the metabolism of a variety of xenobiotics. For example, CutC, a member of the glycyl radical enzyme (GRE) family, is responsible for cleavage of the dietary nutrient choline into a disease-associated metabolite trimethylamine (TMA) (*42*). An additional GRE, *p*-hydroxyphenylacetate decarboxylase, produces *p*-cresol (*28*), which is sulfated by host enzymes into a uremic toxin that also impacts host processing of additional xenobiotics (see **1.1**).

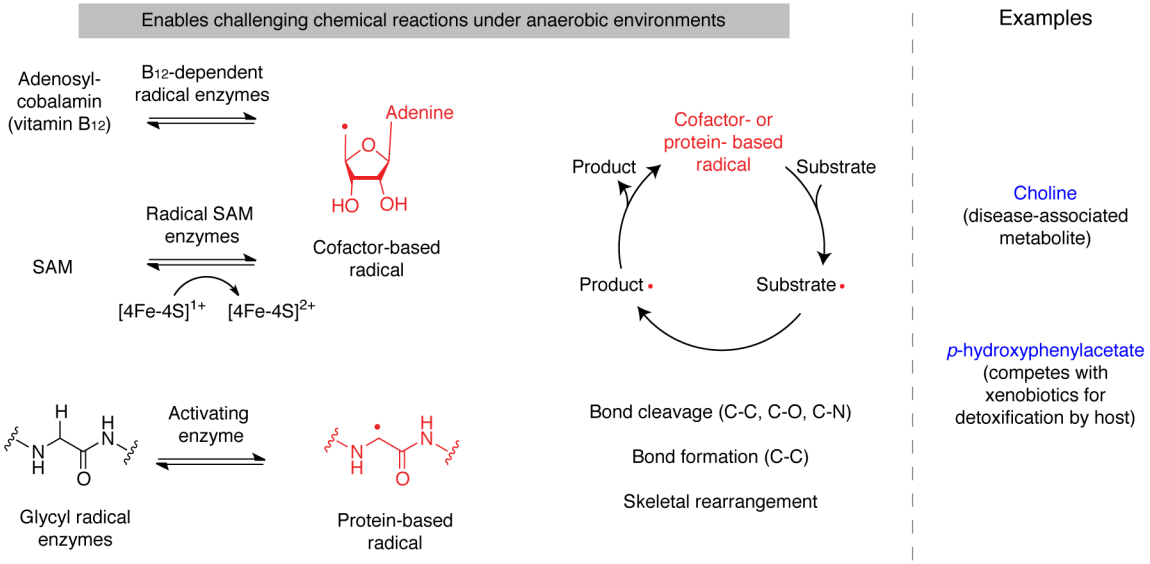


Figure 8: Gut microbial radical enzymes involved in xenobiotic metabolism.
Known substrates of these enzymes are shown in blue, and the consequences of metabolism for the human host are shown in parentheses.

1.3 Gut microbial metabolism of pharmaceuticals

Excluding antibiotics, the human gut microbiota is involved in the transformation of over fifty pharmaceuticals, yielding metabolites with altered biological properties and toxicities (*5, 6, 40*). In some cases, drug metabolites that were teratogenic (*85*), toxic (*60, 61*), and even lethal (*86*) were only observed when drugs were already on the market. Apart from direct metabolism of pharmaceuticals, microbial products can also regulate host xenobiotic metabolic pathways (*87*). Additionally, various drugs alter microbial community composition in the gut (*88, 89*), which affects the composite metabolic potential of the colonized host. Despite the tremendous influence of microbes on the pharmacological properties of drugs spanning many indications and host targets, the responsible microbes and enzymes have only been identified or characterized in a handful of cases. It is thus critical to characterize the mechanisms of microbial drug metabolism and interaction, which vary substantially among individuals and may contribute to differential patient responses.

Known examples of drug metabolism by human-associated microbes are discussed in the following sections. The identity and distribution of relevant microbes, genes, and enzymes involved in drug metabolism are described, although many of these details are unknown. While ongoing studies have illuminated a complex interplay between drugs and microbes that may indirectly influence human health (*88, 90-92*), we focus here on the direct modification of drugs by gut microbes, and the known effects of these microbial metabolites on gut microbial and human physiology.

1.3.1 Anti-inflammatory and gastrointestinal agents

Microbial metabolism of drugs targeted to the GI tract is perhaps unsurprising considering this is a major site of microbial colonization. Nevertheless, the scope of transformations and resultant effects on host health is significant. Microbes can enhance the toxicity and activity of drugs, as is observed with non-steroidal anti-inflammatory drugs (NSAIDs). Many NSAIDs contain a carboxylic acid that is glucuronidated by liver enzymes to facilitate excretion from the body. However, this inactive conjugate can re-enter the gut where it is deconjugated by microbes to release the active component in the gut lumen

(Figure 9A) (*93*). NSAID inhibition of cyclooxygenases in the gut epithelium leads to lower production of prostaglandins, immune compounds that maintain mucosal defense (*94*). In addition, NSAIDs are weakly acidic and can enter epithelial cells leading to intestinal injury (*94*). Therefore, prolonged use of NSAIDs leads to intestinal damage and ulcer development (*95*), which can be attributed to the deconjugation activity of bacterial β-glucuronidase enzymes. However, administration of β-glucuronidase inhibitors in mice alleviated the toxicities associated with the NSAIDS indomethacin, ketoprofen, and diclofenac (*60*). Gut microbial metabolites of gastrointestinal agents can also alter expression of host xenobiotic-metabolizing enzymes. In rat models, gut microbes reduced the proton pump inhibitor omeprazole (*96*), generating metabolites later shown to be possible ligands of aryl hydrocarbon receptor, a transcription factor involved in regulating xenobiotic-metabolizing enzymes (*87*). Additionally, omeprazole use in gastric ulcer prophylaxis was found to increase incidence of *C. difficile* infections by 65% (*97*). Longitudinal studies found that omeprazole use decreased microbial diversity in the gut (*98*), which could result in lowered colonization resistance and increased susceptibility to infection.

Microbial metabolism of anti-inflammatory and gastrointestinal agents can also be beneficial, particularly in the context of prodrug activation. For instance, gut microbial reduction of the *N*-oxide functional group found in the anti-diarrheal prodrug loperamide oxide liberates the active drug loperamide only at the desired site of action (GI tract), which significantly lowers systemic exposure and associated side effects (Figure 9B) (*75*). Another prominent example of prodrug activation is the gut microbial metabolism of anti-inflammatory drugs containing azo-linkages. Sulfasalazine, used for arthritis and inflammatory bowel disease (IBD), is rapidly cleaved by microbial azoreductases to produce sulfapyridine and 5-aminosalicylic acid (5-ASA), the latter of which is the active moiety in the treatment of Crohn's disease and ulcerative colitis (Figure 9C) (*76, 99*). Balsalazide and olsalazine underwent similar metabolism *in vitro* by fecal communities, and metabolism rates matched clinical data from humans (*100*). Various intestinal bacteria can further convert 5-ASA into *N*-acetyl 5-ASA, a metabolite that lacks anti-inflammatory activity, and significant variability in acetylation levels was observed in patients (*101*). Together with differences in the extent of azo reduction, this observation could potentially

explain variable therapeutic efficacy of sulfasalazine in patients. Additionally, *N*-acetyl ASA was shown to inhibit the growth of several microbes including *C. difficile* (*81*), suggesting that this metabolism could alter gut microbiota composition.

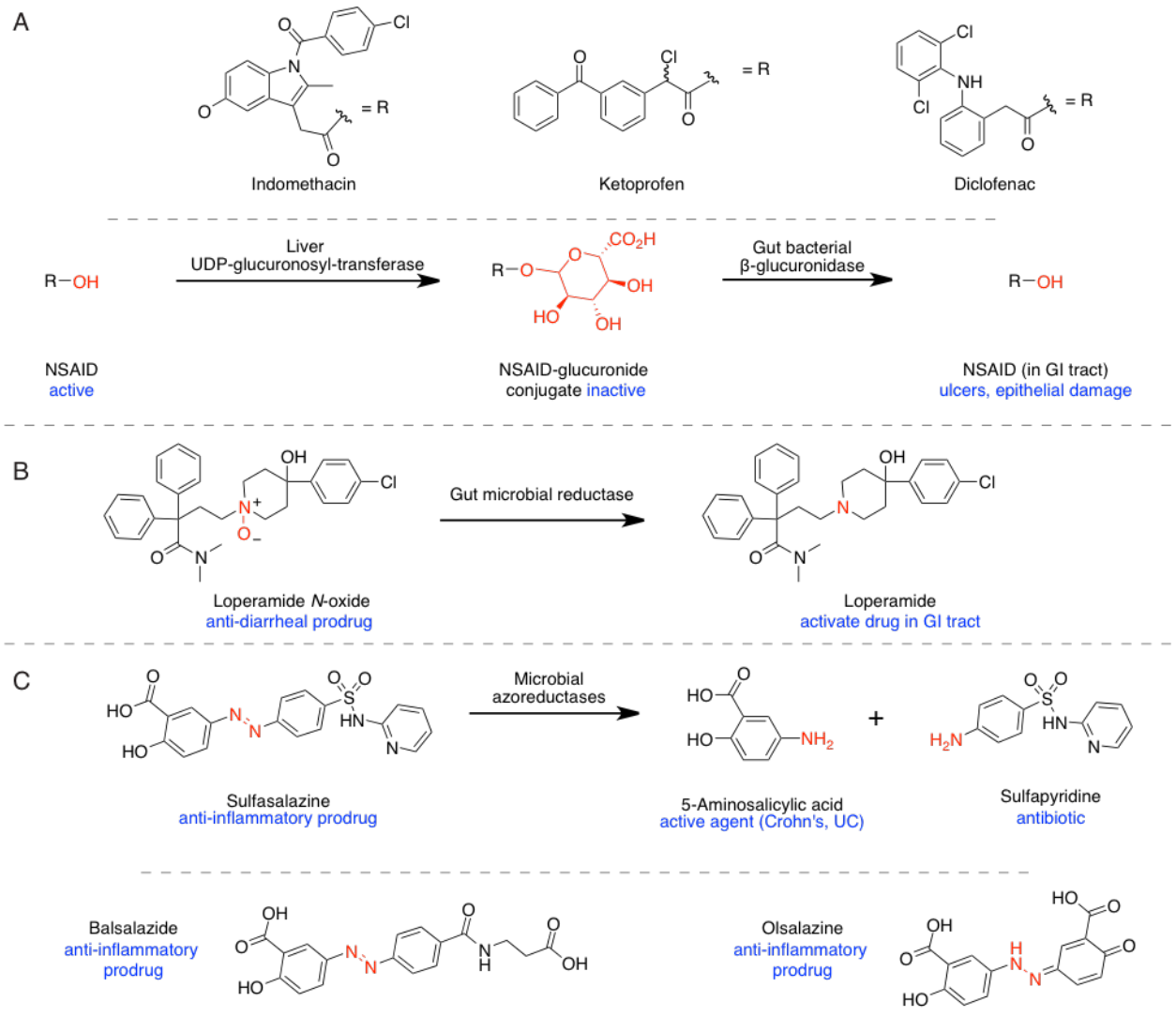


Figure 9: Gut microbial metabolism of anti-inflammatory and gastrointestinal agents. (A) NSAIDs are glucuronidated by host enzymes. In the GI tract, bacterial β-glucuronidases cleave the sugar and regenerate active drug, leading to gut epithelial damage and ulcers. (B) Gut microbial reductases activate the prodrug anti-diarrheal Loperamide-*N*-oxide in the gut, limiting systemic toxicity. (C) Microbial azoreductases reduce azo-linked anti-inflammatory drugs to generate the active agent 5-aminosalicylic acid.

1.3.2 Cancer chemotherapeutics

The contribution of gut microbial metabolism to chemotherapeutic efficacy is well established, and contributes to desired activities as well as toxicity. One of the best-characterized examples of microbial drug metabolism occurs in the context of bacterial metabolism of irinotecan into a toxic, dose-limiting metabolite within the GI tract (*61*). Irinotecan is a prodrug of the topoisomerase inhibitor SN-38. The active drug is glucuronidated by host liver enzymes, generating an inactive conjugate (SN-38G). SN-38G enters the gut via biliary excretion where it is hydrolyzed by gut bacterial β-glucuronidases to regenerate the active chemotherapeutic agent (*61*). SN-38 then enters colonic epithelial cells, causing intestinal damage and severe diarrhea, side effects that limit the use of this otherwise effective drug (Figure 10A). β-glucuronidases are broadly distributed in commensal bacteria (*57*) and are also present in humans. Using an *in vitro* high-throughput screen against *Escherichia coli* β-glucuronidase, researchers successfully identified potent inhibitors of gut bacterial β-glucuronidases (*61*), one of which was able to prevent metabolism of SN-38G in the gut and concomitant toxicity in mice. These inhibitors were also selective for gut bacterial enzymes, and did not affect human β-glucuronidase activity (*61*). Because bacterial β-glucuronidases can deconjugate glucuronides derived from many dietary compounds and drugs, inhibitors of these enzymes may be useful in therapeutic contexts outside of cancer treatment (*60*).

In the last several decades, scientists have found substantial evidence that certain gut bacteria associate with human tumors and are correlated with different therapeutic outcomes (*102, 103*). One possible route by which microbes alter cancer outcomes is through direct metabolism of chemotherapies within the tumor microenvironment. Indeed, recent studies demonstrate that the gut bacteria *E. coli* and *Listeria welshimeri* altered the efficacy of 16 (out of 30 tested) anticancer agents towards cancer cell lines (*104*). In addition, *E. coli* was shown to alter the *in vivo* efficacy of two tested drugs (gemcitabine and CB1954) in murine models (*104*). Expanding on these results, researchers recently found that *Gammaproteobacteria* expressing a long isoform of the cytidine deaminase (CDD_L) enzyme were responsible for metabolizing the pancreatic cancer agent gemcitabine into the inactive metabolite 2’2’-

difluorodeoxyuridine (Figure 10B). (*105*). CDD_L^+ bacteria conferred resistance to pancreatic cancer treatment in animal models, and could be detected in 76% of human pancreatic tumors (*105*), suggesting that gut bacteria modulate cancer drug efficacy within the tumor site.

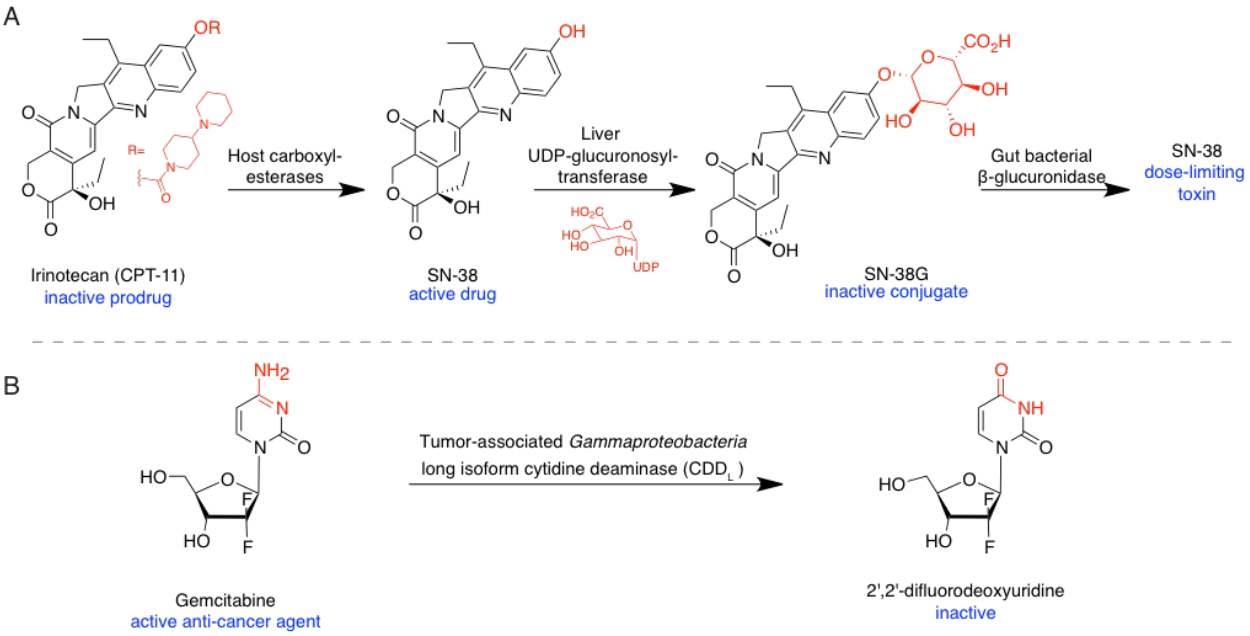


Figure 10: Gut bacterial metabolism of anti-cancer agents.
(A) The inactive prodrug irinotecan is cleaved by host carboxylesterases to generate the active topoisomerase inhibitor SN-38. SN-38 is glucoronidated by host enzymes, yielding an inactive conjugate. In the GI tract, bacterial β-glucuronidases cleave the sugar and regenerate active drug, which leads to dose-limiting diarrhea. (B) Tumor-associated *Gammaproteobacteria* harboring a long isoform of the CDD gene inactivate the pancreatic cancer drug gemcitabine.

While gut bacteria have demonstrated a high propensity for directly metabolizing therapeutic agents, emerging data also demonstrates that specific members of the gut microbiota may impact cancer therapies by interacting with the host immune system. Certain *Bacteroides* and *Bifidobacteria* species were shown to be crucial for cancer immunotherapies by modulating the host immune response to enhance the anti-tumor effects of Ipilimumab and anti-PD-L1 antibodies in mice (*106, 107*). More recently, gut microbial composition was shown to influence patient outcomes to anti-PD-1 immunotherapies in melanoma and epithelial cancers. In particular, *Ruminococcacceae* species (*92*) and

Akkermansia muciniphila (*91*) were more prevalent in patients that responded to anti-PD-1 therapy. Gut microbial induced inflammatory responses also impact cancer therapies that do not directly interact with the immune system. For example, commensal-dependent modulation of myeloid-derived cells increased the efficacy of the DNA-damaging agents cisplatin and oxlaplatin (*108*). These studies highlight the extensive and complex ways in which the gut microbiota can impact cancer outcomes in humans.

1.3.3 Biologics and peptide-based drugs

Although many biologics are delivered intravenously, smaller peptides and even insulin (*109*) are available as oral formulations. Biliary excretion has also been observed for hormones and shorter peptides (*110*), which could enable microbial interaction with peptide-based drugs. Gastrointestinal microbes from rats were shown to degrade the peptide drugs insulin and calcitonin (*55*) and azetirelin (*111*), and addition of protease inhibitors or antibiotics to cecal suspensions reduced metabolism. A panel of 17 peptides was also shown to have variable levels of degradation when incubated with human fecal samples (*54*). As peptide-based therapeutics are growing in popularity, elucidating the impact of gut microbes on their metabolic fate will become increasingly relevant.

1.3.4 Central nervous system (CNS) drugs

In addition to affecting drugs that act locally, gut microbial metabolism can also influence the efficacy of therapeutics that target organ systems distant from this body site. Many prominent examples can be found among CNS drugs. For example, oral levodopa (L-dopa) is used to treat Parkinson's disease, a condition characterized by dopaminergic-neuronal death. L-dopa crosses the blood-brain barrier, where it is decarboxylated by host enzymes to restore dopamine levels (*112*). However, extensive metabolism within the gut by both host and microbial enzymes affects the concentration of drug that reaches the brain. Microbial decarboxylation (*113*) and *p*-dehydroxylation convert L-dopa to *m*-tyramine, which can be further oxidized to *m*-hydroxyphenylacetic acid (Figure 11A) (*114*). Differences in these activities may contribute to the substantial variation observed in patient response to L-dopa (*115*). Although a tyrosine

decarboxylase from a food-associated strain of *Lactobacillus brevis* accepts L-dopa *in vitro* (*116*), the human gut microbes and enzymes responsible for L-dopa metabolism are unknown.

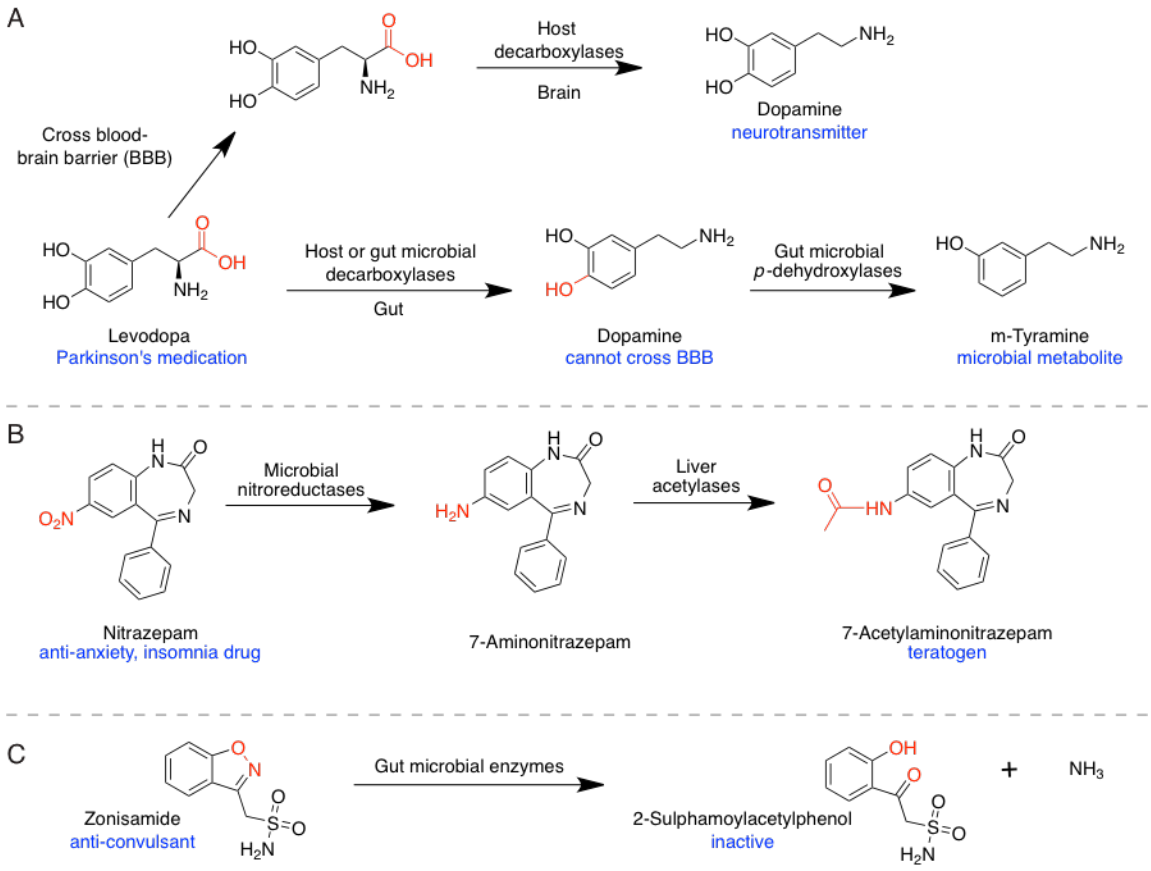


Figure 11: Gut microbial metabolism of central nervous system (CNS) drugs.
(A) Levodopa is a prodrug that is able to cross the blood brain barrier, and is cleaved by host carboxylesterases to replenish dopamine in the brain. Levodopa can also be decarboxylated by host and gut microbial carboxylesterases to generate dopamine outside of the brain. Gut microbes further metabolize dopamine via *p*-dehydroxylation to yield *m*-tyramine. (B) Microbial nitroreductases metabolize the anti-anxiety and insomnia drug nitrazepam into a metabolite that is further acetylated by host enzymes to generate a teratogenic compound. (C) Gut bacterial enzymes inactivate the anti-convulsant drug zonisamide.

Another example of a CNS drug that is modified by gut bacteria is nitrazepam (NZ), an anxiety and insomnia medication whose sequential metabolism by microbes and host generates the teratogen 7-acetylaminonitrazepam (AANZ) (Figure 11B). Microbial involvement in the initial reduction of NZ to 7-

aminonitrazepam was confirmed in culture (*71*), and AANZ was shown to induce fetal abnormalities in the offspring of pregnant rats (*85*). Very high exposure of NZ during human pregnancy also led to an increased rate of congenital disease in children (*117*), and its teratogenic effects are likely attributed to the co-metabolite AANZ. Finally, the anti-convulsant, zonisamide, was transformed into the inactive metabolite 2-sulphamoylacetylphenol by bacterial cultures as well as by the rat gut microbiota (Figure 11C) (*118*). As studies continue to reveal connections between the gut microbiota and various neurological diseases (*119*), it will become increasingly important to identify and characterize additional microbial interactions with CNS-targeted drugs.

1.3.5 Anti-viral agents

Microbes from both the gut and vaginal microbiotas metabolize anti-viral agents. An example that received considerable attention was the co-administration of the herpes zoster anti-viral medication sorivudine with the anti-tumor agent 5-fluorouracil (5-FU), which resulted in 18 patient deaths in Japan within months of release of sorivudine to market (*86*). Gut microbes hydrolyze the glycosidic linkage of sorivudine to release (*E*)-5-(2-bromovinyl)-uracil (BVU) (Figure 12A). BVU is then reduced by host dihydropyrimidine dehydrogenases (DPD), generating an allyl bromide intermediate that forms a covalent thioether adduct with a cysteine in DPD (*120*). As DPD participates in 5-FU degradation, its irreversible inhibition by BVU leads to 5-FU accumulation and toxicity (*121*). Despite the efficacy of sorivudine, its conversion to lethal microbial metabolites promptly led to recall of this drug (*122*).

Recently, pharmaceutical companies have begun to incorporate pre-clinical tests aimed at identifying and characterizing microbial metabolites of new anti-viral drug candidates (*123, 124*). In one example, the experimental hepatitis C virus drug deleobuvir was extensively metabolized upon anaerobic incubation with human fecal samples (*123*). The α,β-unsaturated carboxylic acid of deleobuvir was reduced by microbes to the metabolite CD6168. This metabolite was depleted ~9x-fold in pseudo-GF mice whose microbiotas had been depleted with antibiotics, supporting the role of gut microbes in

deleobuvir metabolism *in vivo* (*123*). Deleobuvir and CD6168 differentially inhibited or activated a number of host cytochrome P450s, which could contribute to variable, adverse drug-drug interactions (*125*).

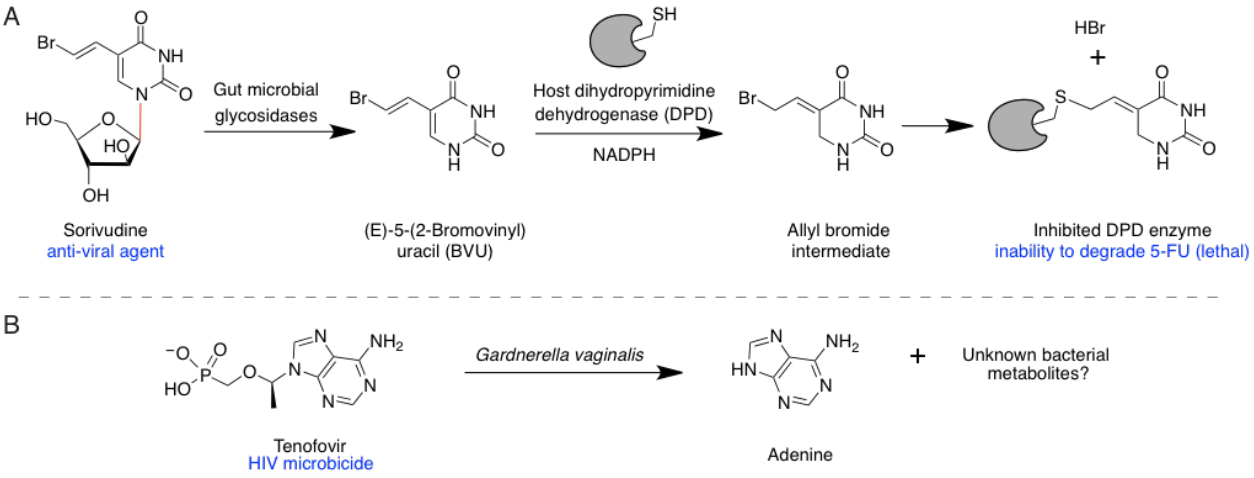


Figure 12: Gut and vaginal microbial metabolism of anti-viral agents.
(A) Gut microbial glycosidases cleave the anti-viral drug sorivudine to generate BVU. This metabolite is reduced by the host DPD enzyme, generating an intermediate which then form a covalent thioether adduct with a conserved cysteine in the enzyme. As DPD is required for detoxification of the cancer therapy 5-FU, its irreversible inhibition by sorivudine metabolites leads to toxicity and death in patients receiving both drugs. (B) Patients with higher levels of the vaginal bacterium *G. vaginalis* responded worse to prophylactic anti-HIV treatment with tenofovir. This bacterium takes up the drug and releases adenine to the media, although other metabolites have not been identified.

Another challenging but increasingly promising strategy for uncovering microbial contributions to anti-viral drug efficacy is to identify key microbes that are associated with differential outcomes in clinical cohorts. A recent study in sub-Saharan Africa found that women with *Lactobacillus*-dominated vaginal microbiotas responded ~40% better to the vaginal HIV microbicide tenofovir (taken prophylactically) than women whose microbiotas were dominated by other anaerobic microbes (*126*). Within these more complex communities of anaerobes, the bacterium *Gardnerella vaginilas* was associated with lower concentrations of active drug in the vaginal mucosa of these women. In anaerobic culture, *G. vaginalis* was shown to take up tenofovir and release adenine to culture media (Figure 12B)

(*126*), providing a link between microbial metabolism and clinically relevant variability in drug efficacy. Further work is required to identify which enzymes are involved in tenofovir metabolism, as well as which other metabolite(s) are produced. Knowledge of the bacteria and enzymes that mediate tenofovir resistance would enable the development of new therapeutic, probiotic (*127*), or vaginal microbiota remodeling approaches to improve vaginal health and increase HIV prevention.

1.3.6 Herbal supplements and traditional medicines

Gut microbial transformation of poorly absorbed constituents of herbal and traditional remedies can lead to potential health benefits or harmful side effects. There are many startling examples of the variable efficacy of traditional medicines, and this phenomenon may be due in part to the complexity of active ingredients, but also to differences in gut microbial metabolism of these treatments. Amygdalin, a mandelonitrile glycoside found in fruit pits was taken in 1960s as an alternative treatment for cancer (*128*). However, not only did amygdalin lack efficacy in clinical trials (*129*), but it was later revealed that gut microbial products of amygdalin were toxic to the host. Microbes were found to cleave the glycoside from amygdalin to release mandelonitrile, which decomposes to release benzaldehyde and cyanide (Figure 13A) (*62, 130*).

Gut bacteria also influence the activity of berberine, a plant-derived bensoisoquinolone alkaloid that is used to treat metabolic disorders including type 2 diabetes, hyperlipidemia, and obesity. Berberine has poor oral bioavailability, and gut microbial reduction to dihydroberberine greatly enhances intestinal absorption by the host (*131*). After passing into host tissues, dihydroberberine is re-oxidized and enters the blood as the active drug (Figure 13B). Co-administration of berberine and antibiotics led to lower levels of dihydroberberine and decreased drug efficacy in mice (*131*). Berberine was also shown to modulate the composition of the rat gut microbiota, particularly through a reduction in microbial diversity and enrichment of bacteria that produce SCFAs (*132*). Despite widespread use, the anti-diabetic mechanism of this supplement is not understood. However, gut microbial composition is known to

contribute to the development of metabolic disorders (*133, 134*), and further investigation is required to determine whether microbial changes are the cause or consequence of positive disease outcomes.

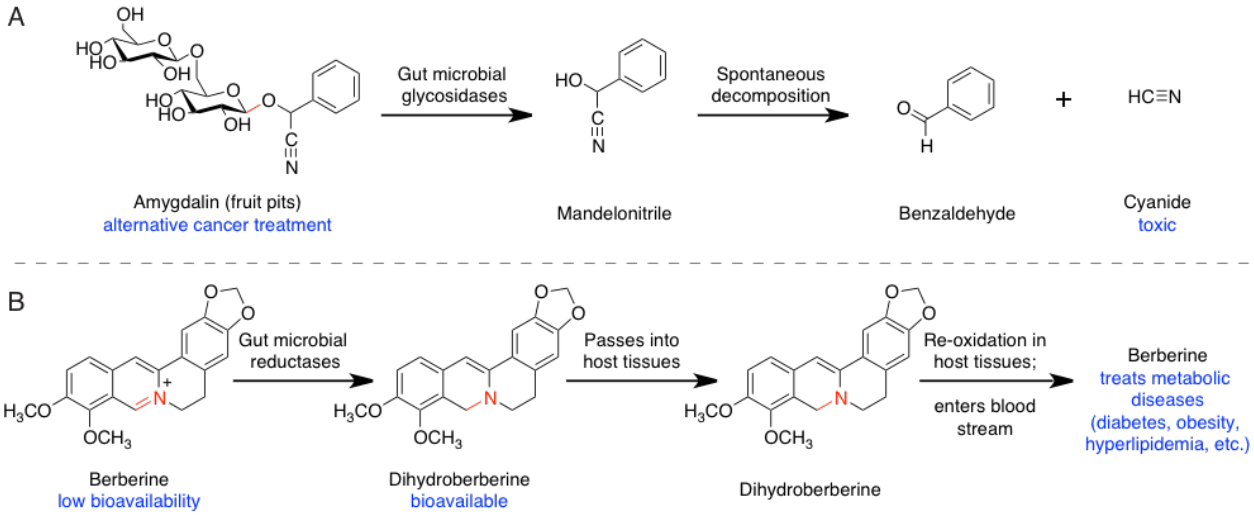


Figure 13: Gut microbial metabolism of traditional remedies and herbal supplements. (A) Gut microbial glycosidases cleave the sugars from amygdalin, a component of fruit pits that was tested as an alternative cancer treatment. The resulting mandelonitrile spontaneously decomposes to generate toxic cyanide. (B) Berberine is used in the treatment of various metabolic disorders, although it is poorly absorbed by the host. Gut microbial reduction of berberine generates a host-absorbable metabolite. Within host tissues, dihydroberberine is re-oxidized into the active compound and enters systemic circulation.

1.3.7 Cardiovascular drugs

Statins are used in the treatment of cardiovascular disease and inhibit 3-hydroxy-3-methyl-glutaryl-coenzyme A (HMG-CoA) reductase, the rate-limiting enzyme in cholesterol biosynthesis. Although statins are among the most widely prescribed drugs on the market, patient response is highly variable, and emerging evidence suggests that gut microbes and their metabolites contribute to this variability. Several exclusively microbial metabolites, including secondary bile acids and coprostanol, a bacterially reduced metabolite of cholesterol, are strong predictors of patient response to simvastatin (*135*). Additionally, in *ex vivo* incubations, human intestinal microbes could demethylate, dehydroxylate, and hydrolytically cleave simvastatin to release dimethylbutanoic acid and produce additional new

metabolites with unknown biological activity (Figure 14) (*136*). As some microbes posses an HMG-CoA reductase that is essential for growth (*137*), statins (atorvastatin, simvastatin, rosuvastatin) exhibit different degrees of antimicrobial activity against a broad range of human-associated microbes (*89*). These drugs could therefore alter gut microbiota composition and function, which may in turn affect drug and steroidal metabolite levels to ultimately impact patient response. As detailed in Section **1.4.3**, *Eggerthella lenta*, a common member of the human gut microbiota, is also known to metabolize and inactivate the cardiac drug digoxin, limiting drug efficacy in patients (*138*).

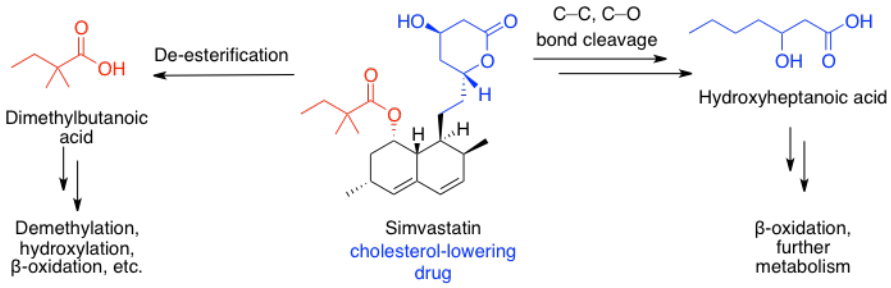


Figure 14: Gut microbial metabolism of simvastatin.
Gut microbes metabolize simvastatin using a range of transformations including de-esterification, C–C bond cleavage, demethylation, hydroxylation, and β-oxidation.

1.4 Gut microbial inactivation of the cardiac medication digoxin

Digoxin, a natural product produced by the *Digitalis* (foxglove) plant, is a medication of substantial historical and clinical importance. *Digitalis* plants have been used to treat a variety of diseases for centuries, with the first recorded medical use dating back to 1250 A.D. in Wales (*139*). In particular, the use of *Digitalis purpurea* extracts for 'dropsy' (now called congestive heart failure) in the 18th century is regarded as the start of modern clinical medicine. The active constituent of the widely used foxglove plant, digoxin, was isolated in the 1930s (*140*), and its target in the host, Na+/K+ ATPases, was elucidated 23 years later (*141*). Originally, digoxin was thought to undergo minimal metabolism in the body. However, it was later found that particular strains of the prominent gut bacterium *E. lenta* metabolize digoxin into the inactivate metabolite (20*R*)-dihydrodigoxin (Figure 15) (*74, 138, 142-144*). In modern day medicine, digoxin is used to treat cardiac failure and atrial fibrillation, although its high toxicity has lead to reduced use in recent years (*145*).

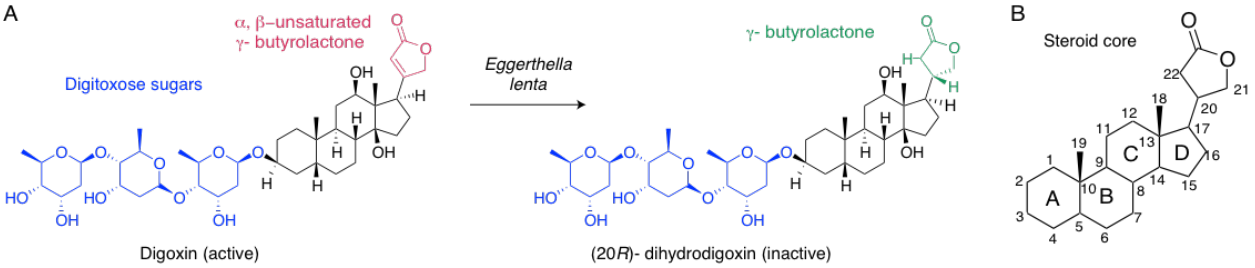


Figure 15: Gut microbial inactivation of the cardiac medication digoxin.
(A) Structural features of digoxin and its microbially reduced metabolite (20*R*)-dihydrodigoxin. (B) Steroid core ring nomenclature and numbering.

1.4.1 Pharmacology of the foxglove plant and digoxin

The *Digitalis* (foxglove) plant has a rich pharmaceutical history. In the 13th century, this plant was first described in a Welsh pharmaceutical book as an effective treatment for headaches, spasms, and epilepsy (*139*). In 1785, William Withering published *An Account of the Foxglove and its Medical Uses*

in which he described the clinical use of *Digitalis* for cardiac indications (*146*). *Digitalis purpurea* was used for this purpose as a formulation until the 1840s, when scientists Homolle and Ouevanne partially purified the active substances, which they termed "digitaline," from the plant leaves (*147*). In the early 20th century, much more was known about the pharmacology of digitaline than its chemical structure. Scientists could only crudely identify a mixture of cardiac glycosides composed of a labile sugar (glycoside) and a non-carbohydrate component (genin), and the molecular weights of the glycosylated compounds were determined with combustion analyses (*147*). In the 1930s, Sydney Smith extracted digoxin from *D. lanata* (a more potent plant than *D. pupurea*), determined its molecular weight, and established the presence of three digitoxose sugars. The structure of digoxin was solved by X-ray crystallography in 1980 (Figure 15A) (*148*) and its ^{1}H and ^{13}C NMR spectra were fully assigned in 1992 (*149*). Although digoxin and the closely related cardiac glycoside digitoxin are most extensively used in therapeutic contexts, over 30 active cardiac glycosides have been isolated from *D. purpurea* alone (*139*).

For cardiac indications, digoxin treatment is initiated by a loading dose (oral or IV; 0.5-0.75 mg), which is followed by a daily oral maintenance dose (0.125 or 0.25 mg) (FDA reference ID 3043958). Approximately 70-80% of orally ingested digoxin is absorbed in the proximal small intestine and is distributed throughout tissues with highest accumulation in the heart, kidneys, and skeletal muscles (*150*). For much of its clinical use, digoxin was thought to be largely unmetabolized in the body except for minimal acid-catalyzed deglycosylation in the stomach. However, with the advent of more sensitive detection methods, in the 1970s the gut microbial metabolite (20*R*)-dihydrodigoxin (referred to hereafter as "dihydrodigoxin") (Figure 15) was found in a substantial portion of patients as further described in Section **1.4.3**. Digoxin is predominantly filtered through the kidneys to the urine, but approximately 30% of the drug and its metabolites are excreted in the feces (*150*). Digoxin is also a substrate of P-glycoprotein (P-gp) efflux pumps, which pump drug from intestinal cells back into the gut lumen. Drug-drug interactions are thus common in patients taking digoxin as many drugs inhibit P-gp, leading to higher concentrations of digoxin in the bloodstream and toxicity (*151*). Digoxin has a very narrow therapeutic window (0.5-2 ng/mL), although toxicity may occur in some patients even within this

therapeutic range. Common symptoms of toxicity include nausea and vomiting, while more severe effects include neurological symptoms, elevated potassium levels, and cardiac arrhythmias, the latter of which can be fatal. Digoxin overdoses are treated with digoxin-specific antibody fragments (Digibind) or activated charcoal (*151*).

Owing to concerns about digoxin toxicity, the Digitalis Investigation Group conducted a randomized, double-blind, placebo-controlled clinical trial in 1997 to evaluate the safety and efficacy of digoxin. The study concluded that digoxin did not affect mortality, although it did reduce hospitalization rates for patients with chronic heart failure (*145*). Since this finding, digoxin use has steadily declined and has been replaced by other therapies such as ACE inhibitors, beta-blockers, and aldosterone antagonists. However, more recent meta-analyses of digoxin clinical trials found that the baseline conditions for patients receiving digoxin were often worse than for other patients receiving other cardiac medications, which could bias previous conclusions regarding digoxin efficacy and safety (*152*). Although digoxin is no longer used as a first line therapeutic for cardiac indications, it is still important for patients that do not respond to other drugs, as well as a growing population of patients with both cardiac failure and atrial fibrillation for whom first-line drugs (such as Ca^{2+} channel blockers) produce dangerous side effects (*152*). Although digoxin use has decreased considerably since the late 1990s (22 million annual prescriptions) (*153*), there were approximately 5 million annual prescriptions in 2010 in the US alone (*154*), demonstrating that digoxin and its microbial metabolism are still clinically relevant.

1.4.2 Mechanism of action of digoxin

Digoxin and other cardiac glycosides are potent inhibitors of Na^+/K^+ ATPases, ion transporters that are found in the plasma membranes of all animals. These transporters are important not only for maintaining the electrochemical gradient of the cell, but also for activating downstream signaling pathways resulting in diverse biological outcomes (*155, 156*). A Na^+/K^+ ATPase is made up of a catalytic α subunit that hydrolyzes ATP and a β subunit that localizes the complex to the membrane (*156*). The Na^+/K^+ ATPase exists in two conformations. In the E1 state, the enzyme binds an ATP molecule and 3

Na^+ ions with high affinity. ATP is then hydrolyzed and the γ-phosphate is transferred onto a conserved aspartate residue of the enzyme (*157*). In the phosphorylated E2 state, the affinity of the enzyme towards Na^+ ions is decreased, leading to extracellular release of these ions. Two K^+ ions then bind to the E2 complex, leading to spontaneous dephosphorylation. Binding of an additional ATP molecule then triggers intracellular release of K^+, and the catalytic cycle can proceed (*157*).

The hydrolysis of each ATP molecule drives the export of 3 Na^+ ions and the import of 2 K^+ ions against their concentration gradients, and results in a net negative charge inside the cell. The electrochemical potential generated by ATPases thus provides a gradient that enables other transporters, such as the Na^+/Ca^{2+} exchanger in the cell to function (Figure 16A) (*158*). Some downstream signaling pathways of Na^+/K^+ ATPases are initiated by changes in intracellular ion concentrations. However, a growing body of research shows that signaling pathways are mediated by different protein-protein interactions between the ATPase and additional membrane proteins, in a manner that is independent from the ion pump function of this enzyme (*158, 159*). Therefore, the expression levels and isoforms of the ATPase as well as membrane protein content can contribute to variable signaling cascades in different host tissues and disease contexts (*159*).

In humans, there are four α isoforms and three β isoforms with differential expression in the body. The α1 isoform is involved in "housekeeping" functions throughout the body, while the α2 isoform (in complex with β1) plays an important role in cardiac and smooth muscle tissue (*156*). Cardiac glycosides preferentially inhibit α2β1 Na^+/K^+ ATPases in cardiac myocytes, leading to higher intracellular Na^+ concentrations. The increased Na^+ levels impact the activity of Na^+/Ca^{2+} exchangers in the cardiac cell, leading to higher intracellular Ca^{2+} concentrations and increased muscular contraction (*156*) (Figure 16B).

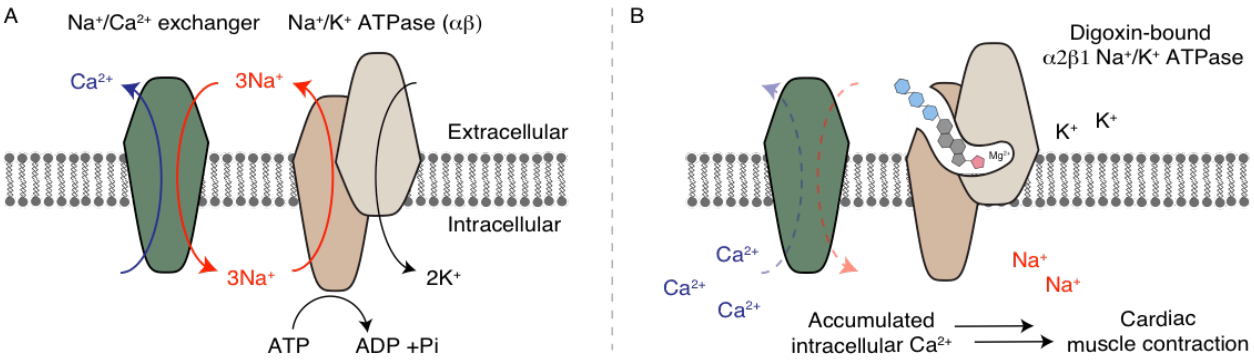


Figure 16: Na^+/K^+ ATPases function and inhibition by digoxin.
(A) Na^+/K^+ ATPases and Na^+/Ca^{2+} exchangers regulate the intracellular concentrations of the respective ions in the cell. (B) Digoxin and other cardenolides potently bind Na^+/K^+ ATPases to inhibit Na^+ export. The increased intracellular Na^+ concentrations leads to lower export of Ca^{2+} by Na^+/Ca^{2+} exchangers. High Ca^{2+} concentrations within cardiac cells lead to increased muscular contraction.

Na^+/K^+ ATPases have a hydrophobic tunnel with two metal binding sites where the steroid core and unsaturated lactone of digoxin bind (Figure 17). Additionally, they have a wide hydrophilic cavity where the first sugar group binds, while the two distal sugars extend outside of the protein (*155, 160*). Molecular docking and co-crystallographic studies were not able to identify any apparent protein-substrate binding interactions within the hydrophobic/lactone binding sites. Intriguingly, cardiac glycosides from the cardenolide family, which all contain 5-membered unsaturated lactones co-crystallize with a Mg^{2+} ion in the metal II site of the ATPase (Figure 17B). Binding to the ATPase is thought to be partly mediated by long-range (6.2 Å) electrostatic interactions between this Mg^{2+} ion and the electron rich carbonyl of the strained, unsaturated lactone of digoxin (*155, 160*). Gut microbial reduction of the lactone ring to yield (20*R*)-dihydrodigoxin changes the hybridization at C20 from sp^2 to sp^3 altering the orientation of the lactone carbonyl within the binding pocket. This results in weakened electrostatic interactions, reducing the affinity for α2β1 ATPase binding by 6.6-fold (IC_{50} digoxin: 0.33 μM; IC_{50} dihydrodigoxin 2.19 μM) (*155*) and rendering the drug inactive.

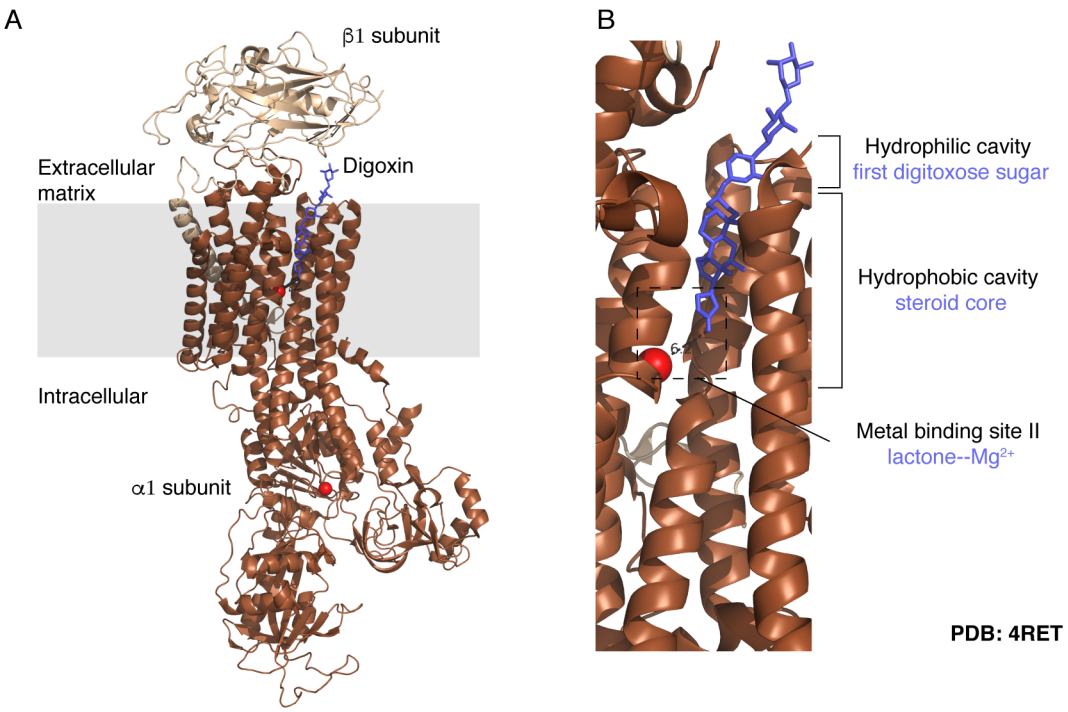


Figure 17: Binding of digoxin to Na^+/K^+ ATPases.
(A) X-ray crystal structure of the E2 α1β1 Na^+/K^+ ATPase in complex with digoxin and a Mg^{2+} ion (PDB 4RET) (*160*). (B) The unsaturated lactone moiety of digoxin is buried in a hydrophobic cavity of the ATPase and interacts with a Mg^{2+} ion in the metal II site. The steroid core and the first digitoxose sugar of digoxin are bound in the hydrophobic and hydrophilic cavities, respectively, and the two terminal sugars extend into the extracellular matrix.

While digoxin has historically been used in the treatment of cardiac conditions, drug-repurposing efforts have recently identified digoxin as an effective treatment for various cancers, HIV, and autoimmune diseases. For some of these indications, inhibition of Na^+/K^+ ATPases by digoxin activates multiple signaling cascades that impact disease symptoms or progression (*161*). As Na^+/K^+ ATPases are overexpressed in many cancers (*162*), cardiac glycosides have increasingly been assessed for their anti-cancer activities, including in four clinical trials for non-small cell lung cancer, ErbB2 positive breast cancer, metastatic breast cancer, and prostate cancer (trial identifiers: NCT01162135, NCT01887288, NCT01763931, NCT02138292) (*163, 164*). In other cases, digoxin binds entirely new targets (*165*). Although the anti-cancer mechanism of digoxin has not been fully resolved, it has been shown to potently

inhibit DNA double strand break (DSB) repair *in vitro*, possibly through direct interaction with DSB proteins (*162*). Digoxin was also shown to inhibit synthesis of HIF-1α, a transcription factor that is implicated in many cancers (*164*). Finally, both digoxin and dihydrodigoxin were shown to specifically bind and inhibit RORγt, a transcription factor involved in autoimmune disorders and inflammation, and digoxin treatment delayed the onset and reduced the severity of autoimmune disease symptoms in mice (*165*).

Although digoxin is more selective for the α2β1 complex, substantial interactions with other Na^+/K^+ ATPase isoforms including α1β1 contribute to side effects and toxicity of the drug in certain contexts (*155, 156*). In patients with heart failure, expression of α1β1 and α3β1 complexes can be reduced as much as 30-40% while α2β1 is unaffected (*166*), which could mitigate some of toxicity for these patients. However, in emerging indications of digoxin, toxicity may be more or less significant depending on which ATPase isoforms (or other digoxin targets) are relevant in the diseased tissue. Additionally, it will be critical to determine how dihydrodigoxin interacts with these new diseases targets, as significant variability in the level of digoxin metabolism is observed clinically.

1.4.3 Inactivation of digoxin by the prominent gut bacterium *Eggerthella lenta*

The first systematic studies of digoxin metabolism in humans were performed in the 1960s by Marcus and co-workers (*167, 168*). These initial efforts showed that upon IV administration of tritiated digoxin, the drug was predominantly excreted in the urine, whereas oral dosing of the drug led to substantial radioactivity (up to 30%) in the stool. Researchers further concluded that apart from minor loss of sugars, digoxin was largely unmodified in the body (*167, 168*). In the 1970s, gas chromatography-mass spectrometry (GC-MS) analysis identified a new metabolite, dihydrodigoxin, that could not be distinguished from the parent compound with previously used techniques (e.g. radioactivity assays, colorimetric assays, thin layer chromatography) (*142*). Metabolites were extracted from the urine and plasma of larger cohorts, derivatized with heptafluorobutyrate, and analyzed by GC-MS to reveal that dihydrodigoxin constituted between 10-35% of the digoxin-derived compounds in tested patients (*142,*

169). Catalytic hydrogenation of digoxin yields a 3:1 ratio of the 20*R*: 20*S* stereoisomers of dihydrodigoxin, which can be distinguished from each other using diagnostic ^{1}H NMR proton peaks of the saturated lactone at C21 and C22 (*170*). Only the 20*R* stereoisomer was identified in the urine of digoxin-metabolizing patients (*170, 171*).

To determine the site of metabolism in the body, digoxin was administered to volunteers through IV, poorly absorbed tablet formulations, or direct injection into the small intestine (*144, 172, 173*). These studies found that highest levels of dihydrodigoxin were produced upon direct injection into the jejunum, followed by oral tablet, and IV. Furthermore, co-administration of digoxin with antibiotics led to depletion or elimination of dihydrodigoxin from the body, and raised serum levels of digoxin, suggesting that gut bacteria in the distal small intestine were responsible for digoxin reduction (*144, 172, 173*). To determine which bacteria could perform this transformation, Dobkin and co-workers isolated and screened over 400 anaerobic bacterial strains from the stool of human volunteers that excreted high levels of dihydrodigoxin. Only two isolates could metabolize digoxin (*138*), both of which were *Eubacterium lentum* (later renamed *Eggerthella lenta*), a Gram-positive anaerobic species of the *Actinobacteria* phylum. *E. lenta* is a common inhabitant of the human gut that was isolated in 1938 from a rectal cancer biopsy (*174*). Closely related *Eubacterium* species did not metabolize digoxin, and only a subset (18/28) of *E. lenta* strains could reduce digoxin to 20*R* dihydrodigoxin. Digoxin metabolism was inhibited by addition of L-arginine, a substrate that is preferentially used by *E. lenta* to generate ATP (*138, 171*). *E. lenta* strains were also present in the patients that did not excrete dihydrodigoxin, demonstrating that the presence of this bacterium was not sufficient for explaining or predicting metabolism in patients (*138*). Additionally, Mathan, Alam, and co-workers found significant variability in dihydrodigoxin production in patients living in rural (South India, Bangladesh) vs. urban (South India, Bangladesh, and New York City) communities, although digoxin-reducing bacteria were present at similar percentages of both groups (*175, 176*). These studies highlight that environmental factors including geography and diet can modulate digoxin-metabolism by *E. lenta*.

While the role of *E. lenta* in digoxin metabolism has been appreciated for decades, challenges in growing the organism and a lack of genetic tools hampered efforts to understand this transformation. Recently, Turnbaugh and co-workers performed RNA sequencing (RNA-Seq) on the digoxin-metabolizing *E. lenta* DSM2243 strain that was grown in the presence or absence of digoxin and identified a digoxin-inducible gene cluster, which they called the cardiac glycoside reductase (*cgr*) operon (*74*). Comparative genomics of the DSM2243 strain with non-metabolizers revealed that this gene cluster was only present in the digoxin-inactivating *E. lenta* isolate (*74*). The *cgr* operon from this organism encodes two proteins that resemble bacterial reductases involved in anaerobic respiration (Figure 18A). Bioinformatic analyses suggest that a membrane-associated cytochrome (Cgr1) transfers electrons through a series of hemes to a predicted flavin-dependent reductase (Cgr2) that converts digoxin to dihydrodigoxin (Figure 18B). The *cgr* genes correlated with digoxin reduction by *E. lenta* in culture, and could also be used as biomarkers to stratify human fecal samples into low- or high- digoxin metabolizing communities. Additionally, L-arginine, which is the preferred energy source for *E. lenta* was found to decrease *cgr* gene expression and digoxin metabolism in culture. This knowledge successfully informed the design of a high protein dietary intervention that reduced digoxin metabolism in GF mice that were mono-associated with cgr^+ *E. lenta* (*74*).

Although substantial progress has been made in characterizing the gut bacteria, genes, and factors that correlate with or influence digoxin metabolism, several key questions remain, including: (1) How is digoxin metabolism distributed among *E. lenta* strains and the general human population? (2) Are the Cgr protein responsible for digoxin reduction, and what is their catalytic mechanism? (3) What are the evolutionary origins or physiological impacts of digoxin metabolism for gut bacteria? In the following Chapters, I describe a variety of bioinformatic, (meta)genomic, microbiological, and biochemical investigations that we have employed to gain a mechanistic understanding of this clinically important gut microbial transformation.

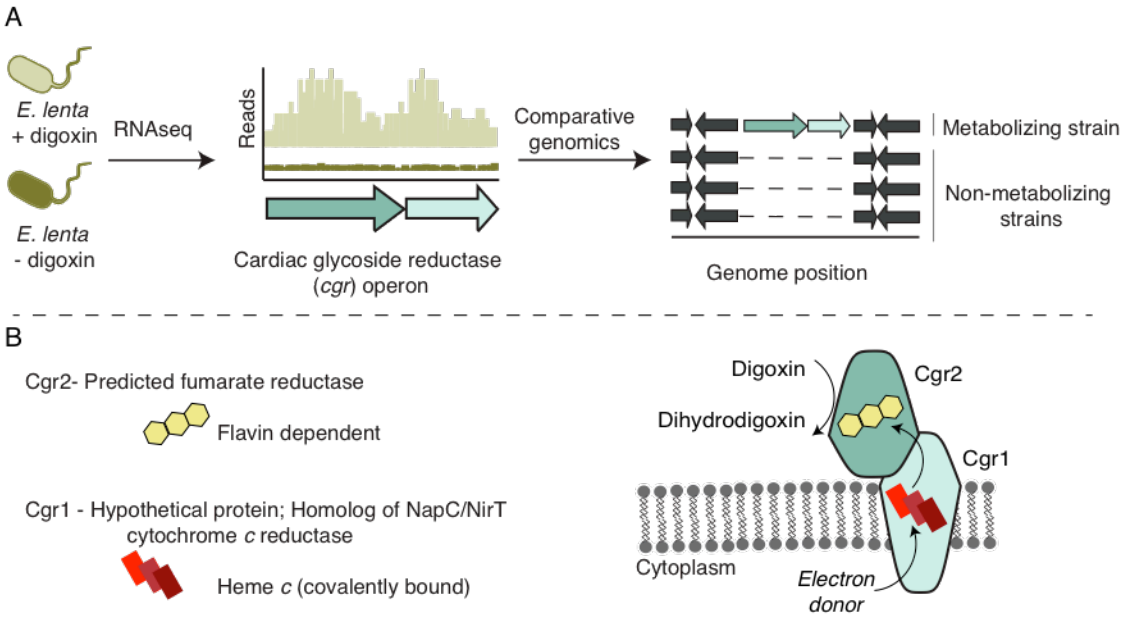


Figure 18: Identification of the cardiac glycoside reductase (*cgr*) operon in a digoxin-metabolizing strain of *E. lenta*.
(A) RNA-seq revealed two genes that were upregulated in a digoxin-metabolizing strain of *E. lenta* grown in the presence of the drug. The cardiac glycoside reductase (*cgr*) genes were only present in the metabolizing strain. (B) Initial bioinformatic hypothesis for digoxin metabolism. Cgr1 is a predicted membrane-bound, cytochrome *c* enzyme. Cgr2 is a predicted flavin-dependent fumarate reductase. Cgr1 may transfer electrons from an electron donor to the flavin of Cgr2 to ultimately reduce digoxin. See Section **3.7** for more details.

1.5 References

1. A. H. Moeller *et al.*, Cospeciation of gut microbiota with hominids. *Nature* **353**, 380-382 (2016).

2. C. G. Buffie *et al.*, Precision microbiome reconstitution restores bile acid mediated resistance to *Clostridium difficile*. *Nature* **517**, 205-208 (2015).

3. G. Sharon *et al.*, Specialized metabolites from the microbiome in health and disease. *Cell Metabolism* **20**, 719-730 (2014).

4. A. Koh, F. De Vadder, P. Kovatcheva-Datchary, F. Backhed, From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites. *Cell* **165**, 1332-1345 (2016).

5. T. Sousa *et al.*, The gastrointestinal microbiota as a site for the biotransformation of drugs. *Int J Pharm* **363**, 1-25 (2008).

6. N. Koppel, V. M. Rekdal, E. P. Balskus, Chemical transformation of xenobiotics by the human gut microbiota. *Science* **356**, eaag2770 (2017).

7. H. Danielsson, B. Gustafsson, On serum-cholesterol levels and neutral fecal sterols in germ-free rats - bile acids and steroids. *Arch Biochem Biophys* **83**, 482-485 (1959).

8. J. C. Lagier *et al.*, Microbial culturomics: paradigm shift in the human gut microbiome study. *Clin Microbiol Infect* **18**, 1185-1193 (2012).

9. L. Ma *et al.*, Gene-targeted microfluidic cultivation validated by isolation of a gut bacterium listed in Human Microbiome Project's Most Wanted taxa. *Proc Natl Acad Sci U S A* **111**, 9768-9773 (2014).

10. A. A. Fodor *et al.*, The "most wanted" taxa from the human microbiome for whole genome sequencing. *PLoS One* **7**, e41294 (2012).

11. Human Microbiome Project Consortium, Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207-214 (2012).

12. Human Microbiome Jumpstart Reference Strains Consortium, A catalog of reference genomes from the human microbiome. *Science* **328**, 994-999 (2010).

13. J. Li *et al.*, An integrated catalog of reference genes in the human gut microbiome. *Nat Biotechnol* **32**, 834-841 (2014).

14. J. Qin *et al.*, A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59-65 (2010).

15. J. C. Clemente, L. K. Ursell, L. W. Parfrey, R. Knight, The impact of the gut microbiota on human health: an integrative view. *Cell* **148**, 1258-1270 (2012).

16. F. Karlsson, V. Tremaroli, J. Nielsen, F. Backhed, Assessing the human gut microbiota in metabolic diseases. *Diabetes* **62**, 3341-3349 (2013).

17. N. Koppel, E. P. Balskus, Exploring and understanding the biochemical diversity of the human microbiota. *Cell Chemical Biology* **23**, 19-30 (2016).

18. J. Aron-Wisnewsky, J. Dore, K. Clement, The importance of the gut microbiota after bariatric surgery. *Nat Rev Gastroenterol Hepatol* **9**, 590-598 (2012).

19. R. Sender, S. Fuchs, R. Milo, Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol* **14**, e1002533 (2016).

20. P. B. Eckburg *et al.*, Diversity of the human intestinal microbial flora. *Science* **308**, 1635-1638 (2005).

21. W. R. Wikoff *et al.*, Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc Natl Acad Sci U S A* **106**, 3698-3703 (2009).

22. R. Pastelin-Palacios *et al.*, Subversion of inate and adaptive immune activation induced by structurally modified lipopolysaccharide from *Salmonella typhimurium*. *Immunology* **133**, 469-481 (2011).

23. M. Deng *et al.*, Lipopolysaccharide clearance, bacterial clearance, and systemic inflammatory responses are regulated by cell type-specific functions of TLR4 during sepsis. *J. Immunol* **190**, 5152-5160 (2013).

24. F. Y. Avci, D. L. Kasper, How bacterial carbohydrates influence the adaptive immune system. *Annu. Rev. Immunol.* **28**, 107-130 (2010).

25. S. K. Mazmanian, C. H. Liu, A. O. Tzianabos, D. L. Kasper, An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell* **122**, 107-118 (2005).

26. E. A. Smith, G. T. Macfarlane, Formation of phenolic and indolic compounds by anaerobic bacteria in the human large intestine. *Microb Ecol* **33**, 180-188 (1997).

27. P. A. Aronov *et al.*, Colonic contribution to uremic solutes. *J Am Soc Nephrol* **22**, 1769-1776 (2011).

28. T. A. Clayton, D. Baker, J. C. Lindon, J. R. Everett, J. K. Nicholson, Pharmacometabonomic identification of a significant host-microbiome metabolic interaction affecting human drug metabolism. *Proc Natl Acad Sci U S A* **106**, 14728-14733 (2009).

29. J. M. Ridlon, D. J. Kang, P. B. Hylemon, Bile salt biotransformation by human intestinal bacteria. *J. Lipid Res.* **47**, 241-259 (2006).

30. J. M. Ridlon, D. J. Kang, P. B. Hylemon, J. S. Bajaj, Bile acids and the gut microbiome. *Curr. Opin. Gastroenterol.* **30**, 332-338 (2014).

31. D. R. Donohoe *et al.*, The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell Metab* **13**, 517-526 (2011).

32. G. den Besten *et al.*, The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J. Lipid Res.* **54**, 2325-2340 (2013).

33. N. Reichardt *et al.*, Phylogenetic distribution of three pathways for propionate production within the human gut microbiota. *ISME J* **8**, 1323-1335 (2014).

34. M. Waldecker, T. Kautenburger, H. Daumann, C. Busch, D. Schrenk, Inhibition of histone-deacetylase activity by short-chain fatty acids and some polyphenol metabolites formed in the colon. *J Nutr Biochem* **19**, 587-593 (2008).

35. Y. Xiong *et al.*, Short-chain fatty acids stimulate leptin production in adipocytes through the G protein-coupled receptor GPR41. *Proc Natl Acad Sci U S A* **101**, 1045-1050 (2004).

36. K. M. Maslowski *et al.*, Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. *Nature* **461**, 1282-1286 (2009).

37. M. Thangaraju *et al.*, GPR109A is a G-protein-coupled receptor for the bacterial fermentation product butyrate and functions as a tumor suppressor in colon. *Cancer Res* **69**, 2826-2832 (2009).

38. B. Wang, L. Hu, T. Siahaan, *Drug delivery: principles and applications*. (John Wiley & Sons Inc., Hoboken, New Jersey, ed. Second edition., 2016), pp. xxii, 697 pages.

39. L. Amin, P-glycoprotein Inhibition for Optimal Drug Delivery. *Drug Target Insights* **7**, 27-34 (2013).

40. P. Spanogiannopoulos, E. N. Bess, R. N. Carmody, P. J. Turnbaugh, The microbial pharmacists within us: a metagenomic view of xenobiotic metabolism. *Nat Rev Microbiol* **14**, 273-287 (2016).

41. P. Gerard, Metabolism of cholesterol and bile acids by the gut microbiota. *Pathogens* **3**, 14-24 (2013).

42. S. Craciun, E. P. Balskus, Microbial conversion of choline to trimethylamine requires a glycyl radical enzyme. *Proc Natl Acad Sci U S A* **109**, 21307-21312 (2012).

43. A. El Kaoutari, F. Armougom, J. I. Gordon, D. Raoult, B. Henrissat, The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nat Rev Microbiol* **11**, 497-504 (2013).

44. B. J. Levin *et al.*, A prominent glycyl radical enzyme in human gut microbiomes metabolizes trans-4-hydroxy-L-proline. *Science* **355**, eaai8386 (2017).

45. A. Martinez-del Campo *et al.*, Characterization and detection of a widely distributed gene cluster that predicts anaerobic choline utilization by human gut bacteria. *MBio* **6**, (2015).

46. J. K. Nicholson, E. Holmes, I. D. Wilson, Gut microorganisms, mammalian metabolism and personalized health care. *Nat Rev Microbiol* **3**, 431-438 (2005).

47. P. H. Degnan, N. A. Barry, K. C. Mok, M. E. Taga, A. L. Goodman, Human gut microbes use multiple transporters to distinguish vitamin B(1)(2) analogs and compete in the gut. *Cell Host Microbe* **15**, 47-57 (2014).

48. R. J. Linhardt, P. M. Galliher, C. L. Cooney, Polysaccharide lyases. *Applied Biochemistry and Biotechnology* **12**, 135-176 (1986).

49. A. Ryan *et al.*, 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 (2014).

50. K. Ellrott, L. Jaroszewski, W. Li, J. C. Wooley, A. Godzik, Expansion of the protein repertoire in newly explored environments: human gut microbiome specific protein families. *PLoS Comput Biol* **6**, e1000798 (2010).

51. K. Kurokawa *et al.*, Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes. *DNA Res* **14**, 169-181 (2007).

52. P. Jancova, P. Anzenbacher, E. Anzenbacherova, Phase II drug metabolizing enzymes. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub* **154**, 103-116 (2010).

53. U. M. Zanger, M. Schwab, Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol Ther* **138**, 103-141 (2013).

54. J. Wang, V. Yadav, A. L. Smart, S. Tajiri, A. W. Basit, Stability of peptide drugs in the colon. *Eur J Pharm Sci* **78**, 31-36 (2015).

55. H. Tozaki *et al.*, Degradation of insulin and calcitonin and their protection by various protease inhibitors in rat caecal contents: Implications in peptide delivery to the colon. *J Pharm Pharmacol* **49**, 164-168 (1997).

56. W. W. Kallemeijn, M. D. Witte, T. Wennekes, J. M. Aerts, Mechanism-based inhibitors of glycosidases: design and applications. *Adv Carbohydr Chem Biochem* **71**, 297-338 (2014).

57. B. D. Wallace *et al.*, Structure and inhibition of microbiome beta-glucuronidases essential to the alleviation of cancer drug toxicity. *Chem Biol* **22**, 1238-1249 (2015).

58. J. E. Ulmer *et al.*, Characterization of glycosaminoglycan (GAG) sulfatases from the human gut symbiont *Bacteroides thetaiotaomicron* reveals the first GAG-specific bacterial endosulfatase. *Journal of Biological Chemistry* **289**, 24289-24303 (2014).

59. G. Lukatela *et al.*, Crystal structure of human arylsulfatase A: the aldehyde function and the metal ion at the active site suggest a novel mechanism for sulfate ester hydrolysis. *Biochemistry* **37**, 3654-3664 (1998).

60. K. S. Saitta *et al.*, Bacterial beta-glucuronidase inhibition protects mice against enteropathy induced by indomethacin, ketoprofen or diclofenac: mode of action and pharmacokinetics. *Xenobiotica* **44**, 28-35 (2014).

61. B. D. Wallace *et al.*, Alleviating cancer drug toxicity by inhibiting a bacterial enzyme. *Science* **330**, 831-835 (2010).

62. J. H. Carter, M. A. Mclafferty, P. Goldman, Role of the gastrointestinal microflora in amygdalin (Laetrile)-induced cyanide toxicity. *Biochem Pharmacol* **29**, 301-304 (1980).

63. T. Niimura, T. Tokieda, T. Yamaha, Partial-purification and some properties of cyclamate sulfamatase. *J Biochem-Tokyo* **75**, 407-417 (1974).

64. A. J. Cooper *et al.*, Cysteine S—conjugate beta-lyases: important roles in the metabolism of naturally occurring sulfur and selenium-containing compounds, xenobiotics and anticancer agents. *Amino Acids* **41**, 7-27 (2011).

65. S. P. Claus, H. Guillou, S. Ellero-Simatos, The gut microbiota: a major player in the toxicity of environmental pollutants? *Npj Biofilms And Microbiomes* **2**, 16003 (2016).

66. J. E. Bakke, J. A. Gustafsson, B. E. Gustafsson, Metabolism of propachlor by the germfree rat. *Science* **210**, 433-435 (1980).

67. S. Suzuki, H. Tomisawa, S. Ichihara, H. Fukazawa, M. Tateishi, A C-S Bond-Cleavage Enzyme of Cysteine Conjugates in Intestinal Microorganisms. *Biochem Pharmacol* **31**, 2137-2140 (1982).

68. I. Rossol, A. Puhler, The *Corynebacterium glutamicum aecD* gene encodes a C—S lyase with alpha, beta-elimination activity that degrades aminoethylcysteine. *J Bacteriol* **174**, 2968-2977 (1992).

69. H. Laue, M. Friedrich, J. Ruff, A. M. Cook, Dissimilatory sulfite reductase (desulfoviridin) of the taurine-degrading, non-sulfate-reducing bacterium *Bilophila wadsworthia* RZATAU contains a fused DsrB-DsrD subunit. *J Bacteriol* **183**, 1727-1733 (2001).

70. S. C. Lee, A. G. Renwick, Sulphoxide reduction by rat intestinal flora and by *Escherichia coli in vitro*. *Biochem Pharmacol* **49**, 1567-1576 (1995).

71. F. Rafii, J. B. Sutherland, E. B. Hansen, C. E. Cerniglia, Reduction of nitrazepam by *Clostridium leptum*, a nitroreductase-producing bacterium isolated from the human intestinal tract. *Clin Infect Dis* **25**, S121-S122 (1997).

72. F. Rafii, C. E. Cerniglia, Reduction of azo dyes and nitroaromatic compounds by bacterial enzymes from the human intestinal tract. *Environmental Health Perspectives* **103**, 17-19 (1995).

73. A. Ryan, E. Kaplan, N. Laurieri, E. Lowe, E. Sim, Activation of nitrofurazone by azoreductases: multiple activities in one enzyme. *Sci Rep* **1**, 63 (2011).

74. H. J. Haiser *et al.*, Predicting and manipulating cardiac drug inactivation by the human gut bacterium *Eggerthella lenta*. *Science* **341**, 295-298 (2013).

75. K. Lavrijsen *et al.*, Reduction of the prodrug loperamide oxide to its active drug loperamide in the gut of rats, dogs, and humans. *Drug Metab Dispos* **23**, 354-362 (1995).

76. M. A. Peppercorn, P. Goldman, Role of intestinal bacteria in metabolism of salicylazosulfapyridine. *Journal of Pharmacology and Experimental Therapeutics* **181**, 555-562 (1972).

77. R. Bentley, T. G. Chasteen, Microbial Methylation of Metalloids: Arsenic, Antimony, and Bismuth. *Microbiol Mol Biol Rev* **66**, 250-271 (2002).

78. T. Kumano, E. Fujiki, Y. Hashimoto, A. Kobayashi, Discovery of a sesamin-metabolizing icroorganism and a new enzyme. *Proc Natl Acad Sci U S A* **113**, 9087-9092 (2016).

79. S. Burapan, M. Kim, J. Han, Demethylation of polymethoxyflavones by human gut bacterium, *Blautia sp.* MRG-PMF1. *J Agric Food Chem* **65**, 1620-1629 (2017).

80. T. Ticak, D. J. Kountz, K. E. Girosky, J. A. Krzycki, D. J. Ferguson, Jr., A nonpyrrolysine member of the widely distributed trimethylamine methyltransferase family is a glycine betaine methyltransferase. *Proc Natl Acad Sci U S A* **111**, E4668-4676 (2014).

81. C. Delomenie *et al.*, Identification and functional characterization of arylamine *N*-acetyltransferases in eubacteria: evidence for highly selective acetylation of 5-aminosalicylic acid. *J Bacteriol* **183**, 3417-3427 (2001).

82. S. E. Rothenberg *et al.*, The role of gut microbiota in fetal methylmercury exposure: Insights from a pilot study. *Toxicol Lett* **242**, 60-67 (2016).

83. W. Buckel, B. T. Golding, Radical enzymes in anaerobes. *Annu Rev Microbiol* **60**, 27-49 (2006).

84. C. A. Kolmeder *et al.*, Comparative metaproteomics and diversity analysis of human intestinal microbiota testifies for its temporal stability and expression of core functions. *PLoS One* **7**, e29913 (2012).

85. S. Takeno, Y. Hirano, A. Kitamura, T. Sakai, Comparative developmental toxicity and metabolism of nitrazepam in rats and mice. *Toxicol Appl Pharmacol* **121**, 233-238 (1993).

86. H. Okuda *et al.*, Lethal drug interactions of sorivudine, a new antiviral drug, with oral 5-fluorouracil prodrugs. *Drug Metabolism and Disposition* **25**, 269-273 (1997).

87. K. Shiizaki, M. Kawanishi, T. Yagi, Microbial metabolites of omeprazole activate murine aryl hydrocarbon receptor in vitro and in vivo. *Drug Metab Dispos* **42**, 1690-1697 (2014).

88. N. R. Shin *et al.*, An increase in the *Akkermansia spp.* population induced by metformin treatment improves glucose homeostasis in diet-induced obese mice. *Gut* **63**, 727-735 (2014).

89. M. Masadeh, N. Mhaidat, K. Alzoubi, S. Al-Azzam, Z. Alnasser, Antibacterial activity of statins: a comparative study of atorvastatin, simvastatin, and rosuvastatin. *Ann Clin Microbiol Antimicrob* **11**, 13 (2012).

90. G. Falony *et al.*, Population-level analysis of gut microbiome variation. *Science* **352**, 560-564 (2016).

91. B. Routy *et al.*, Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science*, 10.1126/science.aan3706 (2017).

92. V. Gopalakrishnan *et al.*, Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science*, 10.1126/science.aan1423 (2017).

93. K. Takasuna *et al.*, Inhibition of intestinal microflora beta-glucuronidase modifies the distribution of the active metabolite of the antitumor agent, irinotecan hydrochloride (CPT-11) in rats. *Cancer Chemother Pharmacol* **42**, 280-286 (1998).

94. H. Matsui *et al.*, The pathophysiology of non-steroidal anti-inflammatory drug (NSAID)-induced mucosal injuries in stomach and small intestine. *J Clin Biochem Nutr* **48**, 107-111 (2011).

95. C. Sostres, C. J. Gargallo, A. Lanas, Nonsteroidal anti-inflammatory drugs and upper and lower gastrointestinal mucosal damage. *Arthritis Research Therapy* **15**, 1-8 (2013).

96. K. Watanabe, S. Yamashita, K. Furuno, H. Kawasaki, Y. Gomita, Metabolism of omeprazole by gut flora in rats. *Journal of Pharmaceutical Sciences* **84**, 516-517 (1995).

97. S. Janarthanan, I. Ditah, D. G. Adler, M. N. Ehrinpreis, *Clostridium difficile*-associated diarrhea and proton pump inhibitor therapy: a meta-analysis. *Am J Gastroenterol* **107**, 1001-1010 (2012).

98. C. T. Seto, P. Jeraldo, R. Orenstein, N. Chia, J. K. DiBaise, Prolonged use of a proton pump inhibitor reduces microbial diversity: implications for *Clostridium difficile* susceptibility. *Microbiome* **2**, 1-11 (2014).

99. U. Klotz, K. Maier, C. Fischer, K. Heinkel, Therapeutic efficacy of sulfasalazine and its metabolites in patients with ulcerative-colitis and Crohns-disease. *New England Journal of Medicine* **303**, 1499-1502 (1980).

100. T. Sousa *et al.*, On the colonic bacterial metabolism of azo-bonded prodrugs of 5-aminosalicylic acid. *J Pharm Sci* **103**, 3171-3175 (2014).

101. R. A. van Hogezand *et al.*, Bacterial acetylation of 5-aminosalicylic acid in faecal suspensions cultured under aerobic and anaerobic conditions. *Eur J Clin Pharmacol* **43**, 189-192 (1992).

102. K. Westphal, S. Leschner, J. Jablonska, H. Loessner, S. Weiss, Containment of tumor-colonizing bacteria by host neutrophils. *Cancer Res* **68**, 2952-2960 (2008).

103. D. L. Mager, Bacteria and cancer: cause, coincidence or cure? A review. *J Transl Med* **4**, 14 (2006).

104. P. Lehouritis *et al.*, Local bacteria affect the efficacy of chemotherapeutic drugs. *Sci Rep* **5**, 14554 (2015).

105. L. T. Geller *et al.*, Potential role of intratumor bacteria in mediating tumor resistance to the chemotherapeutic drug gemcitabine. *Science* **357**, 1156-1160 (2017).

106. M. Vetizou *et al.*, Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science* **350**, 1079-1084 (2015).

107. A. Sivan *et al.*, Commensal *Bifidobacterium* promotes antitumor immunity and facilitates anti-PD-L1 efficacy. *Science* **350**, 1084-1089 (2015).

108. N. Iida *et al.*, Commensal bacteria control cancer response to therapy by modulating the tumor microenvironment. *Science* **342**, 967-970 (2013).

109. S. Clement, P. Dandona, J. G. Still, G. Kosutic, Oral modified insulin (HIM2) in patients with type 1 diabetes mellitus: results from a phase I/II clinical trial. *Metabolism* **53**, 54-58 (2004).

110. R. P. Anderson, T. J. Butt, V. S. Chadwick, Hepatobiliary excretion of bacterial formyl-methionyl peptides in rat. Structure activity studies. *Dig Dis Sci* **37**, 248-256 (1992).

111. I. Sasaki *et al.*, Metabolism of azetirelin, a new thyrotropin-releasing hormone (TRH) analogue, by intestinal microorganisms. *Pharm Res* **14**, 1004-1007 (1997).

112. D. B. Calne *et al.*, Idiopathic Parkinsonism treated with an extracerebral decarboxylase inhibitor in combination with levodopa. *Br Med J* **3**, 729-732 (1971).

113. J. Bergmark *et al.*, Decarboxylation of orally administered L-dopa in the human digestive tract. *Naunyn Schmiedebergs Arch Pharmacol* **272**, 437-440 (1972).

114. B. R. Goldin, M. A. Peppercorn, P. Goldman, Contributions of host and intestinal microflora in the metabolism of L-dopa by the rat. *J Pharmacol Exp Ther* **186**, 160-166 (1973).

115. J. G. Nutt, N. H. Holford, The response to levodopa in Parkinson's disease: imposing pharmacological law and order. *Ann Neurol* **39**, 561-573 (1996).

116. K. Zhang, Y. Ni, Tyrosine decarboxylase from *Lactobacillus brevis*: Soluble expression and characterization. *Protein Expres Purif* **94**, 33-39 (2014).

117. J. Gidai, N. Acs, F. Banhidy, A. E. Czeizel, Congenital abnormalities in children of 43 pregnant women who attempted suicide with large doses of nitrazepam. *Pharmacoepidemiol Drug Saf* **19**, 175-182 (2010).

118. S. Kitamura, K. Sugihara, M. Kuwasako, K. Tatsumi, The role of mammalian intestinal bacteria in the reductive metabolism of zonisamide. *J Pharm Pharmacol* **49**, 253-256 (1997).

119. G. Sharon, T. R. Sampson, D. H. Geschwind, S. K. Mazmanian, The central nervous system and the gut microbiome. *Cell* **167**, 915-932 (2016).

120. T. Nishiyama *et al.*, Mechanism-based inactivation of human dihydropyrimidine dehydrogenase by (E)-5-(2-bromovinyl)uracil in the presence of NADPH. *Mol Pharmacol* **57**, 899-905 (2000).

121. H. Okuda, K. Ogura, A. Kato, H. Takubo, T. Watabe, A possible mechanism of eighteen patient deaths caused by interactions of sorivudine, a new antiviral drug, with oral 5-fluorouracil prodrugs. *J Pharmacol Exp Ther* **287**, 791-799 (1998).

122. H. Gurdon, Japan bans shingles drug after deaths. *BMJ* **309**, 627 (1994).

123. M. McCabe *et al.*, Defining the role of gut bacteria in the metabolism of deleobuvir: *In vitro* and *in vivo* studies. *Drug Metab Dispos* **43**, 1612-1618 (2015).

124. Y. Li, J. Xu, W. G. Lai, A. Whitcher-Johnstone, D. J. Tweedie, Metabolic switching of BILR 355 in the presence of ritonavir. II. Uncovering novel contributions by gut bacteria and aldehyde oxidase. *Drug Metab Dispos* **40**, 1130-1137 (2012).

125. R. S. Sane *et al.*, Contribution of major metabolites toward complex drug-drug interactions of deleobuvir: *In vitro* predictions and *in vivo* outcomes. *Drug Metab Dispos* **44**, 466-475 (2016).

126. N. R. Klatt *et al.*, Vaginal bacteria modify HIV tenofovir microbicide efficacy in African women. *Science* **356**, 938-945 (2017).

127. S. Borges, J. Silva, P. Teixeira, The role of *Lactobacilli* and probiotics in maintaining vaginal health. *Arch Gynecol Obstet* **289**, 479-489 (2014).

128. A. S. Relman, Closing the books on Laetrile. *N Engl J Med* **306**, 236 (1982).

129. C. G. Moertel *et al.*, A clinical trial of amygdalin (Laetrile) in the treatment of human cancer. *N Engl J Med* **306**, 201-206 (1982).

130. G. J. Strugala, A. G. Rauws, R. Elbers, Intestinal first pass metabolism of amygdalin in the rat *in vitro*. *Biochem Pharmacol* **35**, 2123-2128 (1986).

131. R. Feng *et al.*, Transforming berberine into its intestine-absorbable form by the gut microbiota. *Sci Rep-Uk* **5**, 121555 (2015).

132. X. Zhang *et al.*, Modulation of gut microbiota by berberine and metformin during the treatment of high-fat diet-induced obesity in rats. *Sci Rep* **5**, 14405 (2015).

133. P. D. Cani *et al.*, Changes in gut microbiota control metabolic endotoxemia-induced inflammation in high-fat diet-induced obesity and diabetes in mice. *Diabetes* **57**, 1470-1481 (2008).

134. J. Qin *et al.*, A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**, 55-60 (2012).

135. R. Kaddurah-Daouk *et al.*, Enteric microbiome metabolites correlate with response to simvastatin treatment. *PLoS One* **6**, e25482 (2011).

136. A. M. Aura *et al.*, Drug metabolome of the simvastatin formed by human intestinal microbiota *in vitro*. *Mol Biosyst* **7**, 437-446 (2011).

137. Y. Matsumoto *et al.*, A critical role of mevalonate for peptidoglycan synthesis in *Staphylococcus aureus*. *Sci Rep* **6**, 22894 (2016).

138. J. F. Dobkin, J. R. Saha, V. P. Butler, H. C. Neu, J. Lindenbaum, Digoxin-inactivating bacteria: Identification in human gut flora. *Science* **220**, 325-327 (1983).

139. K. Greeff, T. Akera, *Cardiac glycosides*. Handbook of experimental pharmacology (Springer-Verlag, Berlin ; New York, 1981).

140. S. Smith, Digoxin, a new digitalis glucoside. *J. Chem. Soc.*, 508-510 (1930).

141. H. J. Schatzmann, Cardiac glycosides as inhibitors of active potassium and sodium transport by erythrocyte membrane. *Helv Physiol Pharmacol Acta* **11**, 346-354 (1953).

142. E. Watson, D. R. Clark, S. R. Kalman, Identification by gas chromatography-mass spectroscopy of dihydrodigoxin- a metabolite of digoxin in man. *Journal of Pharmacology and Experimental Therapeutics* **184**, 424-431 (1973).

143. U. Peters, L. C. Falk, S. M. Kalman, Digoxin metabolism in patients. *Arch Intern Med* **138**, 1074-1076 (1978).

144. J. Lindenbaum, D. G. Rund, V. P. Butler, D. Tseeng, J. R. Saha, Inactivation of digoxin by the gut flora: Reversal by antibiotic therapy. *New England Journal of Medicine* **305**, 789-794 (1981).

145. Digitalis Investigation Group, The effect of digoxin on mortality and morbidity in patients with heart failure. *N Engl J Med* **336**, 525-533 (1997).

146. W. Withering, *An account of the foxglove*. (Printed by M. Swinney for G. G. J. and J. Robinson, London, Birmingham Eng., 1785).

147. R. C. Elderfield, The chemistry of the cardiac glycosides. *Chemical Reviews* **17**, 187-249 (1935).

148. K. Go, G. Kartha, J. P. Chen, Structure of digoxin. *Acta Cryst* **B36**, 1811-1819 (1980).

149. A. E. Aulabaugh *et al.*, The conformational behaviour of the cardiac glycoside digoxin as indicated by NMR spectroscopy and molecular dynamics calculations. *Carbohydr Res* **230**, 201-212 (1992).

150. E. Iisalo, Clinical pharmacokinetics of digoxin. *Clin Pharmacokinet* **2**, 16 (1977).

151. M. Pincus, Management of digoxin toxicity. *Aust Prescr* **39**, 18-20 (2016).

152. O. J. Ziff *et al.*, Safety and efficacy of digoxin: systematic review and meta-analysis of observational and controlled trial data. *BMJ* **351**, h4451 (2015).

153. A. Ahmed, R. M. Allman, J. F. Delong, Inappropriate use of digoxin in older hospitalized heart failure patients *The Journals of Gerontology* **2**, M138-M143 (2002).

154. I. See, N. Shehab, S. R. Kegler, S. R. Laskar, D. S. Budnitz, Emergency department visits and hospitalizations for digoxin toxicity: United States, 2005 to 2010. *Circ Heart Fail* **7**, 28-34 (2014).

155. K. M. Weigand *et al.*, Na(+),K(+)-ATPase isoform selectivity for digitalis-like compounds is determined by two amino acids in the first extracellular loop. *Chem Res Toxicol* **27**, 2082-2092 (2014).

156. A. Katz *et al.*, Selectivity of digitalis glycosides for Iisoforms of human Na,K-ATPase. *Journal of Biological Chemistry* **285**, 19582-19592 (2010).

157. G. Scheiner-Bobis, The sodium pump. Its molecular properties and mechanics of ion transport. *Eur J Biochem* **269**, 2424-2433 (2002).

158. A. Aperia, E. E. Akkuratov, J. M. Fontana, H. Brismar, Na+-K+-ATPase, a new class of plasma membrane receptors. *Am J Physiol Cell Physiol* **310**, C491-495 (2016).

159. Z. J. Xie, A. Askari, Na+/K+-ATPase as a signal transducer. *European Journal of Biochemistry* **269**, 2434-2439 (2002).

160. M. Laursen, J. L. Gregersen, L. Yatime, P. Nissen, N. U. Fedosova, Structures and characterization of digoxin- and bufalin-bound Na+,K+-ATPase compared with the ouabain-bound complex. *Proc Natl Acad Sci U S A* **112**, 1755-1760 (2015).

161. R. W. Wong, A. Balachandran, M. A. Ostrowski, A. Cochrane, Digoxin suppresses HIV-1 replication by altering viral RNA processing. *PLoS Pathog* **9**, e1003241 (2013).

162. Y. V. Surovtseva *et al.*, Characterization of cardiac glycoside natural products as potent inhibitors of DNA double-strand break repair by a whole-cell double immunofluorescence assay. *J Am Chem Soc* **138**, 3844-3855 (2016).

163. F. Kayali, M. A. Janjua, D. A. Laber, D. Miller, G. Kloecker, Phase II trial of second-line erlotinib and digoxin for nonsmall cell lung cancer (NSCLC). *Open Access Journal of Clinical Trials* **3**, 9-13 (2011).

164. J. Lin *et al.*, A pilot phase II Study of digoxin in patients with recurrent prostate cancer as evident by a rising PSA. *Am J Cancer Ther Pharmacol* **2**, 21-32 (2014).

165. J. R. Huh *et al.*, Digoxin and its derivatives suppress TH17 cell differentiation by antagonizing ROR γ t activity. *Nature* **472**, 486-490 (2011).

166. J. Muller-Ehmsen, J. Wang, R. H. Schwinger, A. A. McDonough, Region specific regulation of sodium pump isoform and Na,Ca-exchanger expression in the failing human heart-right atrium vs left ventricle. *Cell Mol Biol* **47**, 373-381 (2001).

167. F. I. Marcus, G. J. Kapadia, G. G. Kapadia, The metabolism of digoxin in normal subjects. *J Pharmacol Exp Ther* **145**, 203-209 (1964).

168. F. I. Marcus, L. Burkhalter, C. Cuccia, J. Pavlovich, G. G. Kapadia, Administration of tritiated digoxin with and without a loading dose. *Circulation* **34**, 865-874 (1966).

169. D. R. Clark, S. R. Kalman, Dihydrodigoxin: a common metabolite of digoxin in man. *Drug Metab Dispos* **2**, 148-150 (1974).

170. R. H. Reuning, T. A. Shephard, B. E. Morrison, H. N. Bockbrader, Formation of [20*R*]-dihydrodigoxin from digoxin in humans. *Drug Metab Dispos* **13**, 51-57 (1985).

171. L. W. Robertson, A. Chandrasekaran, R. H. Reuning, J. Hul, B. D. Rawal, Reduction of digoxin to 20*R*-dihydrodigoxin by cultures of *Eubacterium lentum*. *Applied and Environmental Microbiology* **51**, 1300-1303 (1986).

172. J. Lindenbaum, D. Tse-Eng, V. P. Butler, D. G. Rund, Urinary excretion of reduced metabolites of digoxin. *American Journal of Medicine* **71**, 67-74 (1981).

173. J. O. Magnusson, B. Bergdahl, Metabolism of digoxin and absorption site. *Br. J. Clin. Pharmac.* **14**, 284-285 (1982).

174. W. E. C. Moore, E. P. Cato, L. V. Holdeman, Eubacterium lenturn (Eggerth) Prevot 1938:Emendation of Description and Designation of the Neotype Strain. *INTERNATIONAL JOURNAL of SYSTEMATIC BACTERIOLOGY* **21**, 299-303 (1971).

175. A. N. Alam, J. R. Saha, J. F. Dobkin, J. Lindenbaum, Interethnic variation in the metabolic inactivation of digoxin by the gut flora. *Gastroenterology* **95**, 117-123 (1988).

176. V. I. Mathan, J. Wiederman, J. F. Dobkin, J. Lindenbaum, Geographic differences in digoxin inactivation, a metabolic activity of the human anaerobic gut flora. *Gut* **30**, 971-977 (1989).

Chapter 2: Cgr2 is a unique enzyme encoded in the *cgr* operon that is responsible for digoxin metabolism

2.1 Bioinformatic analyses and heterologous expression link the *cgr* operon to digoxin reduction

Multiple lines of evidence suggested that the enzymes encoded by the *cgr* operon (Cgr1 and Cgr2) are responsible for digoxin reduction in *E. lenta* DSM2243 (*1*). We thus turned to heterologous expression, culturing, and bioinformatic analyses to definitively assess whether these proteins were involved in drug inactivation and to probe the mechanism of digoxin reduction. As only one sequenced *cgr*$^{+}$ *E. lenta* strain was available prior to our work, we also sought to expand the collection of digoxin-metabolizing bacteria. By performing coupled sequencing and culturing experiments, we aimed to understand the sequence conservation and distribution of *cgr* genes among bacterial isolates and validate their utility as biomarkers for digoxin metabolism. Finally, we wanted to probe the distribution and prevalence of *cgr*$^{+}$ *E. lenta* strains within the human population to understand the clinical implications and scope of digoxin metabolism.

In the addition to the observations that the *cgr* genes are upregulated in the presence of digoxin and are found exclusively in a metabolizing strain of *E. lenta*, bioinformatic analyses further support the hypothesis that the Cgr proteins directly catalyze digoxin reduction. Cgr1 is a putative membrane-anchored protein that is predicted to harbor at least three covalently bound heme groups using three CXXCH motifs and a less common CXXXCH motif (1x) (*2*) (Figure 19). Basic local alignment search tool (BLAST) analyses revealed that Cgr1 belongs to the cytochrom_*c*3_c2 superfamily (Pfam 14537), and shares up to 47% sequence identity to the NapC/NirT (NrfH) protein family. Members of this protein family transfer electrons from the quinone pool in the cell membrane to extracytoplasmic or periplasmic proteins to facilitate reduction of terminal electron acceptors such as nitrite and sulfite (*3, 4*). The similarity between Cgr1 and these proteins points to a putative role in electron transfer. We also identified a close homolog of Cgr1, Elen_2528, in *E. lenta* DSM2243 (92% amino acid identity; PSI-BLAST; e-value <141) as well as in *E. lenta* 1_3_56FAA, a non-metabolizing strain of *E. lenta* (91% amino acid identity; PSI-BLAST; e-value <138). The presence of this highly similar protein in both metabolizing and non-metabolizing strains provides further support that Cgr1 is involved in a more general function (electron transfer) rather than directly mediating reduction of digoxin.

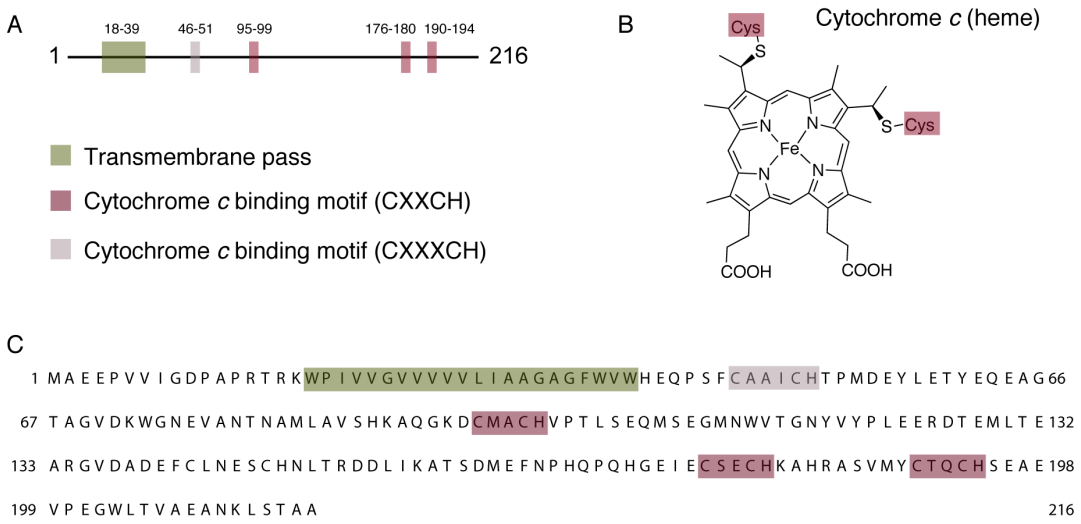
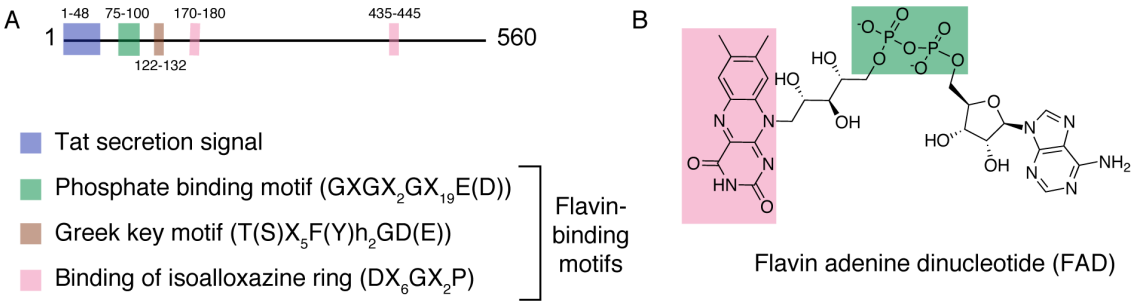


Figure 19: Annotation of Cgr1.
(A) Cgr1 protein architecture with relevant motifs highlighted and corresponding amino acid numbering shown above. (B) Structure of cytochrome *c* highlighting covalent cysteine ligation sites. (C) Full amino acid sequence and annotation of Cgr1.

On the other hand, Cgr2 is unique in *E. lenta* DSM2243; the closest homolog of Cgr2 in this strain (identified through PSI-BLAST) had <28% amino acid identity. Cgr2 harbors a conserved RRXFLK motif and a single pass transmembrane region, and is thus predicted to undergo secretion to the periplasm through the twin arginine translocation (Tat) pathway. This predicted localization suggests a capacity of Cgr2 to metabolize extracellular substrates. Additionally, Cgr2 is predicted to bind flavin adenine dinucleotide (FAD). Of four common FAD-binding folds, Cgr2 most closely resembles the glutathione reductase (GR) superfamily, which adopts a Rossmann fold topology ($\beta_1\alpha_1\beta_2\alpha_2\beta_3$) (*5*). Cgr2 contains all three required FAD binding motifs of this GR enzyme class, with some modifications. It contains two copies of the DX_6GX_2P motif, which makes polar contacts with the isoalloxazine ring of FAD. In addition, the 'greek key' motif of Cgr2 (TX_6YhGD) differs slightly from the canonical motif in GRs ($T(S)X_5F(Y)hhGD$) (*5*) (Figure 20).

Cgr2 is a member of FAD-dependent fumarate reductases (Pfam 00890; Interpro 003953/027477), but lacks close homologs. The most similar characterized homologs of Cgr2 in the NCBI non-redundant database include flavocytochrome *c* fumarate reductases from *Shewanella*

frigidimarina and *S. putrefaciens* MR-1 (25-26% amino acid identity) as well as ketosteroid dehydrogenases from *Rhodococcus jostii Rha1* and *R. erythropolis* (25% amino acid identity). Additional remote protein homologs of Cgr2 were identified through the HHPred prediction tool (https://toolkit.tuebingen.mpg.de/#/tools/hhpred) (*6*), including fumarate reductases from *Shewanella* species (PDB: 1D4D, 1Q08, 1Y0P), ketosteroid dehydrogenases from *Rhodococcus* species (PDB: 4C3X, 4AT2), and L-aspartate oxidases (PDB: 1KNR, 2E5V). All of these enzymes share a similar mechanism involving hydride transfer to/from a flavin cofactor and a bound substrate, and enzyme-mediated proton transfer. While Cgr2 is likely to share this general mechanism, it is highly divergent and lacks most of the required catalytic residues of these enzymes (see Section **2.2**). These results suggest that Cgr2 likely adopts a distinct fold or utilizes different catalytic residues to perform substrate reduction.



C

1 MEYGKCRGIERGMGRRDFLKAATLLGATAAGAGMLAGCAPKSASEAQAQTAPAATGGLDPADVDWK 66
67 YETDVVIVGSGSGGTCAAIEAAEAGADVVVFEKDKAMYGGNSALCGGYMLAAGWSTQEEITGYAGD 132
133 TGEAFANQMLRWSQGLGNQDMIREACLRSGEAVDWMMDTGRTYEGASPLPPVWSCGDTEADVVPRS 198
199 VYNHNAYGATEGHMATLKKRAESLSNIEIEMGCEVAHILKNAEGSVIGVQLADGSFAKARKGVVMA 264
265 CASVDNNLEMSKDLGLMQNVWGLTLEGAGLLAPGNPDMDSNTGDGVRMLREIGAELCMQQAVCMND 330
331 SIYVGGISDWGMSEILGKDVNIHDSSNIDAILVDKTGRRFCQDDAEWGYVMHECAQAAWKQGFTPD 396
397 DPTTGYIFYVYDATGAPFFEMKGHTPDTCDTTFSADSVDGLAEFIGCDPTALASEVERWNSFCEAG 462
463 LDADFGRRANMAPIATPPFYCDVVRPGPMGTFAGAKSNVEAEIIGLDGNPIPRLYGAGCIIGGNVS 528
529 GAFYFGCGWSITNTVVWGREAGRNVAALEPWE 560

Figure 20: Annotation of Cgr2.
(A) Cgr2 protein architecture with relevant motifs highlighted, with amino acid numbering shown above. (B) Structure of flavin adenine dinucleotide (FAD). Highlighted regions correspond to interaction sites with residues in FAD-binding motifs. (C) Full amino acid sequence and annotation of Cgr2.

To definitively test if the Cgr proteins are sufficient for digoxin reduction, we transformed *cgr1* and *cgr2* from *E. lenta* DSM2243 into the model Actinobacterium *Rhodococcus erythropolis* L88 (*7-9*), discussed in more detail in Section **3.1**. After inducing protein expression (15 °C), cultures were incubated with digoxin for 2 days (28 °C), and dihydrodigoxin production was quantified by liquid chromatography tandem mass spectrometry (LC-MS/MS) (Figure 21A). All of the Cgr2 expressing strains showed a significant increase in dihydrodigoxin levels relative to empty vector controls. In contrast, no activity was observed for the strain expressing only Cgr1, and no overexpression of this protein could be detected in clarified lysates, cell pellets, or partially purified membrane fractions (Figure 21B). These results show that Cgr2 is sufficient for digoxin reduction in *R. erythropolis*. Our general model is that Cgr1 and Cgr2, which resemble bacterial terminal electron reductases, likely form a complex that mediates electron transfer from membrane-associated electron donors to ultimately reduce the α,β-unsaturated butyrolactone of digoxin (Figure 18B). Additionally, the ability of Cgr2 to produce dihydrodigoxin in *R. erythropolis* suggests that endogenous redox active proteins and/or metabolites in this heterologous host may fulfill the putative function of Cgr1 as an electron donor.

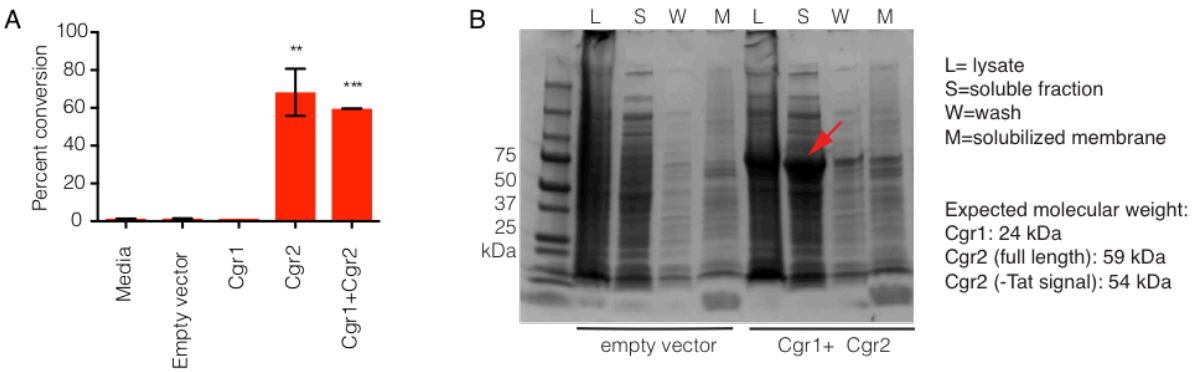


Figure 21: Heterologous expression confirms that Cgr2 is sufficient for digoxin reduction. (A) Whole cell assays in *R. erythropolis* overexpressing Cgr proteins (15 °C, overnight) and incubated with digoxin for 2 days at 28 °C. All strains overexpressing Cgr2 displayed gain-of-function digoxin reduction capacity. Data represents mean ± standard error of the mean (SEM) (n=3). Asterisks denote statistical significance by Student's t test (** p <0.01, *** p <0.001). (B) Representative gel showing overexpression of soluble Cgr2. No Cgr1 overexpression was observed in lysates or in membrane fractions.

2.2 Cgr2 is a divergent flavin dependent reductase

After confirming that Cgr2 is a digoxin reductase, we aimed to identify amino acid residues or sequence motifs that may be important for its activity or fold, and to more systematically assess its relationship to other reductases. We thus constructed a sequence similarity network (SSN) using the top 5,000 most similar Cgr2 homologs from the UniProtKB protein database (http://efi.igb.illinois.edu/efi-est/) (*10*). Within the network, there were seven enzymes that had been biochemically characterized (UniProtKB ID: Q07WU7, Q9Z4P0, 8CVD0, P71864), biochemically and structurally characterized (PDB ID: 1D4D, 1E39), or genetically characterized (Q7D5C1) (*11-18*). At all thresholds at which Cgr2 remained connected to other protein sequences, there was no sub-clustering of any of the characterized enzymes within the SSN (Figure 22A). At alignment thresholds that separated characterized enzymes into putative isofunctional clusters (e value < 10^{-130}), Cgr2 was always present as a 'singleton' with no links to other sequenced proteins (Figure 22B).

We validated our SSN by aligning the sequences of characterized enzymes in the network with additional co-clustered sequences to assess conservation of essential active site residues. All of the characterized enzymes within the SSN share a similar catalytic mechanism, which involves hydride transfer to or from a flavin cofactor and substrate, and proton transfer through a network of conserved proton shuttling residues. Inspection of multiple sequence alignments revealed that Cgr2 contains mutations in the majority of essential catalytic residues found in the characterized enzymes (Figure 23). Cgr2 lacked 6/7 conserved active site residues required for fumarate reductase activity (*11, 19*). Additionally a second predicted fumarate reductase associated with the *cgr* gene cluster (Cac4) (see Section **2.3**) lacked all active site residues required for this enzyme family, suggesting that both enzymes are misannotated as fumarate reductases (Figure 24).

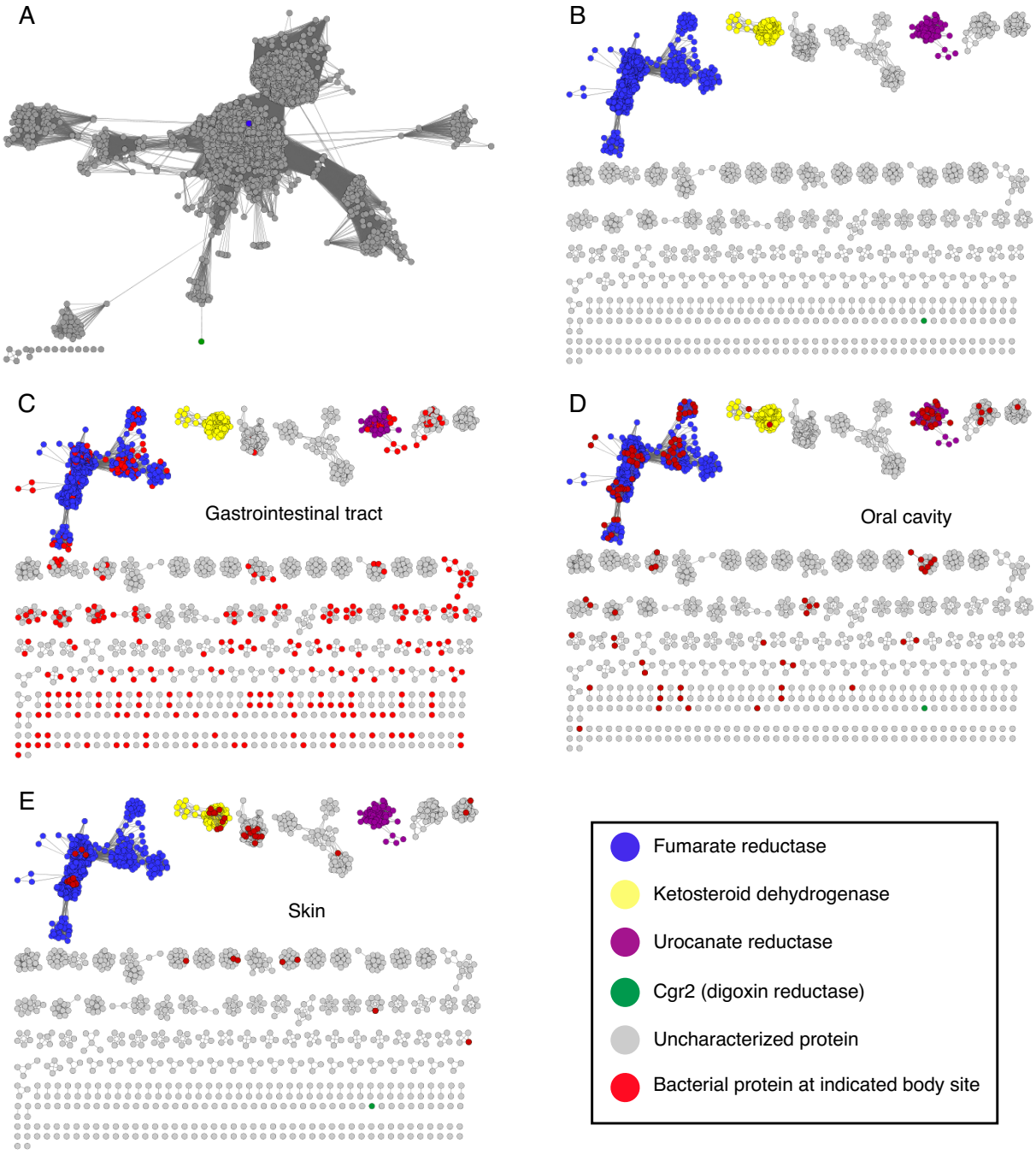


Figure 22: Sequence similarity network (SSN) analysis reveals that Cgr2 is a highly distinct member of a large reductase family that is widespread in human associated microbes. SSNs were constructed using the top 5000 most similar proteins to Cgr2 from the UniprotKB database. (A) Nodes represent proteins with 100% sequence identity and edges connect nodes whose BLAST pairwise similarity score is less than 10^{-50}. (B) SSN in which nodes represent proteins with 100% sequence identity and edges connect nodes whose BLAST pairwise similarity score is less than 10^{-130}. Putatively isofunctional clusters are colored according to biochemical functions. Nodes that encode proteins from bacteria found in the (C) gastrointestinal tract, (D) oral cavity, and (E) skin are colored in red.

Cgr2 also contained mutations in 4/5 conserved active site residues found in urocanate reductases (Figure 25) (*14*). Finally, Cgr2 lacked 3/5 active site residues of ketosteroid dehydrogenases (Figure 26), although two active site amino acids involved in proton abstraction (Tyr) and substrate binding and activation (Gly backbone) were conserved (Y532, G536) (*15, 20*). In whole cell assays in *R. erythropolis*, the putative Cgr2 active site mutant Y532F retained activity, suggesting that Y532 does not serve the same essential role in Cgr2 as it does in ketosteroid dehydrogenases. Mutation of the second putative active site residue G536 led to a loss of activity (Figure 23E), although it is possible that this residue plays a structural rather than catalytic role in Cgr2. Taken together, these observations demonstrate that Cgr2 is distinct from characterized reductases and may adopt a different fold or utilize a distinct set of residues to catalyze the reduction of cardenolides.

Our SSN also revealed that this class of reductase enzymes is often found in human-associated bacteria from diverse body sites (GI tract, oral cavity, skin) (Figure 22C-E). Reductases are particularly widespread in organisms from the GI tract, demonstrating the importance of this type of metabolism in anaerobic environments. Out of the 266 distinct clusters in the SSN (e value < 10^{-130}), 4 clusters had been biochemically validated (2% of total network) and 113 clusters of unknown function (42% of total network) were present in at least one gut bacterial isolate (Figure 22C). These results further highlight the underappreciated chemical and functional diversity of this class of enzymes, especially in the context of the human gut microbiota.

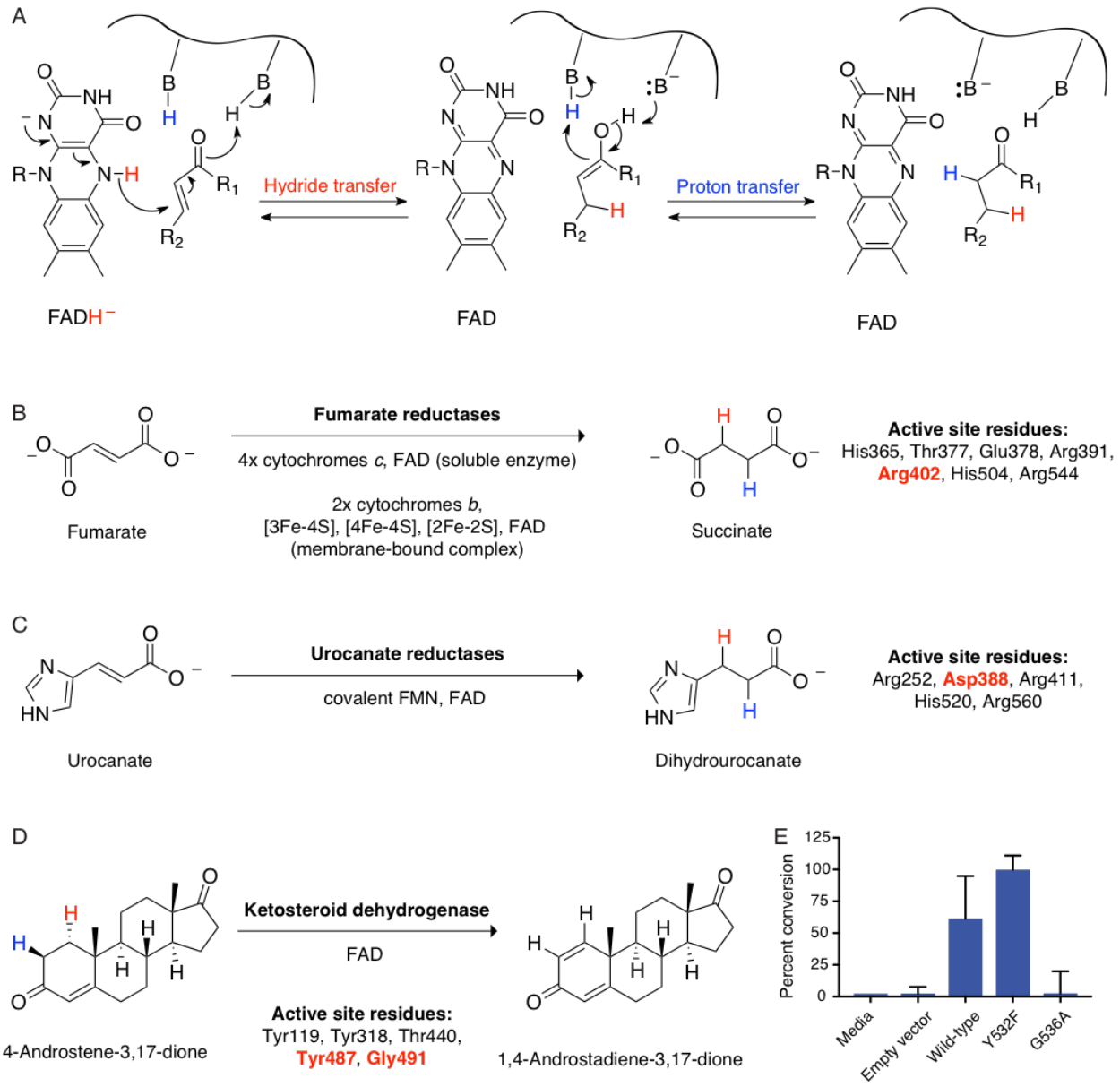


Figure 23: Cgr2 is a distinct flavoprotein reductase.
(A) General mechanism of catalysis by Cgr2 homologs. Cgr2 lacks most of the conserved active site residues found in the most similar related enzymes, including (B) 6/7 residues utilized by fumarate reductases, (C) 4/5 residues utilized by urocanate reductases, and (D) 3/5 residues utilized by ketosteroid dehydrogenases. Active site residues are shown with numbering based on *S. putrefaciens* fumarate reductase, *S. oneidensis* MR-1 urocanate reductase, and *R. erythropolis* SQ1 ketosteroid dehydrogenase. Residues shown in red were conserved in Cgr2. (E) Two residues involved in substrate binding and activation in ketosteroid dehydrogenases are conserved in Cgr2 (Y532, G536). Whole cell assays in *R. erythropolis* overexpressing putative active site mutants in Cgr2 showed that Y532 was not essential for Cgr2 activity towards digoxin. Data represents mean ± SEM (n = 3).

Figure 24: Multiple sequence alignments of fumarate reductases.
UniProtKB ID numbers are shown in parentheses. Active site residues (marked with an asterisk) were conserved in characterized fumarate reductases and clustered proteins from the SSN, but not in Cgr2 and another predicted fumarate reductase (Cac4) associated with the *cgr* gene cluster.

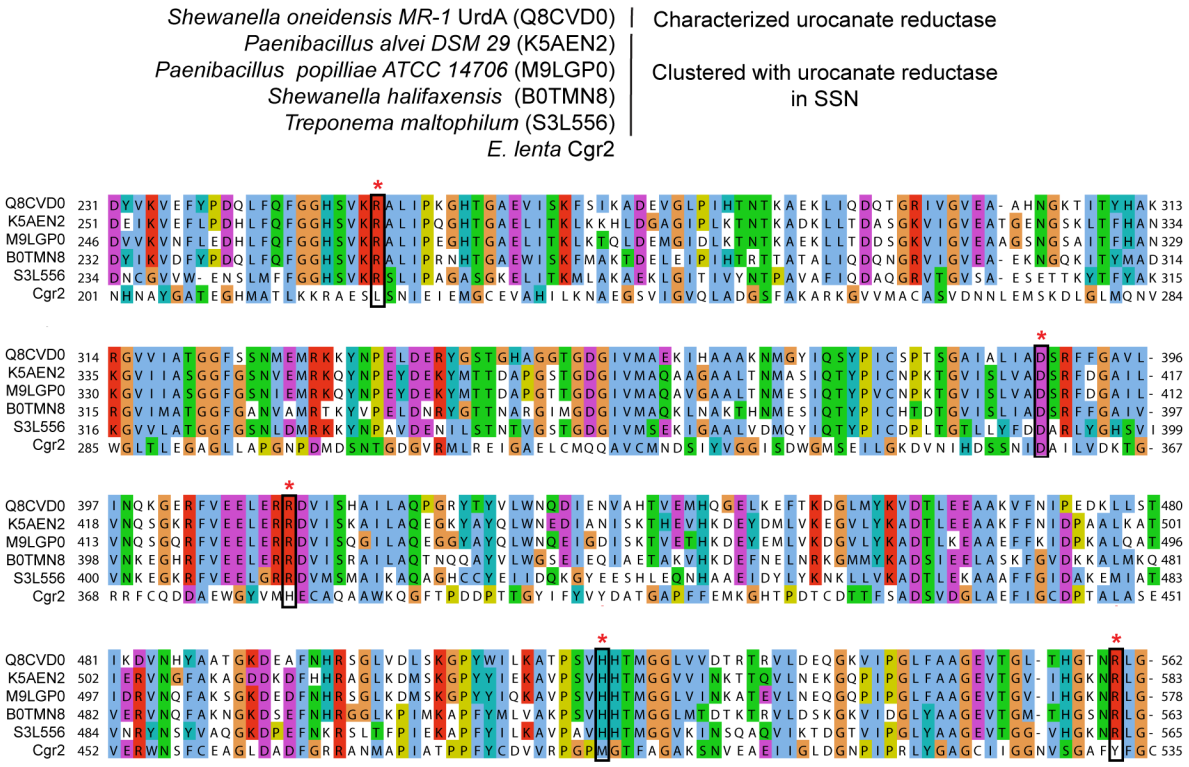


Figure 25: Multiple sequence alignments of urocanate reductases.
UniProtKB ID numbers are shown in parentheses. Active site residues were conserved in characterized urocanate reductases and clustered proteins from the SSN but not in Cgr2.

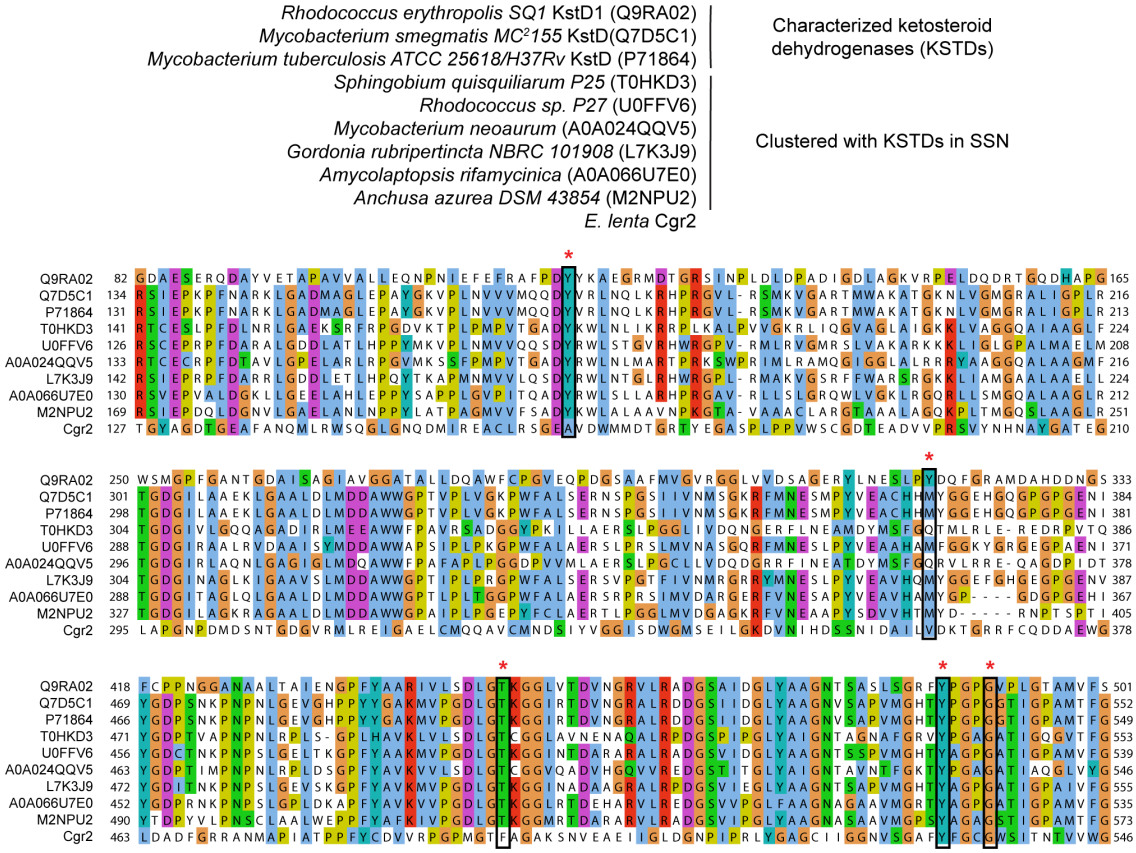


Figure 26: Multiple sequence alignments of ketosteroid dehydrogenases.
UniProtKB ID numbers are shown in parentheses. Active site residues were conserved in characterized ketosteroid dehydrogenases and clustered proteins from the SSN. Two residues involved in substrate binding and activation were conserved in Cgr2 (Y532, G536).

2.3 An expanded *cgr* operon is present in digoxin metabolizing *E. lenta* isolates

Despite the high incidence of dihydrodigoxin production observed clinically, only one digoxin-reducing strain, *E. lenta* DSM2243, had been isolated and sequenced (*21*). While the *cgr* operon was predictive of digoxin metabolism in this isolate (*1*), it remained unclear whether digoxin metabolism was exclusively associated with the *cgr* operon in patients, or if other pathways existed. Moreover, the presence of additional *cgr* sequences could help identify important catalytic residues. We thus aimed to obtain and test a library of human-associated isolates to assess the broader relevance of digoxin metabolism in humans. Dr. Jordan Bisanz from the Turnbaugh laboratory at UCSF curated, isolated, and sequenced a collection of 25 *E. lenta* isolates and closely related *Coriobacteriaceae* species (Figure 27A) obtained from 22 individuals in six countries (France, USA, UK, Canada, Spain, China). These strains were isolated from stool, colonic biopsies, rectal tumors, and blood samples of both healthy subjects and patients. Whole genome sequencing revealed the presence of seven additional cgr^+ strains within the collection. We grew each strain in the presence of digoxin and used LC-MS/MS to quantify the biotransformation to dihydrodigoxin. Indeed, all cgr^+ strains reduced digoxin while cgr^- strains did not, confirming that the *cgr* operon is predictive of digoxin metabolism (Figure 27B). *Cgr* presence and digoxin metabolism did not correlate with *E. lenta* phylogeny (Figure 27C).

Unexpectedly, the *cgr* operon was part of an expanded 8-gene cluster that was highly conserved in all digoxin-metabolizing strains (>98% at amino acid level). The expanded operon included the *cgr* operon at the 5' terminus and 6 additional genes that formed a "cgr associated gene cluster" (*cac*) at the 3' terminus. Reanalysis of RNA-seq data (*1*) revealed that within this expanded operon, only *cgr1, cgr2,* and the LuxR type transcriptional regulator *cac3* were upregulated in *E. lenta* strains grown in the presence of digoxin (156x, 174x, 1.5x, respectively). The remaining *cac* genes were not significantly upregulated, implying that despite their co-occurrence with the *cgr* operon, these genes are not involved digoxin metabolism. Cac4, a second annotated flavin-dependent fumarate reductase within the expanded gene cluster lacks all known catalytic and binding residues for this enzyme class (Figure 28). Like Cgr2, Cac4 has a predicted FAD-binding Rossmann fold and Tat secretion signal, but shares only 23% sequence

identity to Cgr2. Thus, it is likely to metabolize a different substrate than fumarate and digoxin. It is conceivable that Cac4 could also interact with Cgr1 to obtain the necessary reducing equivalents for a reductive transformation. The expanded operon also contains three hypothetical proteins, one of which (Cac6) is homologous to stomatin, prohibitin, flotillin, HflK/C (SPFH) proteins that are often associated with lipid rafts or functional microdomains in bacteria (*23*). These microdomains are comprised of different lipids or modified cholesterol and sterol compounds (*23*), and could perhaps interact with steroidal substrates or products of the *cgr* operon. However, additional work is required to understand the activities of these enzymes and whether they interact with the Cgr proteins.

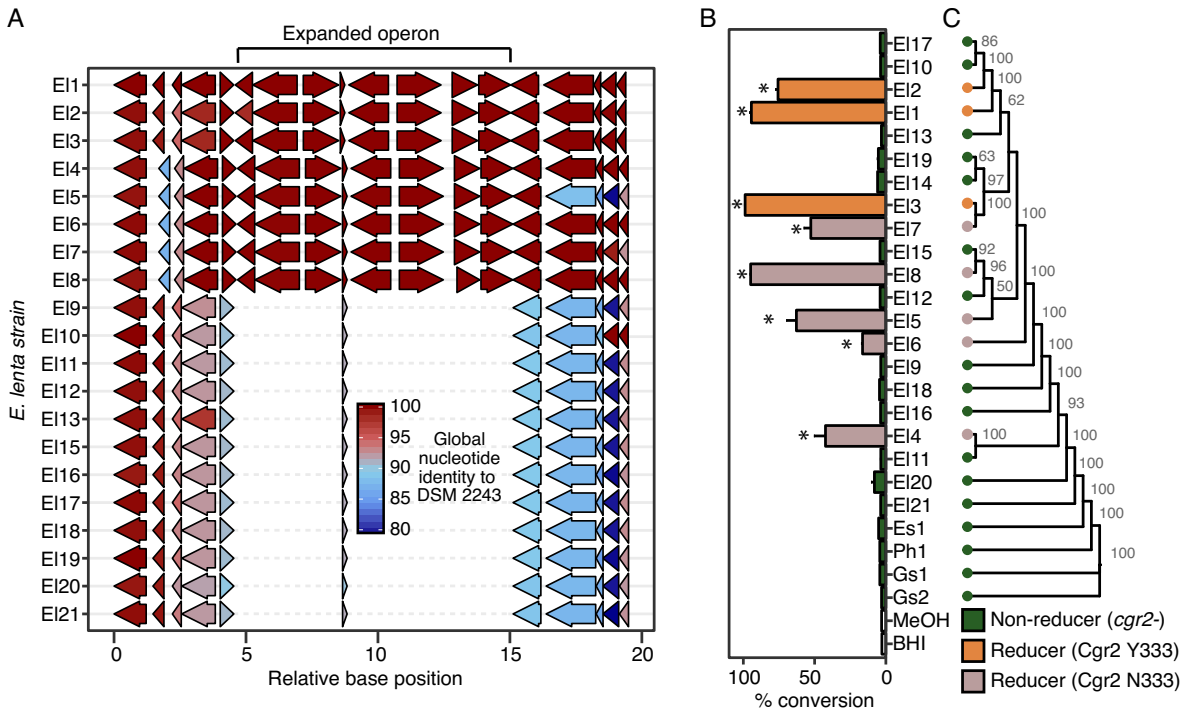


Figure 27: Comparative genomics expands the boundaries of the *cgr* operon
(A) Whole genome sequencing revealed an expanded, highly conserved 10.4 kb locus including the *cgr* operon and six additional genes in all digoxin-reducing strains of *E. lenta*. (B) Survey of digoxin reduction in 21 strains of *E. lenta* (El#), 2 strains of *Gordonibacter* spp. (Gs#), *E. sinensis* (Es1), and *Paraeggerthella hongkongesis* (Ph1) revealed 8 strains capable of reducing digoxin to dihydrodigoxin (* $p < 0.05$, ANOVA with Dunnett's test vs. vehicle controls). For full strain names, refer to Table 2 in Section **2.6**. (C) Digoxin reduction did not correlate with phylogeny in *E. lenta* species. Tree was constructed using 400 conserved genes (*22*). Bootstrap values are indicated at nodes.

**Genomic sequencing and analysis performed by Dr. Jordan Bisanz.*

The GC content of *cgr1* and *cgr2* is 64.2% and 63.9%, respectively, which is similar to the genomic average for *E. lenta* (64.2%) (*21*). On the other hand, the average GC content of the *cac* genes is 54.9%. In addition to the presence of the short *cac2* gene in all non-metabolizing strains, these observations could suggest that one or more horizontal gene transfer (HGT) events have taken place to insert (or remove) portions of the *cgr* or *cac* genes into this genomic locus. However, no apparent mobile genetic elements or flanking repeats could be found within the expanded cluster, and the *cac1* and *cac3-6* genes are not found in any other sequenced organisms in the NCBI database. Therefore, we could not definitively conclude whether the *cac* genes were acquired through HGT in metabolizing strains, or whether non-metabolizing strains had lost the *cgr* and *cac* operons over time.

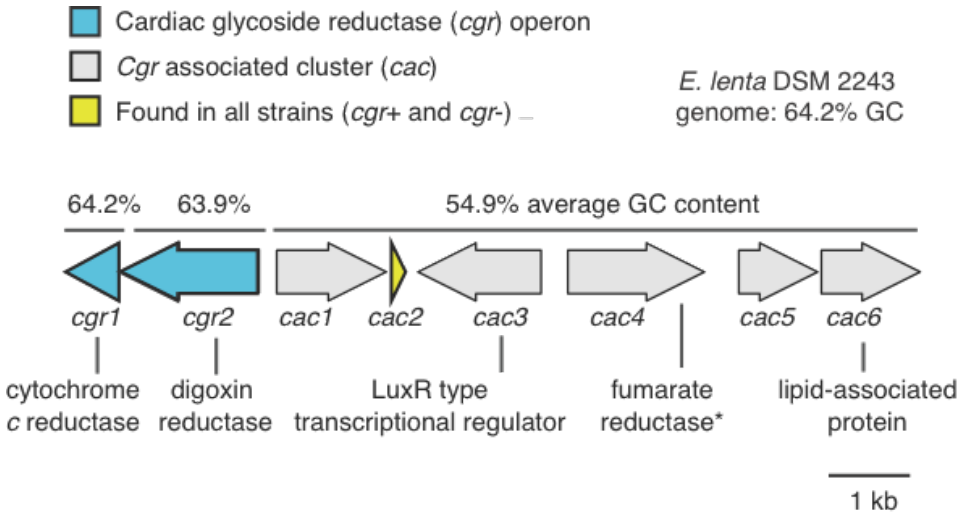


Figure 28: An expanded operon is present in all digoxin-reducing *E. lenta* strains.

Similarly to the *cac* genes, the *cgr* operon was remarkably conserved, with >96% amino acid sequence identity for Cgr1 and >98% amino acid sequence identity for Cgr2 across strains. Only two coding mutations were found in Cgr2, including an infrequent M381/V381 dimorphism (M381 only present in DSM2243 strain) and a frequent Y333/N333 dimorphism (3 and 5 strains, respectively) (Figure 29A). Within Cgr2 sequences that were reconstructed from human gut metagenomes (Section **2.4**), only the Y333 and N333 variants were detected (9 and 5 sequences, respectively) (Figure 29B). *E. lenta* strains harboring the N333 variant generally displayed lower digoxin-reducing activity in culture than strains encoding a Y333 variant of Cgr2 (Figure 29C). Additionally, phylogenetic clustering of Cgr2 variants suggests that Y333 may have derived from N333 (Figure 27C). This data could imply that over time, Cgr2 evolved to have enhanced catalytic efficiency, or that the two variants may have different substrates.

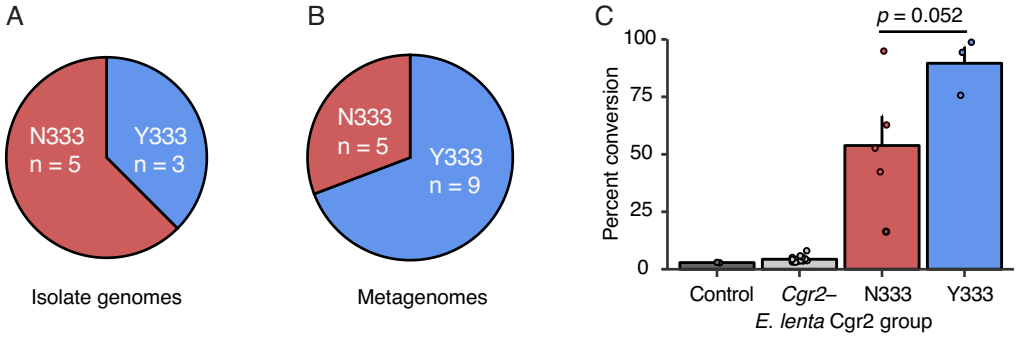


Figure 29: A common Y/N333 polymorphism in Cgr2 leads to altered digoxin metabolism. Prevalence of Cgr2 variants in (A) *E. lenta* genomes and (B) metagenomes. (C) Comparison of digoxin metabolism in culture by *E. lenta cgr2*$^{-}$ (n = 13 strains), *E. lenta* Cgr2 Y333 (n = 3 strains), *E. lenta* Cgr2 N333 (n = 5 strains), or digoxin in BHI media (control). Each point represents the mean percent conversion to dihydrodigoxin of each individual strain cultured in triplicate. Bars represent the mean ± SEM percent conversion per group. Statistical significance between Y333 and N333 groups was calculated using two-tailed Welch's t test (p = 0.052). **Sequencing and metagenomic analyses performed by Dr. Jordan Bisanz.*

2.4 *E. lenta* and *cgr2* are prevalent in human gut microbiomes

Eggerthella is a non-motile, anaerobic, Gram-positive bacterial genus that commonly colonizes the human gut and is one of the five most abundant genera within the *Actinobacteria* phylum (*24*). *Eggerthella* species are also highly associated with host health, particularly through the metabolism of bioactive host-derived and xenobiotic compounds (*24*). *E. lenta* strains can reductively metabolize a broad range of bioactive substrates including host bile acids (*25*) and many plant-derived compounds such as catechins (*26*), demethylated dietary lignans (*27*), isoflavones (*28*), and digoxin. More recently, *E. lenta* has been increasingly implicated as an opportunistic pathogen, especially in polymicrobial GI infections and in bacteremia (*29*). Despite their importance in both healthy and diseased states, *E. lenta* strains have been poorly characterized because they are slow-growing and challenging to detect in complex mixtures due to their recalcitrant cell walls and high GC content (*29*). Recently, using *E. lenta* specific qPCR probes, researchers detected $7x10^5$ CFU/g feces on average (19 individuals), although high variability was observed ($1x10^3$ - $1x10^6$) (*24*). To more comprehensively understand the distribution of this important xenobiotic-modifying bacterium in humans, we extended these analyses to a larger population. As *E. lenta* abundance alone is a poor predictor of digoxin metabolism in the clinic (*30*), we additionally quantified the digoxin-metabolizing gene *cgr2*.

Dr. Jordan Bisanz analyzed the prevalence and abundance of *cgr2* and *E. lenta* in 1872 human gut metagenomes using Metaquery (*31*) and in 228 human fecal samples (from 158 individuals) using quantitative PCR (qPCR). In order to correlate *E. lenta* abundance to *cgr2* abundance, a single copy marker gene "*elnmrk1*" that was present in all sequenced *E. lenta* strains was used as a proxy for this species (*32*). *E. lenta* and *cgr2* were detected in 42% and 28% of individual metagenomes, respectively (*31*). *Cgr1* and the other *cac* genes were detected at similar levels to *cgr2*, suggesting that these genes co-occur in *E. lenta* strains within the broader population (Figure 30A). In some metagenomes, *elnmrk1* was detected in the absence of *cgr2*, indicating the presence of non-metabolizing strains. *Cgr2* was highly correlated with *elnmrk1*, which suggests that this gene is restricted to *E. lenta* (Figure 30B). Although

some *cgr2*$^{+}$ *elnmrk1*$^{-}$ samples were detected at lower abundances of *cgr2*, this may represent a subpopulation of *E. lenta* strains that lacks the *elnmrk1* gene.

These results were recapitulated in fecal samples. However, as qPCR analysis is more sensitive and can detect lower abundance species, *elmrk1* and *cgr2* were detected in 82% and 75% of samples, respectively (Figure 30C). The majority of human gut metagenomes and fecal samples had *cgr* ratios (*cgr*: *elnmrk1*) < 1, which suggests that many individuals likely harbor a mixed community of *cgr*$^{+}$ and *cgr*$^{-}$ strains of *E. lenta* in their GI tract (Figure 30B, Figure 30D; insets). Overall, both the qPCR- and metagenomic sequencing-based analyses show that *E. lenta* and *cgr2* are widely distributed, but low in abundance in the human microbiota.

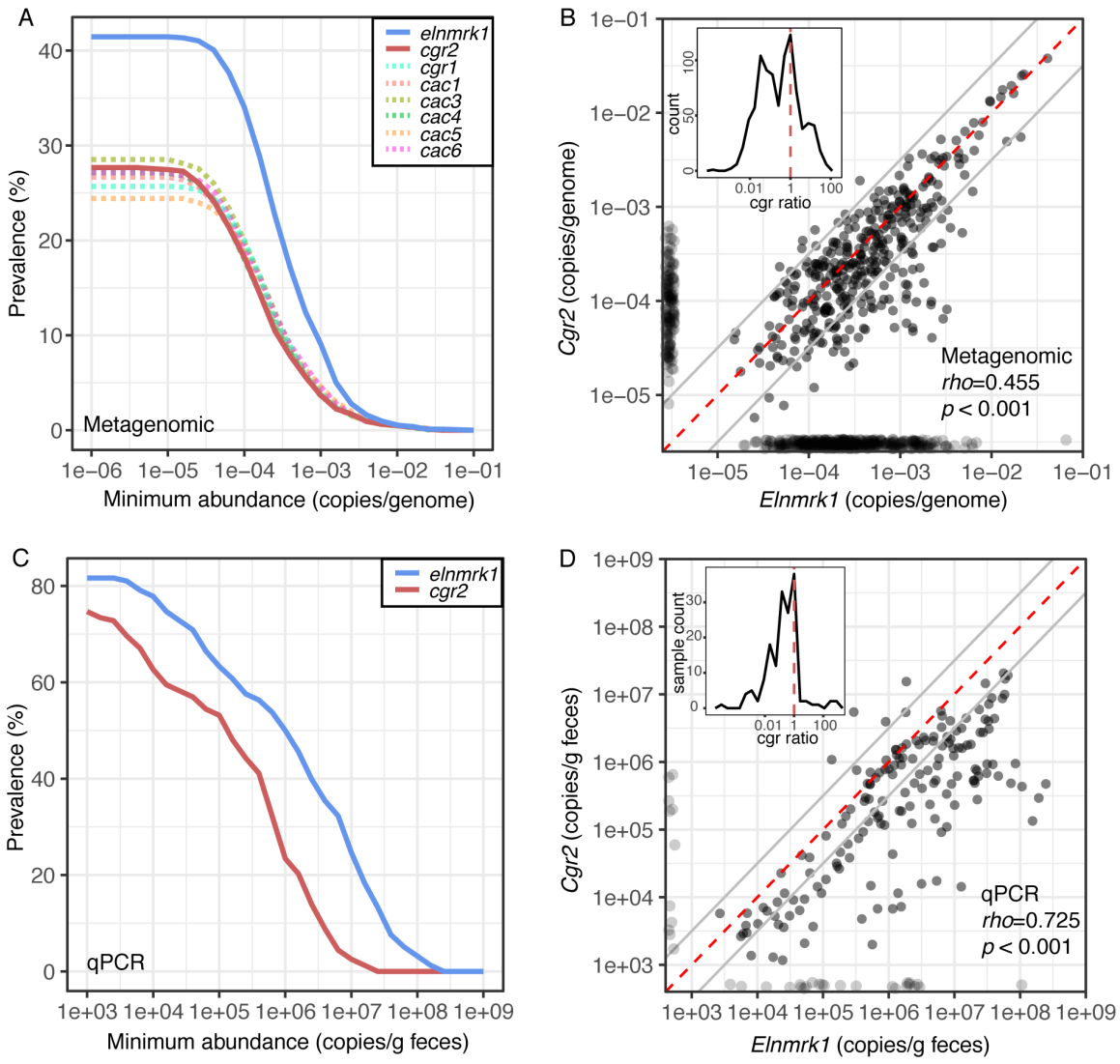


Figure 30: *E. lenta* and *cgr2* are prevalent in the human gut microbiota.
(A) Abundance of *cgr*, *cac*, and *elnmrk1* genes in 1872 human metagenomes. *Cgr2* and *E. lenta* were found in 28% and 42% of individuals, respectively, although often at low abundance. (B) *E. lenta* and *cgr2* abundances in gut metagenomes were tightly correlated suggesting that *cgr2* is only found in *E. lenta* strains. Expected linear relationship shown in red +/- half log deviation in dashed lines. (Inset) Histogram of *cgr/elnmrk1* ratio (*cgr*-ratio) of human metagenomes. (C) Detection of *cgr2* and *elnmrk1* genes in fecal samples of 158 individuals (228 samples) by qPCR. *Cgr2* and *E. lenta* were found in 75% and 82% of individuals, respectively near the detection limit (1E3 copies/g feces). (D) *E. lenta* and *cgr2* abundances in human fecal samples were correlated, as in metagenomic analyses. Expected linear relationship shown in red +/- half log deviation in dashed lines. (Inset) Histogram of *cgr*-ratio in human fecal samples.

**Analyses performed by Dr. Jordan Bisanz.*

2.5 Discussion

In many cases, variability in patient response to drugs cannot be explained by host genetics alone (*33*), highlighting the need to characterize the highly disparate and variable metabolic activities of the human microbiota. For over three decades, the human gut bacterium *E. lenta* has been linked to cardiac drug inactivation (*30*), although the identity, specificity, and distribution of the responsible enzymes were unknown. Our studies unambiguously demonstrate that the highly unique *E. lenta* enzyme Cgr2 is responsible for inactivating digoxin, a pharmaceutical agent that has been used for over two centuries for cardiac treatment. We have demonstrated that the *cgr* genes are highly conserved, are restricted to *E. lenta* strains, and are widespread in the human gut microbiota.

However, the prevalence of *cgr2*$^{+}$ *E. lenta* strains (~75% at lower abundances) is substantially higher than the incidence of dihydrodigoxin production observed clinically (10-35%) (*34, 35*). Although the presence of *cgr2* predicts whether *E. lenta* strains produce dihydrodigoxin in culture, additional factors likely dictate the extent of digoxin metabolism in humans. Further animal and human studies will thus be crucial to understand the minimum colonization levels required to alter digoxin concentrations and efficacy within the body and the impacts of different diets on this metabolism. The studies are especially important in light of the observations that dietary protein and L-arginine in particular have been shown to affect *cgr* operon expression and activity (*1, 36*). In addition, bacterial genetic diversity may play a role, although our findings suggest that *cgr2* and its associated genes are highly conserved, with only one common variant of Cgr2 (Y333/N333) that results in differential metabolism of digoxin in culture. This high sequence conservation is surprising as digoxin is rarely encountered outside of cardiac therapy, and ingested drug concentrations are very low to minimize toxicity (*37, 38*). It is possible that the *cgr* operon originated in bacteria found in environments where cardenolides are abundant (e.g. GI tracts of animals that ingest cardenolide-producing plants, plant associated bacteria), and where the selective pressure may be more significant (*39*). However, these genes have yet to discovered outside of the human gut microbiota. While we do not yet understand the factors that maintain the conservation of *cgr2* and its

associated genes within *E. lenta* strains, our studies suggest that the metabolic pathways of even low abundance human gut bacteria can significantly contribute to patient health and disease outcomes.

As we move towards an era of personalized medicine, it will be crucial to consider both host genetics as well as gut microbial genetics at the strain level. In an idealized scenario, microbial drug-metabolizing genes can serve as clinical biomarkers to predict metabolite levels and patient responses, or to stratify patients into treatment groups based on their susceptibility to drug metabolism. However, our results demonstrate that other factors may need to be considered. In the case of digoxin, not only do *E. lenta* strains vary in the presence or absence of *cgr2*, but a single, naturally occurring amino acid substitutions leads to decreased metabolism. Thus, highly accurate and sensitive methods will be required that can detect individual nucleotide variation of genes within complex microbiome samples.

Apart from shedding light on a clinically important transformation, our study of Cgr2 has also provided valuable insight into additional chemistry made possible by host-associated microbial communities. The gut microbiome encodes over 3 million genes (*40*), over half of which have unknown functions (*41*), but which may play a role in xenobiotic metabolism. Flavin-dependent reductases appear to be particularly widespread in human-associated microbes, with over 100 uncharacterized, divergent homologs of Cgr2 alone (Figure 22). More broadly, reductive transformations are abundant in anaerobic gut microbes, and many unsaturated FDA-approved drugs and dietary compounds are potential substrates for the gut microbiota. Studying these unique, putative reductase enzymes thus provides an opportunity to uncover exciting new biochemistry of this highly diverse microbial community. Coupled with studies that assess the prevalence and distribution of these metabolic pathways among members of the human microbiota, this information will clarify the emerging roles that gut microbes play in human health and disease.

2.6 Experimental

Bioinformatics

The full length Cgr1 and Cgr2 protein sequences from *E. lenta* DSM2243 were used as a query for BLASTP (*42*) using the NCBI non-redundant protein sequence database (searches performed 9/26/17). These sequences were also used as queries for PSI-BLAST (*42*) to search the *E. lenta* DSM2243 genome (taxid: 479437; searches performed 11/21/17) and all sequenced *Eggerthella* genomes (taxid:1643822; search performed 01/09/18). Cgr2 was used to query the HHPred prediction tool https://toolkit.tuebingen.mpg.de/#/tools/hhpred) to identify additional remote protein homologs using hidden Markov models (*6*). The search was performed on 9/26/17 using the PDB_mmCIF70_27_Aug database. Signal sequences and transmembrane regions were predicted using Phobius (http://phobius.sbc.su.se/), PRED-TAT (http://www.compgen.org/tools/PRED-TAT/), and SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/).

Construction of *cgr1* and *cgr2* vectors in *E. coli*

E. lenta DSM2243 was grown in 5mL of BBL Brain Heart Infusion (BHI) media (BD) + 1% arginine at 37 °C. After 2 days, the culture was pelleted and genomic DNA (gDNA) was purified with the UltraClean® Microbial DNA Isolation Kit (MoBio) according to protocol. The *cgr* operon was amplified from 50 ng of gDNA in a 50 µL reaction volume with 0.5 µM of each primer (Table 1) and Phusion High-Fidelity PCR master mix with HF buffer (New England Biolabs). The following thermocycling parameters were used: denaturation at 98 °C for 3 minutes; 35 cycles of 98 °C for 15 seconds, 71 °C for 20 seconds, and 72 °C for 75 seconds; and a final extension at 72 °C for 5 minutes. The PCR reaction was analyzed by agarose gel electrophoresis (1% agarose gel), and the *cgr* amplicon was excised and purified with the Illustra GFX PCR DNA and Gel Band Purification kit (GE Healthcare). *Cgr1* and *cgr2* variants were amplified in 20 µL PCR reactions using 1 ng of purified *cgr* operon as template, 0.5 µM primer pairs and Phusion High-Fidelity PCR master mix with HF buffer (New England Biolabs) (Table 1). PCR conditions were as follows: denaturation at 98 °C for 2 minutes; 35 cycles of 10 seconds at 98°C, 10

seconds at specified annealing temperature, and 72 °C for the specified extension time; and a final extension at 72 °C for 5 minutes. *Cgr* amplicons were digested in a 30 µL reaction with 1.5 µL of each restriction enzyme (New England Biolabs) for 2.5 hours at 37 °C. pTip vectors were similarly digested and the linearized vector was excised from a 1% agarose gel and purified. Insert and vector pairs were ligated at a 1:3 ratio at room temperature for 2 hours with T4 DNA ligase (New England Biolabs). 2.5 µL of the ligation reaction was transformed into chemically competent One Shot Top10 *E. coli* cells (Thermo Fisher Scientific) and plated on LB agar with ampicillin. Plasmid inserts were sequenced using the primers listed in Table 1.

Table 1: Primers and constructs for heterologous expression of Cgr1 and Cgr2 in *Rhodococcus erythropolis.*

Restriction sites are bolded. For= forward primer, Rev= reverse primer.

Construct	For/ Rev	Sequence	Restriction sites	Vector	Anneal temp (°C)	Extend time (sec)
cgr operon	For	ACTGACCCATGGATGGAATA CGGAAAGTGCC	n/a	n/a	71	75
	Rev	GTTTTACTGCAGTTACGCCG CCGTCGAA				
Cgr1 + Cgr2	For	TGAC**GAATTC**TAATGGAATA CGGAAAGTGCCG	EcoRI, BglII	pTipQT2	70	90
	Rev	TTATA**AGATCT**CGCCGCCGT CGAAAG				
Cgr1	For	TCGAA**CATATG**ATGGCTGAG GAACCTGTGG	NdeI, XhoI	pTipQT1	65	60
	Rev	ATAA**CTCGAG**TCACGCCGCC GTCGAAA				
Cgr2	For	ACTGAC**CCATGG**GCATGGA ATACGGAAAGTGCC	NcoI, HindIII	pTipQC2	65	60
	Rev	ATTAG**AAGCTT**TCACTCCCA CGGCTCGAG				
Sequencing primers	For	CGTGGCACGCGGAAC	n/a	All pTip vectors	n/a	n/a
	Rev	GTGCAGGTTTCGCGTG				

Heterologous expression of Cgr proteins in *R. erythropolis* L-88

All *Rhodococcus* strains and expression vectors were obtained from the National Institute of Advanced Industrial Science and Technology (Tokyo, Japan). 40 ng of DNA were added to 400 µL of *R.*

erythropolis L-88 electrocompetent cells in 30% PEG 1000 (Sigma-Aldrich) in a 2 mm gap electroporation cuvette (VWR). Cells were transformed in a MicroPulser electroporator (Bio-Rad) with a 2.5 kV pulse (time constant ~4.8 - 5.2), rescued with 0.6 mL of LB (Lennox) broth (Alfa Aesar), and incubated for 4 hours at 28 °C, 175 rpm. Cells were plated onto LB agar plates + antibiotic (17 µg/mL chloramphenicol for pTipQC plasmids; 8 µg/mL tetracycline for pTipQT plasmids) and incubated at 28 °C for 5-7 days. Single colonies were inoculated into 50-75 mL of LB + antibiotic (34 µg/mL chloramphenicol or 8 µg/mL tetracycline) and grown for 3-5 days at 28 °C, 175 rpm until reaching saturation. For gain of function studies, 50 mL of LB and antibiotic were inoculated to a starting OD_{600} of 0.2 and grown at 28 °C, 175 rpm. When cultures reached an OD_{600} of 0.6 (~6-8 hours), protein expression was induced with thiostrepton (Sigma-Aldrich) at a final concentration of 0.01 µg/mL, and cultures were incubated at 15°C, 175 rpm. In cultures where Cgr1 was overexpressed, media was supplemented with the heme precursor δ-amino levulinic acid hydrochloride (50 µg/mL final) (Frontier Scientific). After 16-20 hours, digoxin (as a solution in DMF) was added to cultures at a final concentration of 10 µM and incubated for either 7 days at 15 °C, or 2 days at 28 °C, 175 rpm. Culture supernatants were extracted and analyzed as described below. For large-scale purifications, 2 L of LB-chloramphenicol in a 4 L baffled flask were inoculated to a starting OD_{600} of 0.02 and grown to an OD_{600} of 0.6 (~18-25 hours). Protein expression was induced with 0.01 µg/mL thiostrepton, and cultures were incubated at 15 °C, 175 rpm for approximately 21 hours before harvesting cells by centrifugation (10,800 rpm x 20 min). Cell pellets were frozen and stored at –80 °C.

Cell lysis and partial membrane purification

All lysis and purification steps were carried out at 4 °C. Harvested cells were resuspended in 5 mL/g of cell pellet in lysis buffer (50 mM Tris, pH 8, 1 mM $MgCl_2$, 25 mM imidazole) containing Pierce EDTA-free protease inhibitor cocktail (Thermo Fisher Scientific). Cells were passaged through a cell disruptor (Avestin EmulsiFlex-C3) five times at 15,000-25,000 psi and centrifuged for 10 minutes at 2,000 g to remove unbroken cells. The supernatant was ultracentrifuged for 45 minutes at 90,000 g.

Pellets containing membrane and cell wall components were resuspended in wash buffer (50 mM Tris, 500 mM NaCl, pH 8) using a Turrax homogenizer, and centrifuged for an additional 45 minutes at 90,000 g. Washed pellets were solubilized in wash buffer supplemented with 1% *n*-dodecyl β-D-maltoside detergent **(**Sigma-Aldrich) with a Turrax homogenizer. Solution was nutated for 4 hours and centrifuged for 45 minutes at 37,000 g; supernatant contained solubilized membrane proteins. Samples were analyzed by SDS-PAGE analysis using a 4–15% Mini-PROTEAN® TGX™ Precast Gel (Bio-Rad) and Precision Plus Protein™ All Blue protein standards (Bio-Rad).

Extraction and LC-MS/MS detection of digoxin and dihydrodigoxin

Bacterial cultures were centrifuged (10 min x 4000 rpm) and 1 mL of supernatant was extracted three times with 1 mL of dichloromethane. Pooled organic fractions were concentrated using a rotary evaporator. Samples were resuspended in 1 mL of 50% methanol in water and diluted 10x prior to LC-MS/MS analysis. Metabolites were detected on an Agilent 6410 Triple Quad LC/MS using electrospray ionization in negative ion mode. The mass spectrometer settings were as follows: gas temperature (300°C), gas flow (10 L/min), nebulizer pressure (25 psi), capillary voltage (4000 V), and chamber current (0.1 μA). Digoxin was monitored using a 779.4 → 649.3 m/z transition with a fragmentor voltage of 250V and collision energy of 52, and dihydrodigoxin was monitored using a 781.4 →521.3 m/z transition with a fragmentor voltage of 200V and collision energy of 20. Standard curves were prepared using 0.01-1 μM samples of each compound. Digoxin was purchased from Sigma-Aldrich, and dihydrodigoxin standard was obtained through hydrogenation of digoxin as previously described (*1*). Liquid chromatography was performed on an Acclaim Polar Advantage II column with a flow rate of 0.125 mL/min, 5 μL sample injection, solvent A (10% methanol + 1 mM ammonium hydroxide) and solvent B (100% methanol + 1 mM ammonium hydroxide), and a gradient: 70-100% B over 10 minutes, 100% B for 1.5 min, 100-70% B over 3.5 min, and 70% B for 7 min.

Construction of sequence similarity network (SSN)

A SSN was generated using the EFI-EST tool (http://efi.igb.illinois.edu/efi-est/) (*10*). The full length (native) Cgr2 protein sequence was used as an input to generate a network with 5000 of the closest homologs from the UniProtKB protein database. An initial alignment score cutoff of *e*-value < 10^{-66} was used to generate a SSN with 2018 nodes (with 100% identity) and 317,130 edges. The SSN was imported into Cytoscope v 3.2.1 (*43*) and visualized with the 'Organic layout' setting. Seven characterized enzymes were present within the network (UniProtKB IDs: fumarate reductases: P83223, P0C278, Q07WU7, Q9Z4P0; urocanate reductase: Q8CVD0; 3-oxosteroid-1-dehydrogenases: P71864, Q7D5C1). The alignment score cutoff was increased to *e*-value < 10^{-130}, until enzymes with known functions were separated from each other into putatively isofunctional clusters; at this threshold, Cgr2 appears as a singleton. The network shown in Figure 22A was generated with a cutoff of *e*-value < 10^{-50}, at which nearly all protein sequences form one cluster. Multiple sequence alignments were generated in Geneious and visualized in Jalview (clustalx coloring). Cgr2 was aligned with characterized enzymes and additional selected genes within the corresponding clusters of the SSN to validate that the clusters contained putatively isofunctional proteins as determined by conservation of active site residues involved in substrate binding, activation and proton transfer (*11, 14, 15, 19, 20*).

Bacterial culturing

Eggerthella lenta and related strains were curated, isolated, and sequenced by the Turnbaugh lab (*32*). Strains were grown in BBL BHI media (BD) under an atmosphere of 2-5% H_2, 2-5% CO_2, and balance N_2. Strains were streaked onto BHI agar plates supplemented with 1% arginine (w/v) in an anaerobic chamber (COY Laboratory Products). Individual colonies were inoculated into 16 x 125 mm Hungate tubes (Chemglass Life Sciences) containing 5-10 mL of BHI supplemented with 1% arginine and grown at 37°C for 2-3 days. Cardiac glycoside substrates were dissolved at a concentration of 10 mM in dimethylformamide (DMF) and added to a final concentration of 10 µM. Starter cultures were diluted into 10 mL of BHI + substrate to a starting of OD_{600} of 0.05 and grown anaerobically at 37°C (in

triplicate) for 2 days. Culture supernatants were extracted and analyzed by LC-MS/MS as previously described.

Table 2: *Eggerthella lenta* and *Coriobacteriaceae* strains analyzed.

ID	Bacterial species	Strain ID
El1	*Eggerthella lenta*	DSM2243
El2	*Eggerthella lenta*	11c
El3	*Eggerthella lenta*	DSM11767
El4	*Eggerthella lenta*	CC8/6 D5 4
El5	*Eggerthella lenta*	AB12 #2
El6	*Eggerthella lenta*	AB8 #2
El7	*Eggerthella lenta*	32-6-I NA
El8	*Eggerthella lenta*	DSM11863
El9	*Eggerthella lenta*	FAA1-3-56
El10	*Eggerthella lenta*	14A
El11	*Eggerthella lenta*	22C
El12	*Eggerthella lenta*	28B
El13	*Eggerthella lenta*	DSM15644
El14	*Eggerthella lenta*	Valencia
El15	*Eggerthella lenta*	AN51LG
El16	*Eggerthella lenta*	MR1 #12
El17	*Eggerthella lenta*	FAA1-1-60A
El18	*Eggerthella lenta*	CC8/2 BHI2
El19	*Eggerthella lenta*	RC4/6F
El20	*Eggerthella lenta*	CC7/5 D5 2
El21	*Eggerthella lenta*	W1 BHI 6
Es1	*Eggerthella sinensis*	DSM16107
Gs1	*Gordonibacter* sp.	28C
Gs2	*Gordonibacter pamelaeae*	3C
Ph1	*Paraeggerthella hongkongesis*	RC2/2 A

E. lenta and *cgr2* prevalence analyses

The prevalence (gene copies/cell) of *E. lenta* (*elenmrk1*) and *cgr2* was determined in 1872 human gut metagenomes using Metaquery2 (*31, 32*). Metagenome hits were required to have at least 90% coverage and 90% nucleotide identity to *cgr2* and *elnmrk1* queries. For qPCR analyses, DNA was extracted from 228 human fecal samples (158 individuals) and analyzed by using the BioRad Universal Probes Supermix, 200 nM primers (Table 3) and a 60 °C annealing temperature (10 μL reactions, triplicate).

Table 3: qPCR primers and probes.

Gene	Forward primer	Reverse primer	Double-dye probes
Elnmrk1	GTACAACATGCTCCTTGCGG	CGAACAGAGGATCGGGATGG	[6FAM]TTCTGGCTGCACCGTTCGCGGTCCA[BHQ1]
Cgr2	GAGGCCGTCGATTGGATGAT	ACCGTAGGCATTGTGGTTGT	[HEX]CGACACGGAGGCCGATGTCG[BHQ1]

2.7 References

1. H. J. Haiser *et al.*, Predicting and manipulating cardiac drug inactivation by the human gut bacterium *Eggerthella lenta*. *Science* **341**, 295-298 (2013).

2. D. Aragao *et al.*, Structure of dimeric cytochrome *c3* from *Desulfovibrio gigas* at 1.2 A resolution. *Acta Crystallogr D Biol Crystallogr* **59**, 644-653 (2003).

3. M. Kern, O. Einsle, J. Simon, Variants of the tetrahaem cytochrome c quinol dehydrogenase NrfH characterize the menaquinol-binding site, the haem c-binding motifs and the transmembrane segment. *Biochem J* **414**, 73-79 (2008).

4. G. L. Kemp *et al.*, Kinetic and thermodynamic resolution of the interactions between sulfite and the pentahaem cytochrome NrfA from Escherichia coli. *Biochem J* **431**, 73-80 (2010).

5. O. Dym, D. Eisenberg, Sequence-structure analysis of FAD-containing proteins. *Protein Science* **10**, 1712-1728 (2001).

6. V. Alva, Nam, S., Soding, J., Lupas, A.N., The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. *Nucleic Acids Res* **44**, W410-W415 (2016).

7. Y. Mitani, X. Meng, Y. Kamagata, T. Tamura, Characterization of LtsA from *Rhodococcus erythropolis*, an enzyme with glutamine amidotransferase activity. *J Bacteriol* **187**, 2582-25891 (2005).

8. N. Nakashima, T. Tamura, Isolation and characterization of a rolling-circle-type plasmid from *Rhodococcus erythropolis* and application of the plasmid to multiple-recombinant-protein expression. *Appl Environ Microbiol* **70**, 5557-55568 (2004).

9. N. Nakashima, T. Tamura, A novel system for expressing recombinant proteins over a wide temperature range from 4 to 35 degrees C. *Biotechnol Bioeng* **86**, 136-148 (2004).

10. J. A. Gerlt *et al.*, Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. *Biochim Biophys Acta* **1854**, 1019-1037 (2015).

11. D. Leys *et al.*, Structure and mechanism of the flavocytochrome *c* fumarate reductase of *Shewanella putrefaciens* MR-1. *Nature Structural Biology* **6**, 1113-1117 (1999).

12. A. Brzostek, T. Sliwinski, A. Rumjiowska-Galewicz, M. Korycka-Machata, J. Dziadek, Identification and targeted disruption of the gene encoding the main 3-ketosteroid dehydrogenase in *Mycobacterium smegmatis*. *Microbiology* **151**, 2393-2402 (2005).

13. M. K. Doherty *et al.*, Identification of the active site acid/base catalyst in a bacterial fumarate reductase: a kinetic and crystallographic study. *Biochemistry* **39**, 10695-10701 (2000).

14. A. V. Bogachev, Y. V. Bertsova, D. A. Bloch, M. I. Verkhovsky, Urocanate reductase: identification of a novel anaerobic respiratory pathway in Shewanella oneidensis MR-1. *Mol Microbiol* **86**, 1452-1463 (2012).

15. J. Knol, K. Bodewits, G. I. Hessels, L. Dijkhuizen, R. van der Geize, 3-Keto-5alpha-steroid Delta(1)-dehydrogenase from Rhodococcus erythropolis SQ1 and its orthologue in Mycobacterium tuberculosis H37Rv are highly specific enzymes that function in cholesterol catabolism. *Biochem J* **410**, 339-346 (2008).

16. E. L. Rothery *et al.*, Histidine 61: an important heme ligand in the soluble fumarate reductase from *Shewanella frigidimarina*. *Biochemistry* **42**, 13160-13169 (2003).

17. S. L. Pealing *et al.*, Sequence of the gene encoding flavocytochrome *c* from *Shewanella putrefaciens*: a tetraheme flavoenzyme that is a soluble fumarate reductase related to the membrane-bound enzymes from other bacteria. *Biochemistry* **31**, 12132-12140 (1992).

18. P. S. Dobbin, J. N. Butt, A. K. Powell, G. A. Reid, D. J. Richardson, Characterization of a flavocytochrome that is induced during anaerobic respiration of Fe3+ by *Shewanella frigidimarina* NCIMB400. *Biochem J* **342**, 439-448 (1999).

19. G. A. Reid, C. S. Miles, R. K. Moysey, K. L. Pankhurst, S. K. Chapman, Catalysis in fumarate reductase. *Biochim Biophys Acta* **1459**, 310-315 (2000).

20. A. Rohman, N. van Oosterwijk, A. M. Thunnissen, B. W. Dijkstra, Crystal structure and site-directed mutagenesis of 3-ketosteroid Delta1-dehydrogenase from *Rhodococcus erythropolis* SQ1 explain its catalytic mechanism. *J Biol Chem* **288**, 35559-35568 (2013).

21. E. Saunders *et al.*, Complete genome sequence of *Eggerthella lenta* type strain (IPP VPI 0255). *Stand Genomic Sci* **1**, 174-182 (2009).

22. N. Segata, D. Börnigen, X. C. Morgan, C. Huttenhower, PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes. *Nat Comms* **4**, 2304 (2013).

23. M. Bramkamp, D. Lopez, Exploring the existence of lipid rafts in bacteria. *Microbiol Mol Biol Rev* **79**, 81-100 (2015).

24. G. S. Cho *et al.*, Quantification of *Slackia* and *Eggerthella* spp. in human feces and adhesion of representatives strains to Caco-2 cells. *Front Microbiol* **7**, 658 (2016).

25. A. S. Devlin, M. A. Fischbach, A biosynthetic pathway for a prominent class of microbiota-derived bile acids. *Nat Chem Biol* **11**, 685-690 (2015).

26. M. Kutschera, W. Engst, M. Blaut, A. Braune, Isolation of a catechin-converting human intestinal bacteria. *Journal of Applied Microbiology* **111**, 165-175 (2011).

27. T. Clavel *et al.*, Intestinal bacterial communities that produce active estrogen-like compounds enterodiol and enterolactone in humans. *Applied and Environmental Microbiology* **71**, 6077-6085 (2005).

28. F. Rafii, The role of colonic bacteria in the metabolism of the natural isoflavone daidzin to equol. *metabolites* **5**, 56-73 (2015).

29. B. J. Gardiner *et al.*, Clinical and microbiological characteristics of *Eggerthella lenta* bacteremia. *J Clin Microbiol* **53**, 626-635 (2015).

30. J. F. Dobkin, J. R. Saha, V. P. Butler, H. C. Neu, J. Lindenbaum, Digoxin-inactivating bacteria: Identification in human gut flora. *Science* **220**, 325-327 (1983).

31. S. Nayfach, M. A. Fischbach, K. S. Pollard, MetaQuery: a web server for rapid annotation and quantitative analysis of specific genes in the human gut microbiome. *Bioinformatics* **31**, 3368-3370 (2015).

32. J. E. Bisanz, H. J. Haiser, E. Allen-Vercoe, P. J. Turnbaugh, Comparative genomics of *Eggerthella lenta*: a prevalent opportunistic pathogen of interest to host-xenobiotic interactions. *In preparation*.

33. A. E. Rettie, G. Tai, The pharmacogenomics of Warfarin: Closing in on personalized medicine. *molecular interventions* **6**, 223-227 (2006).

34. E. Watson, D. R. Clark, S. R. Kalman, Identification by gas chromatography-mass spectroscopy of dihydrodigoxin- a metabolite of digoxin in man. *Journal of Pharmacology and Experimental Therapeutics* **184**, 424-431 (1973).

35. D. R. Clark, S. R. Kalman, Dihydrodigoxin: a common metabolite of digoxin in man. *Drug Metab Dispos* **2**, 148-150 (1974).

36. V. I. Mathan, J. Wiederman, J. F. Dobkin, J. Lindenbaum, Geographic differences in digoxin inactivation, a metabolic activity of the human anaerobic gut flora. *Gut* **30**, 971-977 (1989).

37. T. W. Smith, V. P. Butler, E. Haber, Determination of therapeutic and toxic serum digoxin concentrations by radioimmunoassay. *N Engl J Med* **281**, 1212-1216 (1969).

38. M. Gheorghiade, K. F. Adams, W. S. Colucci, Digoxin in the management of cardiovascular disorders. *Circulation* **109**, 2959-2964 (2004).

39. A. A. Agrawal, G. Petschenka, R. A. Bingham, M. G. Weber, S. Rasmann, Toxic cardenolides: chemical ecology and coevolution of specialized plant-herbivore interactions. *New Phytol* **194**, 28-45 (2012).

40. J. Qin *et al.*, A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59-65 (2010).

41. Human Microbiome Project Consortium, Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207-214 (2012).

42. S. F. Atlschul, Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., Gapped BLAST and PSI-BLAST: a new generation of database search programs. *Nucleic Acids Res* **25**, 3389-3402 (1997).

43. P. Shannon, Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504 (2003).

Chapter 3: Biochemical characterization of Cgr2, a [4Fe-4S] and FAD-dependent cardiac glycoside reductase.

3.1. Heterologous expression and activity of Cgr1 and Cgr2 towards digoxin

Using culturing, genomic, and bioinformatic analyses, we previously found that the *cgr* operon is highly associated with digoxin reduction by *E. lenta* and that heterologously expressed Cgr2 is sufficient for this metabolism in cells (**Chapter 2**). However, the role of Cgr1 and the mechanistic details of this transformation remained unknown. We thus aimed to biochemically characterize these proteins and their role in digoxin metabolism, understand the optimal conditions and potential cofactor(s) required for activity, and glean mechanistic insights into this metabolic reaction. By confirming the role of these enzymes in digoxin metabolism, we would validate the use of *cgr* genes as candidate biomarkers for assessing the distribution and clinical potential for drug metabolism, as well as shed light on highly unique gut microbial proteins.

In order to determine the activity of the Cgr enzymes, we sought to reconstitute their activity *in vitro*. A range of expression conditions and protein constructs were tested, taking into account the predicted cofactor requirements and cellular localization of these enzymes: Cgr1 is annotated as a membrane anchored cytochrome *c* while Cgr2 is a predicted flavoprotein that is predicted to undergo secretion to the periplasm via that Tat pathway (Section **2.1**). Heterologous expression of Cgr1 and Cgr2 was first attempted in *E. coli*. Various constructs were prepared for each enzyme, including full-length proteins and truncations, in which membrane anchors or secretion sequences were excised. Cgr1 constructs were co-transformed with pEC86, a plasmid encoding cytochrome maturation factors that can increase protein yield and allow for aerobic expression of heterologous cytochromes (*1*). Hexahistidine (His6x) affinity tags were used on either the N- or C- termini. However, as the His6x motif can sequester heme groups and potentially impede formation of the Cgr1 holoprotein, N-terminal Maltose-Binding Protein (MBP) fusion constructs were also prepared. Finally, PelB, an *E. coli*-recognized signal sequence that promotes secretion to the periplasm was inserted at the N-termini of the coding sequences of Cgr1 and Cgr2. In addition to the full-length Cgr2 construct, a cytoplasmic Cgr2 construct was prepared in which the first 48 amino acids (corresponding to the Tat secretion signal) were removed. Finally, various constructs of Cgr1 and Cgr2 were co-expressed.

Expression was attempted for all constructs at various temperatures (15-37 °C), isopropyl β-D-1-thiogalactopyranoside (IPTG) induction concentrations, and expression times, as well as in multiple *E. coli* strains (BL21(DE3), Rosetta, and Tuner cells). Media was supplemented with hemin, or with the heme precursor δ-amino levulinic acid (δ–ALA), as production of this metabolite is the rate-limiting step in heme biosynthesis. In all of these constructs and conditions, Cgr1 and Cgr2 protein overexpression (as assessed by SDS-PAGE analysis) was only observed for MBP-fusions. However, a stable MBP truncation was generated *in vivo* (37 kDa), and cleavage of the MBP fusion with Factor Xa protease led to rapid and complete degradation of the target proteins (Figure 31). These results suggested that Cgr1 and Cgr2 were either unfolded or unstable in the absence of the MBP tag, and were likely expressed only because of the robust stability and ease of expression of MBP (*2*). In addition, flavin and heme cofactors could not be detected by LC-MS in either Cgr2-MBP or Cgr1-MBP protein preps, respectively, suggesting that these fusion proteins were unlikely to be functional.

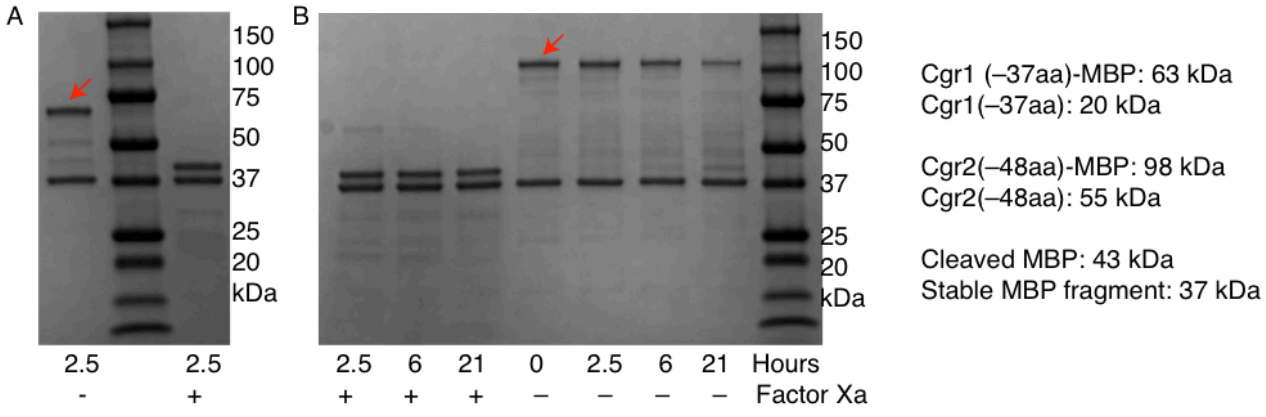


Figure 31: Heterologous expression of Cgr-MBP fusion proteins in *E. coli*.
Purification of N-terminal (A) Cgr1-MBP and (B) Cgr2-MBP fusion proteins. Cleavage with Factor Xa protease at 4 °C led to complete degradation of target proteins, while MBP (43 kDa) was recovered. The band at 37 kDa represents a stable MBP truncation generated *in vivo*. Red arrows points to Cgr-MBP fusion protein prior to cleavage.

We reasoned that Cgr1 and Cgr2 co-expression may be required for protein stability and activity, and postulated that a failure to heterologously express cytochromes *c* of Gram-positive origin (Cgr1) in the Gram-negative host *E. coli* could be due to an incompatibility of cytochrome *c* maturation factors (*3*). We thus moved to the Gram-positive expression host *Rhodococcus erythropolis* L88 (*4*). This host has similar GC content to *E. lenta* (67% and 64%, respectively), can grow and express proteins at a wide temperature range (4-35 °C), and contains homologs of the cytochrome maturation enzymes necessary for Cgr1 holo-protein formation (*1*). Using the inducible pTip expression system (*5, 6*), we heterologously expressed various Cgr1 and Cgr2 constructs in *R. erythropolis*, either individually or in combination. After inducing protein expression, cultures were incubated with digoxin, and supernatants were extracted with DCM and analyzed by LC-MS/MS (Figure 32A). We detected dihydrodigoxin only in strains expressing Cgr2 (14-69% conversion), whereas Cgr1 expression did not result in digoxin reduction. SDS-PAGE analysis of these strains revealed Cgr2 in the soluble fraction, regardless of construct (Figure 32B), whereas Cgr1 could not be detected in lysates or in membrane fractions by Coomassie staining (Figure 32C), or by a more sensitive heme-peroxidase gel stain assay (Figure 32D) (*1*). These results again showed that Cgr2 is necessary and sufficient for reducing digoxin *in vivo* (see Section **2.1**).

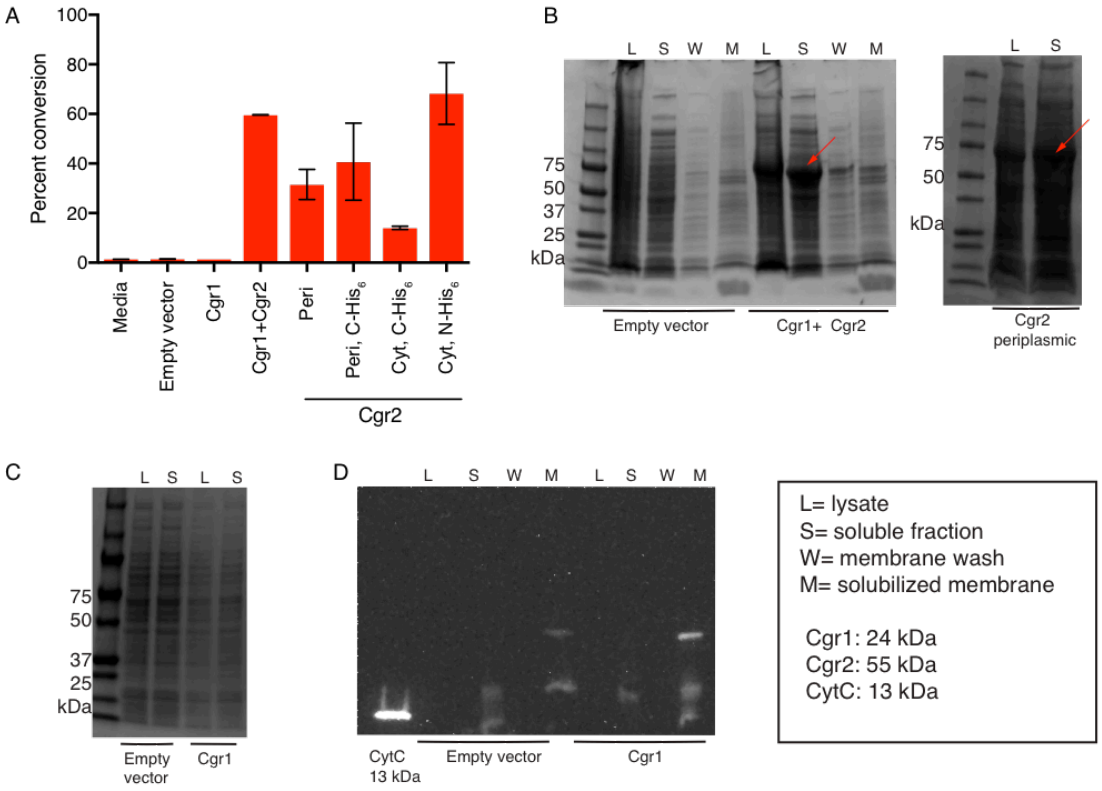


Figure 32: Heterologous expression of Cgr proteins in *R. erythropolis*.
(A) Whole cell assays in *R. erythropolis* expressing various Cgr constructs demonstrated that Cgr2 is sufficient for reducing digoxin. Data represents the mean ± SEM (n = 3). Peri = periplasmic, Cyt = cytoplasmic expression. (B) SDS-PAGE analysis showed overexpression of soluble Cgr2 constructs, whereas (C) no overexpression was observed for Cgr1 constructs. (D) No distinct cytochrome *c* containing bands were observed in Cgr1-expressing strains (relative to empty vector controls) using a heme-peroxidase gel stain assay (*1*). A 13 kDa cytochrome *c* from bovine heart (CytC) was used as a positive control.

3.2. Cgr2 requires [Fe-S] cluster reconstitution and FAD for *in vitro* stability and activity

After establishing a direct biochemical link between Cgr2 and digoxin metabolism in culture, we next aimed to reconstitute its activity *in vitro*. We attempted to purify multiple tagged versions of Cgr2, including His6x tagged and untagged full-length (predicted periplasmic localization) and cytoplasmic constructs in which the Tat secretion signal was removed (–48aa). Although we observed robust overexpression of Cgr2 in cells, subsequent lysis and purification led to substantial protein degradation. For example, lysis by sonication led to near complete proteolysis of Cgr2 (Figure 33A), possibly as a result of heating the sample. Cell disruption methods were thus used to recover intact protein (Figure 33B). Of all tested constructs, cytoplasmically expressed Cgr2(–48aa)-NHis6x construct lacking the Tat secretion signal gave the highest yield (Figure 33B-D) and activity in cells (Figure 32A), and was used for further characterization studies.

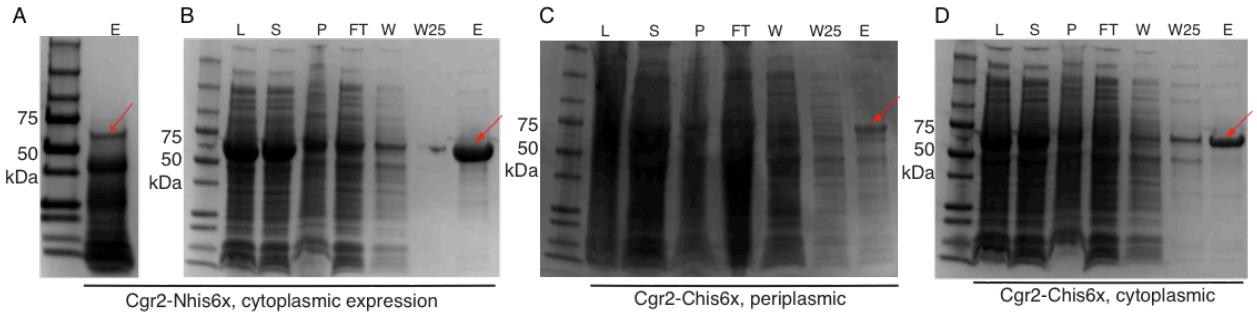


Figure 33: Purification of Cgr2 constructs from *R. erythropolis*.
(A) SDS-PAGE analysis of heterologously expressed Cgr2(–48aa)-NHis6x (expected mass = 55 kDa) purified on HisPur Ni-NTA resin**.** Certain lysis methods (e.g. sonication) led to substantial protein degradation as compared to (B) mechanical cell disruption. SDS-PAGE analysis of (C) Cgr2-CHis6x and (D) Cgr2(–48aa)-Chis6x that were purified on HisPur Ni-NTA resin. L = lysate, S = supernatant, P = pellet, FT = flow through, W = wash, W25 = wash with 25 mM imidazole, E = elution with 200 mM imidazole.

In addition to its propensity for proteolytic degradation, purified Cgr2 was thermally unstable (Figure 35A), displayed low and variable activity *in vitro* (Figure 34C), and did not co-purify with flavin as anticipated based on its annotation as a flavin-dependent reductase (Section **3.5**). Furthermore, protein

preparations were light brown in color and ultraviolet-visible (UV-vis) spectroscopy revealed a modest absorbance peak around 420 nm (Figure 36B). Together, these results suggested that essential metal or organic cofactors were likely missing from or present in low abundance in purified Cgr2. Reductase systems resembling Cgr2 often contain metallocofactors that deliver reducing equivalents to the active site flavin cofactor, including cytochromes *c* in soluble enzymes and oxygen-sensitive iron-sulfur ([Fe–S]) clusters in membrane-bound enzymes (*7-9*). Additionally, [Fe-S] clusters are typically coordinated by cysteine residues, and the mature Cgr2 protein contains 16 cysteines. The cysteine content of the mature Cgr2 (3.1%) was substantially higher than the average for bacterial (1%) and human (~2.3%) proteins, suggesting that these residues may serve an important role in Cgr2 (*10*). Therefore we reasoned that Cgr2 might contain one or more [Fe-S] cluster(s). However, we were unable to detect any canonical [2Fe-2S], [3Fe-4Fe], or [4Fe-4S] cluster binding motifs within the Cgr2 sequence (Table 6 in Section **3.8**) (*11-27*). These observations led us to hypothesize that Cgr2 contains [Fe-S] cluster(s) that may be important for protein stability and/or function, and which are ligated by a non-canonical [Fe-S] cluster-binding motif.

We next sought to determine which cofactors and metal centers were required for Cgr2 to catalyze digoxin reduction. As the presumed partner reductase Cgr1 could not be obtained, methyl viologen (MV) that had been pre-reduced with sodium dithionite was used as an artificial electron donor to initiate anaerobic Cgr2 reduction of digoxin *in vitro*. Reactions were monitored by measuring the decrease in absorbance at 600 nm, which corresponds to substrate-dependent MV oxidation (Figure 34A) (*28*). As this colorimetric assay is a direct measure of electron transfer and not product formation, *in vitro* reactions were also quenched and analyzed by LC-MS/MS to confirm that the decrease in absorbance was accompanied by increased dihydrodigoxin production. In this assay, negligible MV oxidation and dihydrodigoxin production was observed with purified Cgr2 (Figure 34B-C). Anaerobic chemical reconstitution of [Fe-S] clusters in Cgr2 using ammonium iron (II) sulfate and sodium sulfide (12.5 equivalents each, overnight, 4 °C) greatly enhanced digoxin reductase activity. In addition to reconstitution of the [Fe-S] clusters, supplementation of FAD, but not FMN, was required for Cgr2 activity (Figure 34B). Additionally, exposure of reconstituted protein to aerobic conditions led to a loss of

activity, suggesting that Cgr2 contains a labile [Fe-S] cluster (e.g. [4Fe-4S]) that may decompose upon exposure to oxygen. Together, these experiments demonstrated that Cgr2 is an oxygen sensitive, [Fe-S] cluster- and FAD-dependent digoxin reductase.

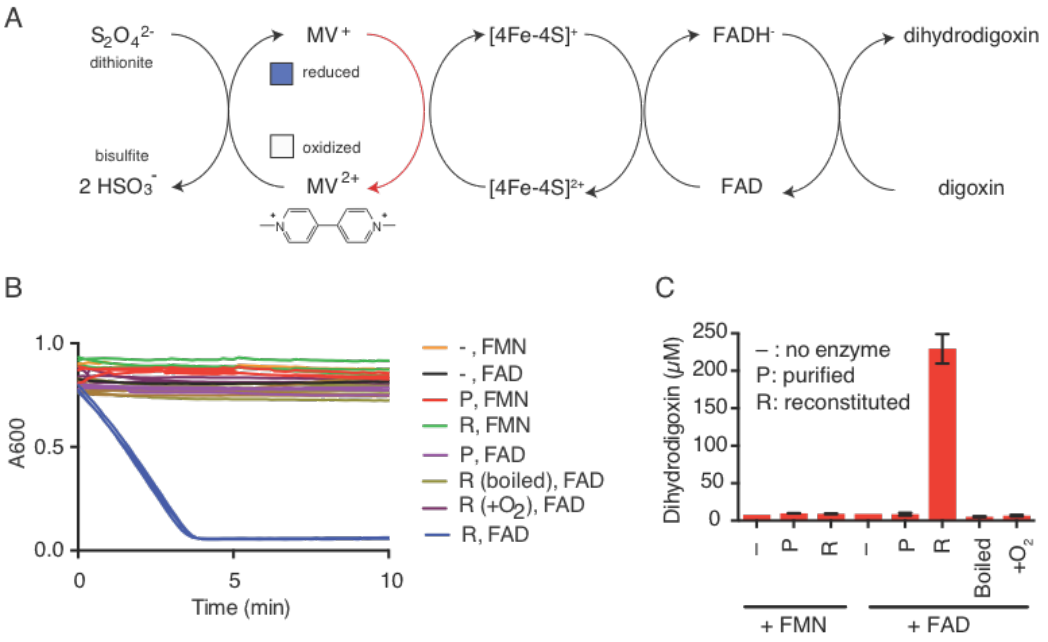


Figure 34: *In vitro* digoxin reduction assay using methyl viologen as an electron donor. (A) MV that had been pre-reduced with sodium dithionite was used as an electron donor *in vitro*. Reduced MV is blue, whereas oxidized MV is colorless, and a decrease in absorbance at 600 nm corresponded to substrate-dependent MV oxidation over time. (B) *In vitro* assays with Cgr2 preparations and digoxin. (C) *In vitro* assays were quenched in methanol at 10 minutes and analyzed by LC-MS/MS. [Fe-S] cluster reconstitution, FAD, and anaerobic conditions were required for Cgr2 activity. Data represents the mean ± SEM (n = 3).

In addition to enhancing Cgr2 activity, [Fe-S] cluster reconstitution affected protein stability and oligomeric state. Reconstituted Cgr2 had significantly improved thermal stability, with melting temperatures between 40 - 50 °C as compared to melting temperatures < 37 °C for purified protein (Figure 35A-B). In addition, reconstitution led to a shift from an apparent dimeric state to monomeric protein as analyzed by analytical size exclusion chromatography (Figure 35C), suggesting that [Fe-S] reconstitution is important for protein structure and that [4Fe-4S] cluster formation could result in a more compact protein. When SDS-PAGE analysis was performed in the absence of a reducing agent, higher

molecular weight oligomers were observed for Cgr2. Upon reduction with dithiothreitol (DTT), both purified and reconstituted Cgr2 were predominantly monomeric (Figure 35D). These data suggest that Cgr2 dimerization could result from intermolecular disulfide bonds formation between exposed cysteines that may endogenously be involved in [Fe-S] cluster ligation. These results also suggest that Cgr2 is partially unfolded, or that our reconstitution protocol does not achieve complete [Fe-S] cluster formation, resulting in free cysteines.

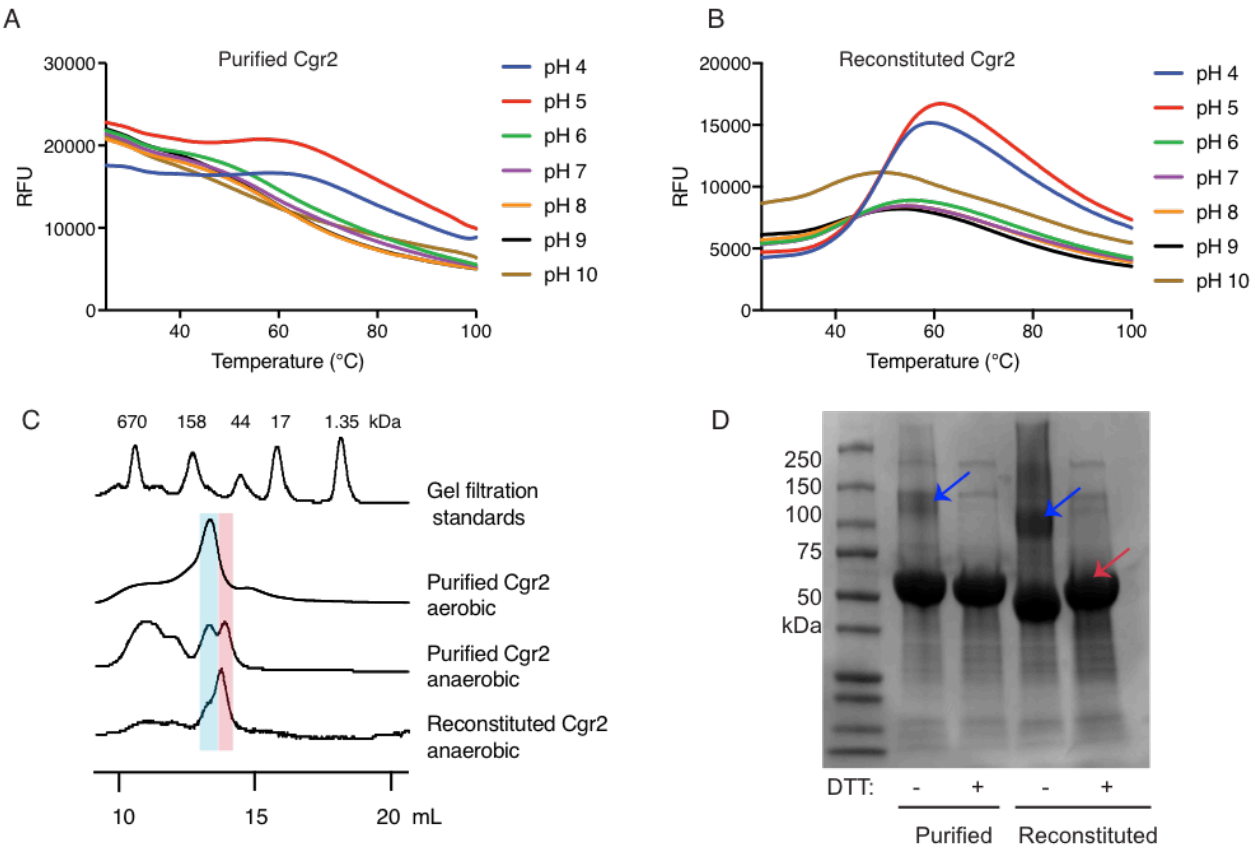


Figure 35: [Fe-S] cluster(s) affect Cgr2 stability and oligomerization.
(A) Thermal melt curves displaying relative fluorescence of Sypro Orange bound to purified and (B) reconstituted Cgr2 in various pH buffers. Peaks represent melting temperature (Tm) with values < 37 °C before reconstitution and 40 - 50 °C after reconstitution. (C) Analytical size exclusion chromatography traces performed under aerobic and anaerobic conditions. Colored bars highlight molecular weights corresponding to dimeric (blue) or monomeric (pink) Cgr2. (D) Native gel analysis of aerobically purified and anaerobically reconstituted Cgr2 (~55 kDa; pink arrows) in the presence or absence of dithiothreitol (DTT). In the absence of reducing agents, higher molecular weight oligomers (~dimers; blue arrows) are present that may result from intermolecular disulfide bonds.

3.3. Characterization of [Fe-S] clusters in Cgr2

3.3.1 UV-vis and EPR spectroscopy reveal the presence of [4Fe-4S] cluster(s) in Cgr2

Having demonstrated that [Fe-S] reconstitution was essential for the activity and stability of Cgr2, we next attempted to determine the exact nature of this metallocofactor. Prior to reconstitution, aerobically purified Cgr2 contained between 0.2 - 0.6 equivalents of iron and sulfide. We also detected low levels of UV-vis absorption peaks at ~330 and 420 nm (Figure 36B), which are characteristic of [2Fe–2S] clusters that may be formed through oxidative damage of [4Fe-4S] clusters (*29*). After anaerobic reconstitution, Cgr2 exhibited a broad peak around 400 nm that decreased in absorbance upon addition of an excess amount of the reducing agent sodium dithionite (Figure 36B). Exposure of reconstituted Cgr2 to oxygen led to [Fe-S] cluster decomposition as evidenced by a decrease in the absorbance at 400 nm (Figure 36C). These spectral properties are characteristic of redox-active, oxygen-sensitive [4Fe-4S] clusters. The oxygen-sensitivity of these clusters could also explain our previous issues with stability and activity of aerobically expressed and purified Cgr2 (Figure 35).

To further elucidate the chemical nature and extent of assembly of redox-active [Fe-S] cofactors in Cgr2, we turned to electron paramagnetic resonance (EPR) spectroscopy in collaboration with Dr. Maria-Eirini Pandelia (Brandeis University). This technique detects unpaired electrons in both organic and inorganic molecules and can differentiate between the various types of [Fe-S] clusters present in proteins as well as provide information about cluster orientation and redox state (*30*). In the absence of a reducing agent, purified Cgr2 was EPR-silent (Figure 36D), excluding the presence of mono- or tri-nuclear [Fe-S] centers. When reduced with sodium dithionite, its EPR spectrum exhibited a signal with axial symmetry and with principal g-components of 2.045 and 1.94. This signal increased in intensity upon reconstitution of Cgr2 with iron and sulfide (10 K) and was barely detectable at higher temperatures (40 K) (Figure 36D). Both the principal g-values and the relaxation properties (temperature dependence) of this signal are characteristic of low-potential tetranuclear $[4Fe\text{-}4S]^{1+}$ centers (*11, 30, 31*). These observations indicate that Cgr2 contains $[4Fe\text{-}4S]^{2+}$ cluster(s) that can undergo reduction to the

corresponding $[4Fe\text{-}4S]^{1+}$ state. These redox properties also suggest that the [4Fe-4S] clusters serve a catalytic (electron transfer) rather than purely structural role in Cgr2.

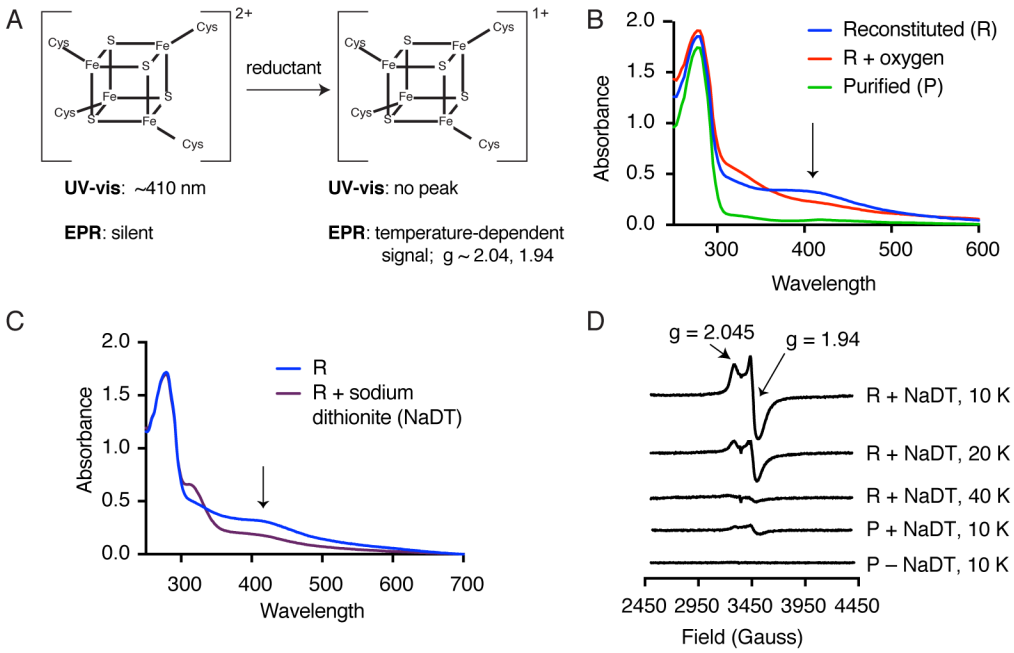


Figure 36: Cgr2 contains redox active, oxygen-sensitive [4Fe-4S] cluster(s).
(A) Typical spectroscopic properties of [4Fe-4S] clusters. (B) UV-vis absorption spectra of Cgr2 revealed an oxygen-sensitive peak centered around 400 nm that increased upon [Fe-S] cluster reconstitution, supporting the presence of [4Fe-4S] clusters in Cgr2. (C) UV-vis spectra of reconstituted Cgr2 in the absence or presence of reducing agent sodium dithionite (NaDT) revealed that the [Fe-S] clusters in Cgr2 are redox active. (D) EPR spectra of sodium dithionite-reduced Cgr2 reconstituted with 12.5 equivalents of iron ammonium sulfate hexahydrate ($(NH_4)_2Fe(SO_4)_2 \cdot 6H_2O$) and sodium sulfide ($Na_2S \cdot 9H_2O$). G-values and decreased EPR signal intensity at higher temperatures (10-40 K) indicated the presence of low potential $[4Fe\text{-}4S]^{1+}$ clusters. Experimental conditions were microwave frequency 9.38 GHz, microwave power 0.2 mW, modulation amplitude 0.6 mT, and receiver gain 40 dB.

3.3.2 Determining the metal dependence and [4Fe-4S] cluster stoichiometry of Cgr2

Next, we sought to determine the stoichiometry of [4Fe-4S] clusters in Cgr2. Using colorimetric assays (*32*), we determined that purified Cgr2 contained 0.2 - 0.6 equivalents of iron and sulfide prior to reconstitution (over different purifications). However, our inability to effectively remove excess iron and sulfide following reconstitution (8-12 equivalents each) prevented accurate iron and sulfur content determination in the most active Cgr2 preparations. Various attempts to remove excess iron and sulfur from Cgr2 after reconstitution, including incubation with EDTA, anaerobic or aerobic FPLC in a variety

of buffers, dialysis, or proteolytic cleavage of the His6x tag led to cluster decomposition, protein aggregation, and concomitant loss of *in vitro* activity (Figure 37). With the exception of desalting on a PD-10 column, further purification could not be performed on reconstituted Cgr2 without compromising protein activity or stability. These data suggested that the [Fe-S] cluster(s) of Cgr2 are susceptible to decomposition. One explanation for this labile nature is that the [4Fe-4S] clusters may be near the surface of the protein, and in the absence of the presumed reductase partner Cgr1 and are not shielded from oxidative or mechanical stresses that damage the cluster (*33*).

As exact determination of total iron and sulfur content was not possible, we assessed Cgr2 [Fe-S] cluster stoichiometry using EPR. However, as only redox-active [4Fe-4S] clusters in the 1+ state are detectible by this method, we would be unable to detect other metal centers, including [4Fe-4S] clusters that are not redox active (e.g. clusters involved in a structural role) or divalent Fe^{2+} species. Using a Cu^{2+}-EDTA standard under non-saturating conditions, we determined that purified Cgr2 contained 0.02 - 0.03 $[4Fe\text{-}4S]^{1+}$ clusters per protein monomer prior to reconstitution. After reconstitution with iron and sulfide, the intensity of the EPR signal increased to 0.13 - 0.25 $[4Fe\text{-}4S]^{1+}$ clusters per Cgr2 over different experiments (Figure 36D). We next optimized [Fe-S] cluster incorporation by varying reconstitution conditions including iron species ($Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$ or $FeCl_3$), buffer components (HEPES, Tris, pH 7-8), concentration of the reducing agent DTT (0.125, 2 mM), iron and sulfur equivalents (0-8x), and reconstitution time (5-21 hours). We then correlated the *in vitro* activity and $[4Fe\text{-}4S]^{1+}$ content of different Cgr2 preparations to determine conditions that resulted in EPR signal saturation. While EPR analysis showed a significant increase in cluster content when comparing purified to reconstituted protein, all reconstitution conditions yielded comparable [4Fe-4S] spin values (0.20 - 0.25 clusters), which could suggest that only one reducible, EPR-sensitive cluster is present in Cgr2 (Figure 40A). However, Cgr2 was most active when reconstituted using the highest Fe and S equivalents and longest reconstitution times (Figure 38B). We also noticed that Cgr2 was significantly more active when reconstituted with Fe^{2+} rather than Fe^{3+} species.

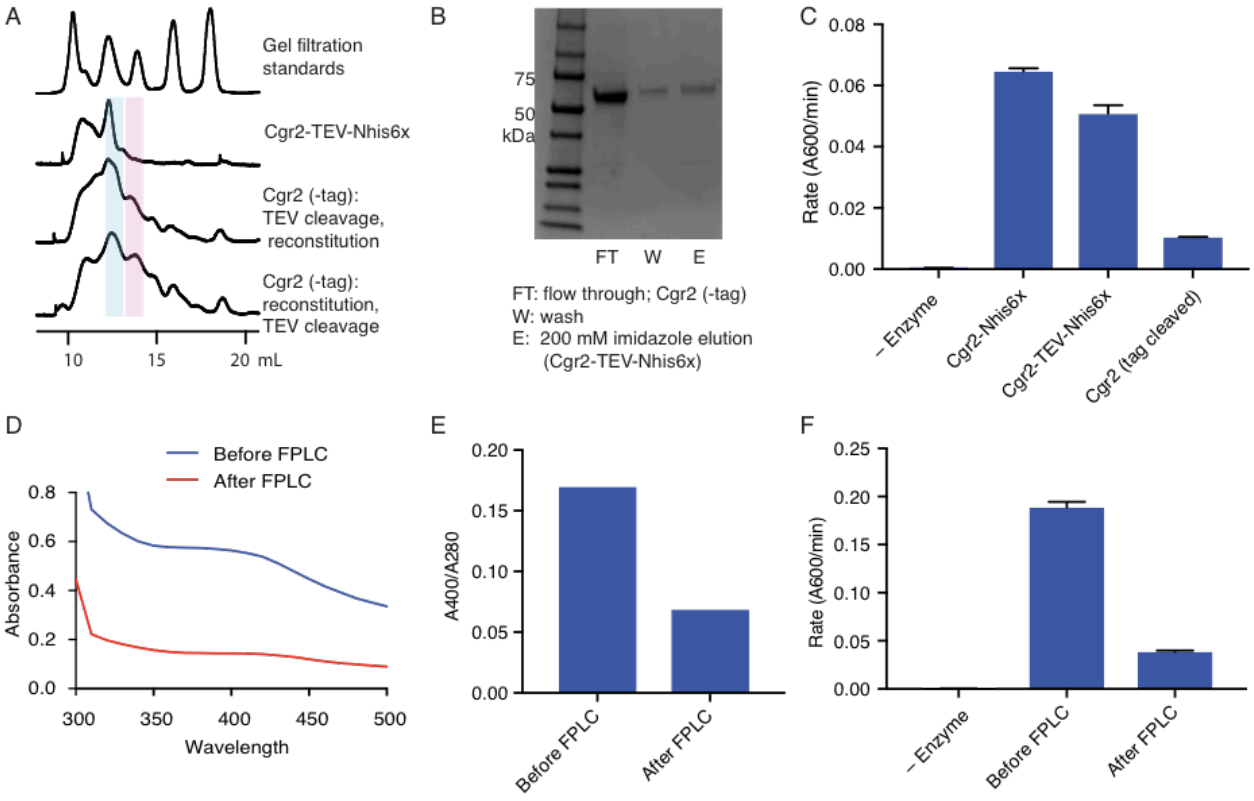


Figure 37: Purification after reconstitution leads to loss of activity and stability of Cgr2. (A) Anaerobic, analytical size exclusion chromatography of Cgr2 constructs. Cleavage of the His6x tag from Cgr2 constructs led to protein aggregation. Pink bars correspond to monomeric protein, and blue bars correspond to dimeric protein. (B) SDS-PAGE analysis of untagged Cgr2 protein following TEV protease cleavage of His6x tag. (C) Cleaved protein (-tag) had significantly lower *in vitro* activity than unprocessed Cgr2-TEV-Nhis6x construct. (D) UV-vis spectra and (E) A400/A280 ratios revealed a loss of [Fe-S] cluster following preparative scale anaerobic size exclusion chromatography of Cgr2-Nhis6x. (F) Cgr2-Nhis6x was less active *in vitro* following size exclusion chromatography.

This prompted us to test whether excess Fe^{2+} could account for the differential Cgr2 activity observed across different reconstitution procedures. Indeed, we found that a range of divalent metal cations (8 equivalents each of Fe^{2+}, Mn^{2+}, Mg^{2+}) stimulated the activity of Cgr2 *in vitro* (Figure 38C) although they did not alter protein stability (Figure 38D). Notably, binding of digoxin to its target in human cells (Na^{+}/K^{+} ATPase) is thought to be mediated by long-range (6.2 Å) electrostatic interactions between a Mg^{2+} ion and the electron rich, partially negatively charged oxygen atom of the unsaturated lactone (*34, 35*). It is possible that a divalent metal cation similarly positions or activates digoxin in the

active site of Cgr2. Additional spectroscopic and structural experiments (Section **5.2**) are warranted to determine whether Cgr2 indeed contains a divalent metal center and to elucidate what role it plays in the transformation of digoxin to dihydrodigoxin.

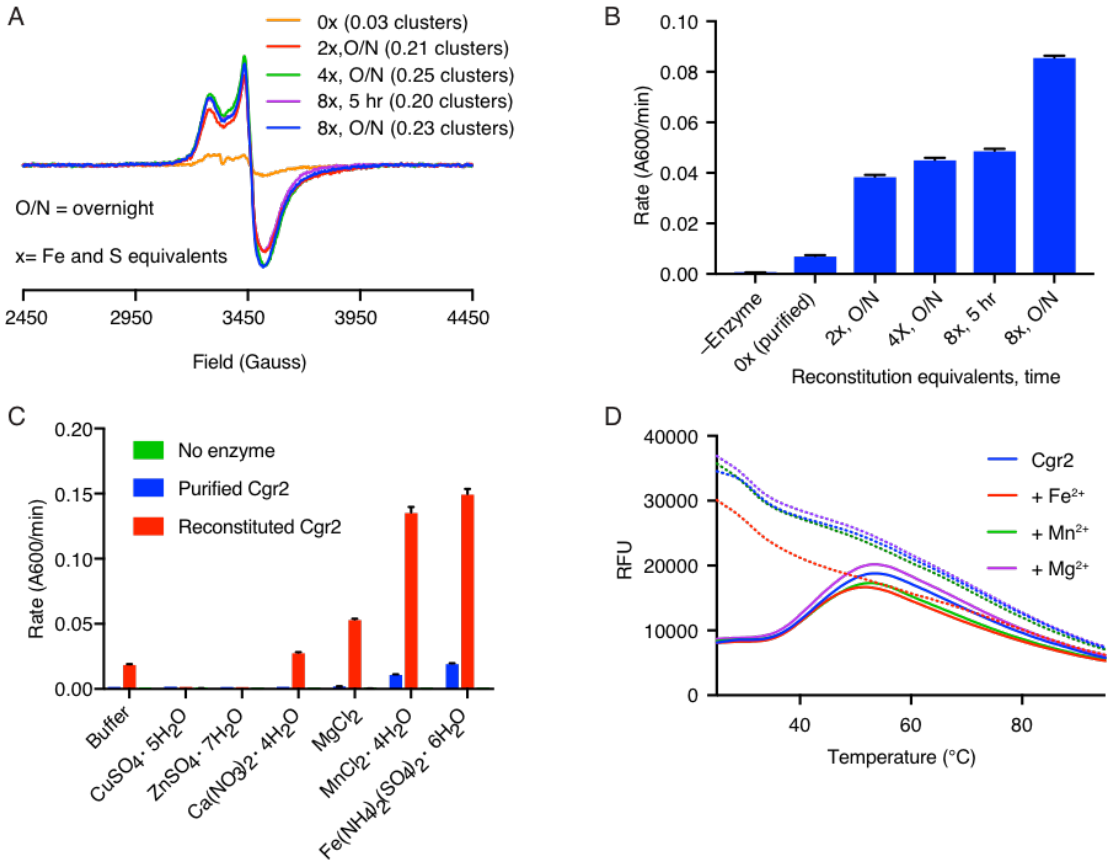


Figure 38: Metal dependence and [4Fe-4S] cluster content of Cgr2.
(A) EPR spectra of dithionite-treated Cgr2 samples that had been reconstituted with 0 (purified), 2, 4, or 8 equivalents of iron and sulfide for 5 hours or overnight (O/N). Number of EPR-active clusters per Cgr2 monomer under each reconstitution condition is shown in parentheses. Samples contained 150 µM protein, 0.3 mM sodium dithionite, and measurements were conducted at 10 K. Spin quantitation was determined against a 150 µM Cu^{2+}-EDTA standard measured under non-saturating conditions. (B) *In vitro* reaction rates of Cgr2 reconstituted under different conditions revealed increasing activity with higher reconstitution equivalents. Data represents mean ± SEM (n = 3). (C) Divalent metal cations (Fe^{2+}, Mg^{2+}, Mn^{2+}) stimulated the activity of Cgr2 *in vitro*. Data represents mean ± SEM (n = 3). (D) Thermal melt curves displaying relative fluorescence of Sypro Orange bound to purified (dashed line) and reconstituted (solid line) Cgr2 in the presence of divalent metals (50 mM HEPES, 100 mM NaCl, pH 8; 8 equivalents of metals relative to protein). Divalent metal cations that stimulated Cgr2 activity did not affect protein stability.

3.4. Identification of cysteine residues important for Cgr2 activity

As discussed in Section **3.2**, the presence of [Fe-S] cluster(s) in Cgr2 was unexpected, as known cluster binding motifs were not detected within its amino acid sequence using current bioinformatic tools (Table 6). Therefore, we attempted to use site-directed mutagenesis to determine whether any of the cysteine residues of Cgr2 were involved in [4Fe–4S] cluster assembly. Single cysteine to alanine mutants were prepared for all 16 cysteines in the mature Cgr2 peptide and tested for digoxin reduction activity in whole cell assays in *R. erythropolis*. All single point mutants were soluble and were obtained in comparable yields to wild-type Cgr2 (Figure 39A, Figure 39C). Single point mutations in 6 cysteine residues resulted in a significant decrease in dihydrodigoxin production relative to wild-type Cgr2 both in whole cells (Figure 39B) and *in vitro* using purified and reconstituted proteins (Figure 39D). EPR analysis of these six point mutants revealed comparable levels of $[4Fe\text{-}4S]^{1+}$ clusters (0.17 - 0.33 clusters/Cgr2) relative to wild-type (0.23 clusters/Cgr2) (Figure 39E), which may argue against the involvement of these cysteines in cluster ligation. However, substitution of a single [4Fe-4S] cluster ligand is not always sufficient to prevent cluster formation (*36*), and free cysteines within a protein have been shown to complement mutants with incompletely coordinated [Fe-S] clusters (*37, 38*). Finally, as EPR spectroscopy can only detect redox active clusters, we cannot rule out the possibility that a subset of these cysteine residues ligate an additional EPR-silent cluster within Cgr2. Additional studies including Mössbauer spectroscopy and structural characterization (Section **5.2**) will thus be required to definitively determine both the location and stoichiometry of the metal centers in Cgr2.

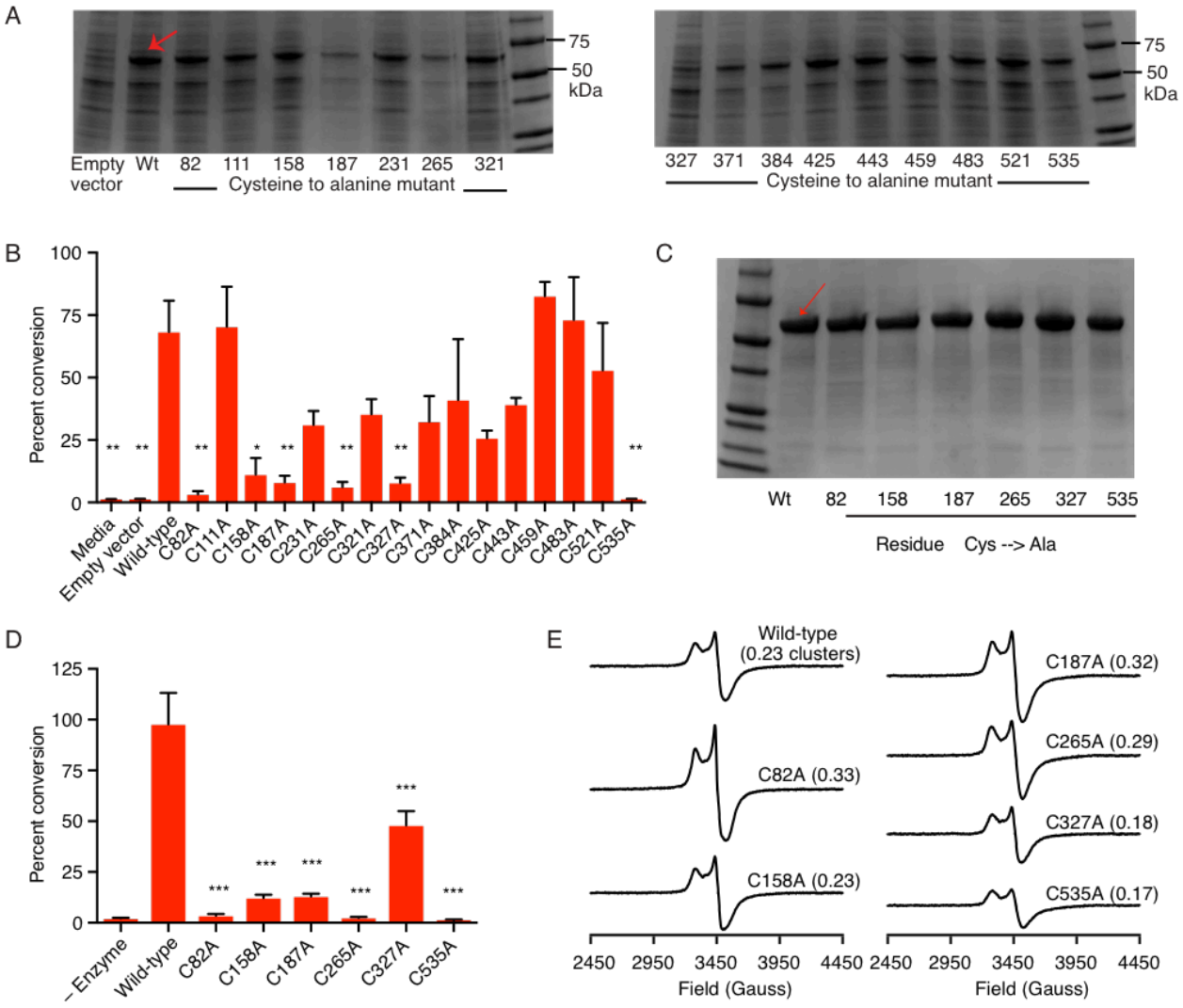


Figure 39: Identification of six cysteine residues important for Cgr2 activity.
(A) SDS-PAGE analysis of clarified lysate from *R. erythropolis* cells transformed with empty pTipQC vector or expressing cytoplasmic wild-type Cgr2 (wt) or individual cysteine to alanine point mutants (~55 kDa). All point mutants were soluble. (B) LC-MS/MS analysis of whole cell assays of *R. erythropolis* expressing individual cysteine to alanine point mutants and incubated with digoxin demonstrated that 6 cysteine residues are important for activity. Data represents mean ± SEM (n = 3). Asterisks indicate statistical significance of each variant compared to wild-type Cgr2 by Student's t test (* $p < 0.05$, ** $p < 0.01$). (C) SDS-PAGE analysis of purified wild-type or single point mutants of Cgr2. (D) LC-MS/MS analysis of *in vitro* assays quenched at 15 min. Data represents mean ± SEM (n = 3). Asterisks indicate statistical significance of each variant compared to wild-type Cgr2 by Student's t test (*** $p < 0.001$). (E) $[4Fe\text{-}4S]^{1+}$ clusters were detected by EPR in all Cgr2 point mutants treated with sodium dithionite. Spin quantitation against a Cu^{2+}-EDTA standard revealed similar levels of $[4Fe\text{-}4S]^{1+}$ clusters per Cgr2 monomer for all variants. Number of clusters shown in parentheses.

As we could not definitively assign a [4Fe-4S] cluster-binding role for the six cysteines identified to be important for Cgr2 activity, we explored whether these residues were involved in other aspects of protein structure or function. We reasoned that these residues could be critical for maintaining protein stability, for example through participation in disulfide bond formation. However, all six point mutants displayed similar (aerobic) melting profiles to wild-type Cgr2 (Figure 40A), ruling out the possibility that protein misfolding or instability was responsible for decreased activity. As we had observed that divalent metal cations enhanced the *in vitro* activity of Cgr2, we next tested whether these cysteine residues might influence binding to another metal center. In fact, Fe^{2+} stimulated the *in vitro* activity of only 3 out of 6 impaired mutants (C158A, C187A, C327A) (Figure 40B). It is possible that the three remaining cysteine residues (C82, C265, C535) could influence divalent metal cation binding in Cgr2, as discussed in more detail in Section **3.7**.

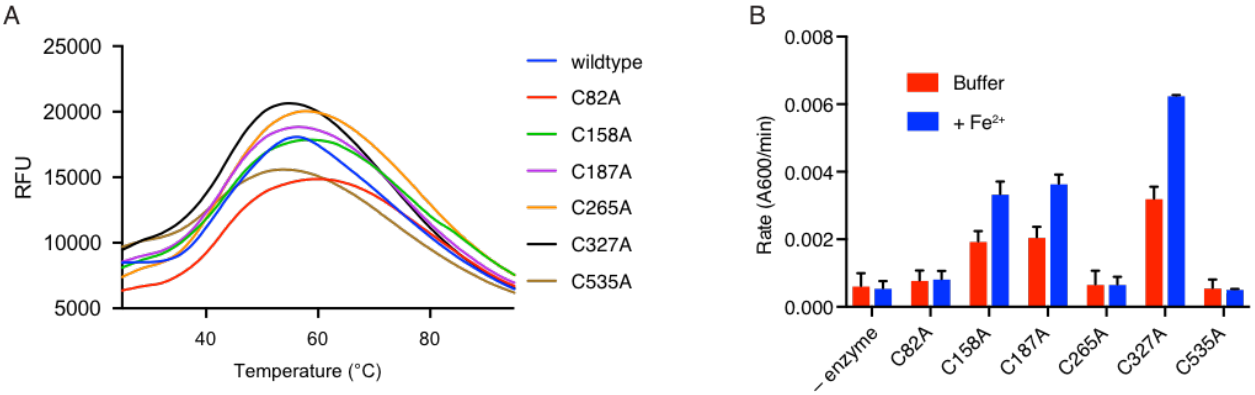


Figure 40: *In vitro* stability and activity of Cgr2 cysteine to alanine point mutants.
(A) Thermal melt curves displaying relative fluorescence of Sypro Orange bound to reconstituted Cgr2 variants (50 mM HEPES, 100 mM NaCl, pH 7). (B) Fe^{2+} stimulated the *in vitro* activity of three cysteine residues, potentially implicating C92, C265, and C535 in metal binding. Data represents mean ± SEM (n = 3).

3.5. FAD dependence of Cgr2

In addition to its metal cofactor dependence, Cgr2 also required FAD for *in vitro* reduction of digoxin (Figure 34C). However, recombinant Cgr2 from *R. erythropolis* did not co-purify with flavin, and various attempts to flavinylate Cgr2 both *in vivo* and *in vitro* did not result in appreciable cofactor binding. This included supplementation of FAD or riboflavin into heterologous expression media, lysis in FAD-containing purification buffers, or dialysis with FAD following purification. Additionally, no significant thermal stabilization was observed upon incubation of the protein with FAD and/or with digoxin (Figure 41A), suggesting that binding of the protein to flavin was minimal. We also explored the possibility that [Fe-S] cluster formation might need to occur either before or concomitantly with flavinylation. However, attempts to reconstitute Cgr2 in the presence of FAD were also unsuccessful, and desalting on a PD-10 column or size exclusion chromatography effectively removed all flavin from Cgr2. Even after Cgr2 reconstitution, no significant thermal stabilization was observed in the presence of FAD (Figure 41B-C), and a substantial excess of FAD was required for maximal *in vitro* activity (Figure 41D), demonstrating a poor binding affinity of the cofactor to Cgr2. We cannot rule out the possibility that *E. lenta* uses an alternative flavin analog *in vivo*, which has higher binding to Cgr2. However, a more likely explanation is that our *in vitro* system may be lacking several key components that could alter the overall protein fold of Cgr2 and increase its affinity for FAD. For example, the FAD binding site in our Cgr2 constructs may be exposed in the absence of fully reconstituted metallocofactor, the presumed binding partner Cgr1, and/or membrane components. Despite this poor binding, FAD was nevertheless essential for conversion of digoxin to dihydrodigoxin by Cgr2 *in vitro*.

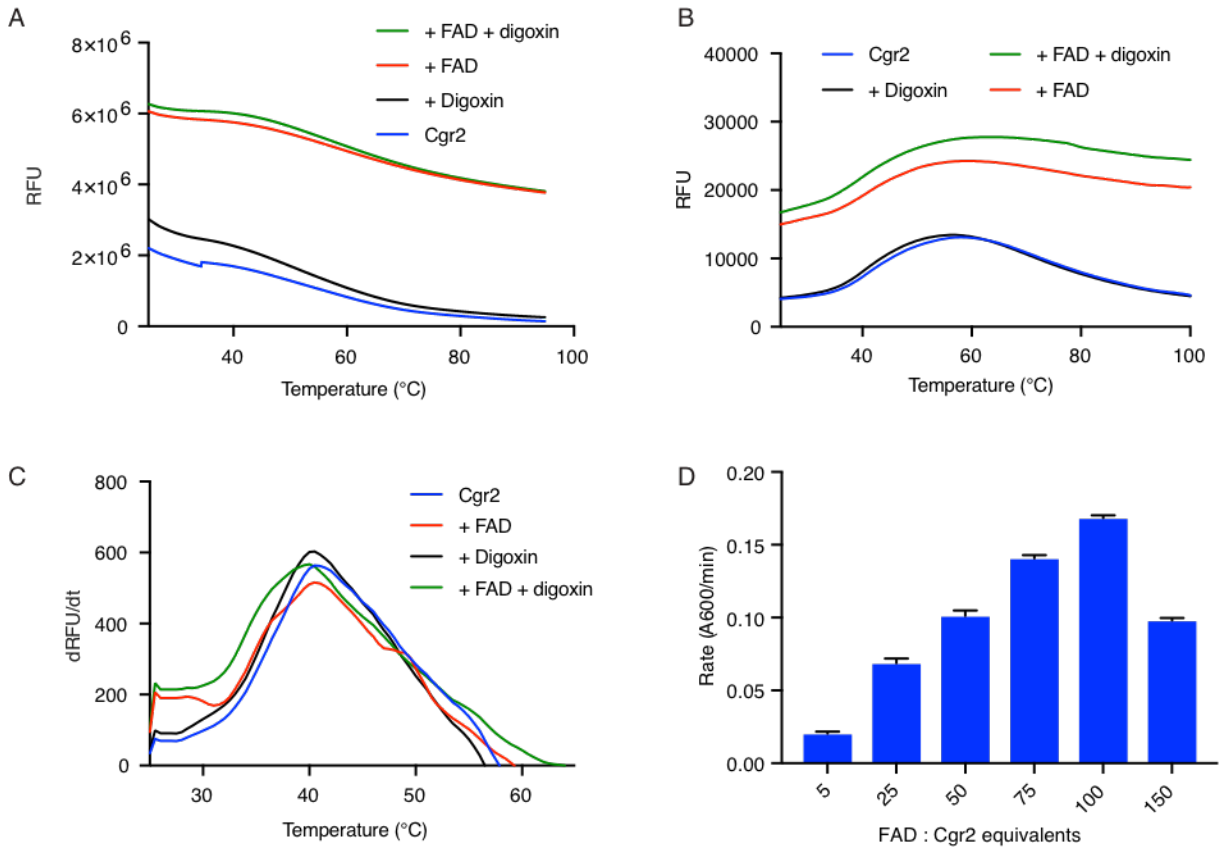


Figure 41: Thermal stability and *in vitro* activity assays reveal that Cgr2 binds FAD poorly. Thermal melt curves displaying relative fluorescence of Sypro Orange bound to (A) purified and (B) reconstituted Cgr2 in the presence of 10x excess of FAD, digoxin, or both (50 mM HEPES, 100 mM NaCl, pH 8). (C) First derivative of fluorescence for reconstituted samples from panel (B). No appreciable thermal stabilization was observed for Cgr2 in the presence of flavin or substrate. (D) Initial *in vitro* rate of digoxin-dependent methyl viologen oxidation, with respect to FAD concentration. Highest rates were observed when using 100 equivalents of FAD (500 µM).

3.6. NMR characterization confirms production of (20*R*)-dihydrodigoxin by Cgr2

After optimizing our *in vitro* assay, we wanted to experimentally confirm that Cgr2 generates the clinically relevant diastereomer of dihydrodigoxin. Unlike catalytic hydrogenation of digoxin which yields a 3:1 ratio of 20*R*:20*S* dihydrodigoxin diastereomers, only the 20*R* diastereomer was detected in *E. lenta* culture media and in the urine of patients (*39, 40*). We thus performed a large scale (2 mg) *in vitro* Cgr2-mediated digoxin reduction reaction and analyzed the extracted product by ^{1}H NMR spectroscopy. Comparison of the substrate (digoxin) (Figure 42) and product (dihydrodigoxin) (Figure 43) spectra in deuterated methanol revealed a clear disappearance of the vinylic proton peak (H22, 5.9 ppm, singlet, 1H) following *in vitro* incubation with Cgr2. In addition, new peaks were detected in the product spectra corresponding to H20 (2.78 ppm, 1H), H21 (4.5 ppm and 4.04 ppm, apparent triplets, 2H), and H22 (2.45 ppm and 2.54 ppm, doublet of doublets, 2H) positions (Figure 44). Analysis of the product by homonuclear correlation spectroscopy (COSY) revealed correlations between proton resonances at H20 and H21, H20 and H22, and H20 and H17 (Figure 45). These data support our assignment of the new ^{1}H NMR proton peaks, and demonstrate that Cgr2 only produces one diastereomer of dihydrodigoxin. We also collected ^{1}H NMR spectra of the product dissolved in deuterated acetone to directly compare the resonances to published work (*41*) and determine which diastereomer was generated by Cgr2. The product spectra only had peaks corresponding to H21 (4.01 and 4.41 ppm) and H22 (2.35 and 2.49 ppm) of the 20*R* diastereomer, and lacked peaks corresponding to H21 (4.05 and 4.35 ppm) and H22 (2.29 and 2.53 ppm) of the 20*S* diastereomer (Figure 46). All together, these data demonstrate that Cgr2 only produces the clinically relevant *20R* diastereomer of dihydrodigoxin

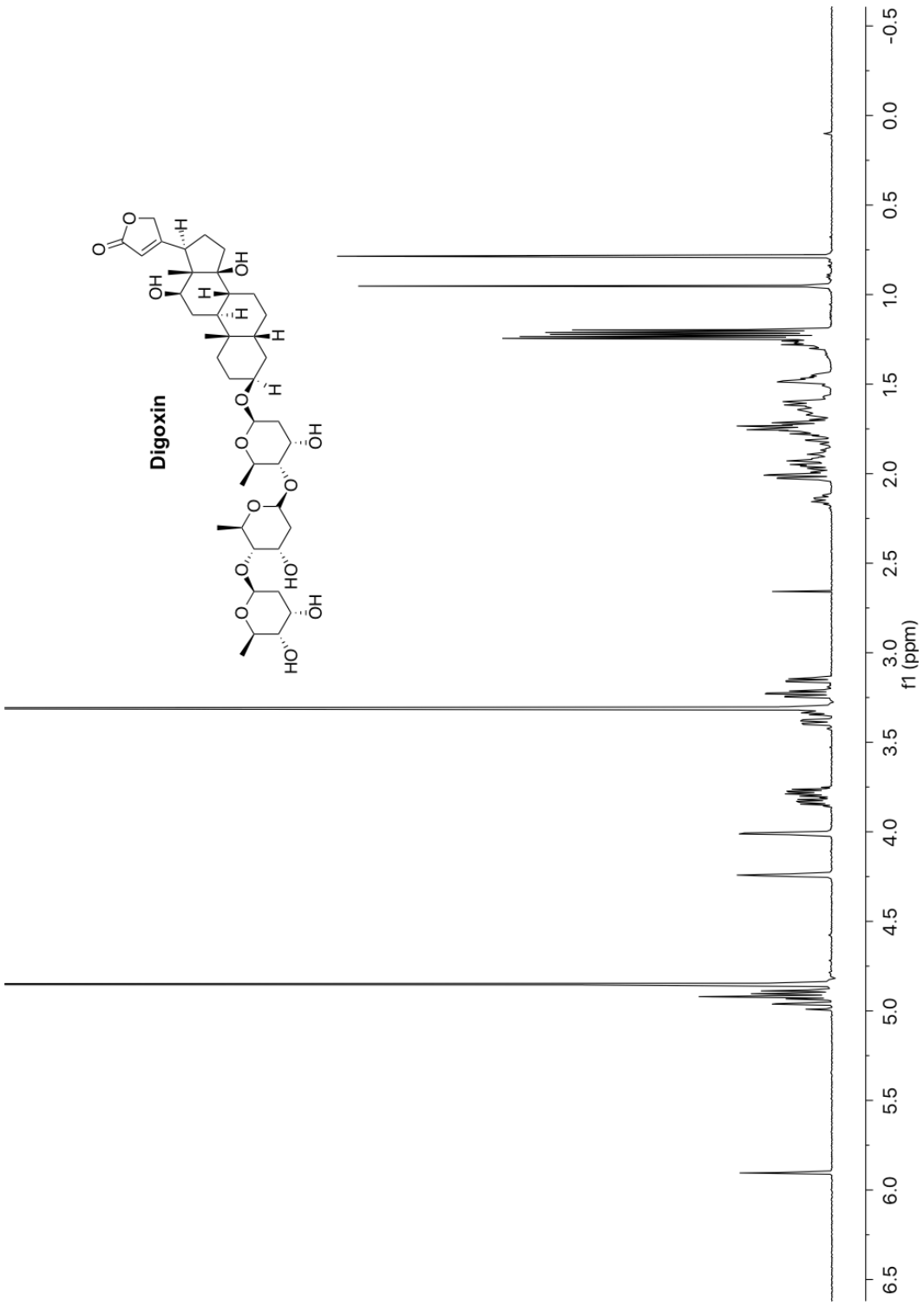


Figure 42: ^{1}H NMR spectrum of digoxin (recorded in CD_3OD at 600 MHz).

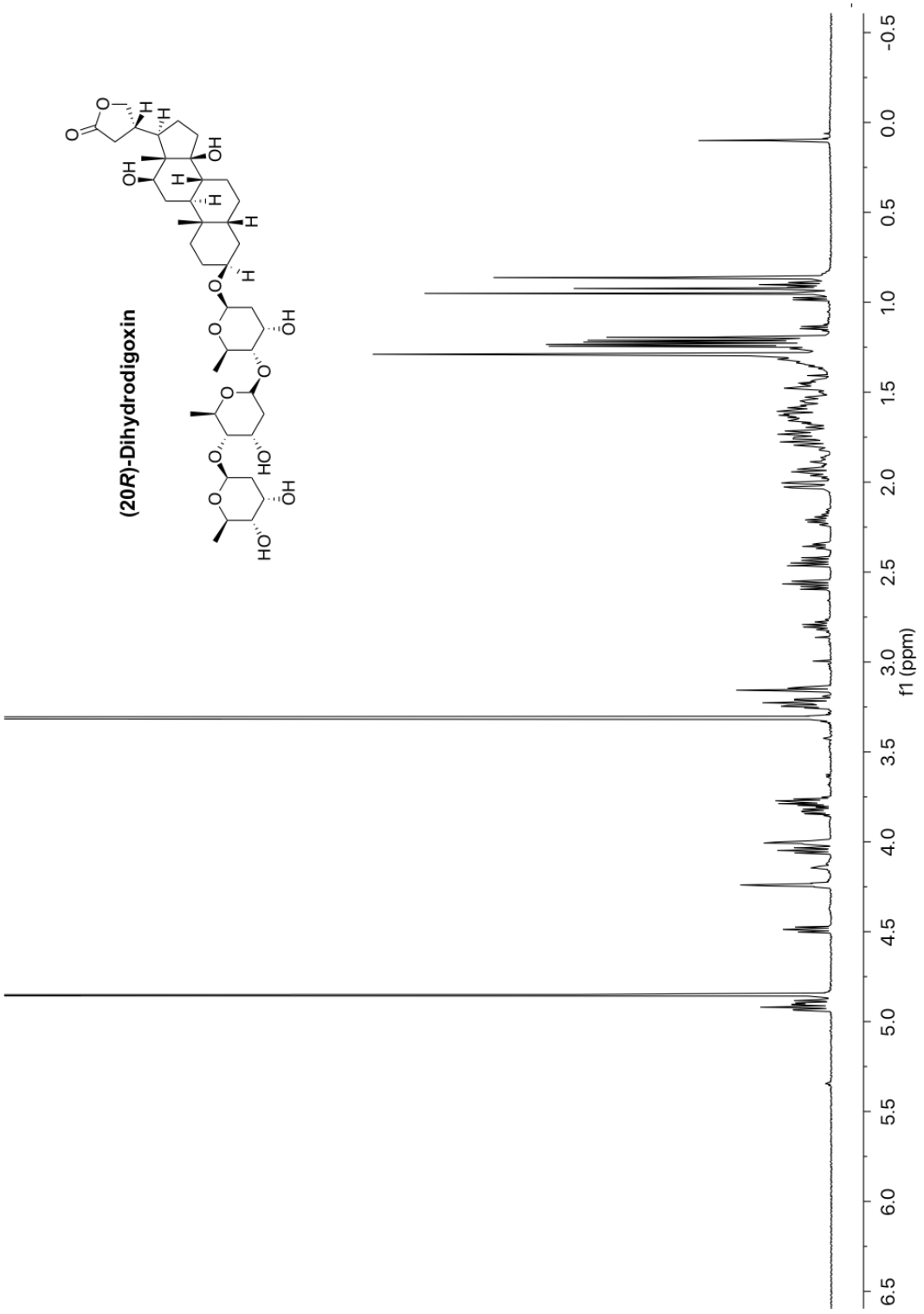


Figure 43: ^{1}H NMR spectrum of dihydrodigoxin generated by Cgr2 (recorded in CD_3OD at 600 MHz).

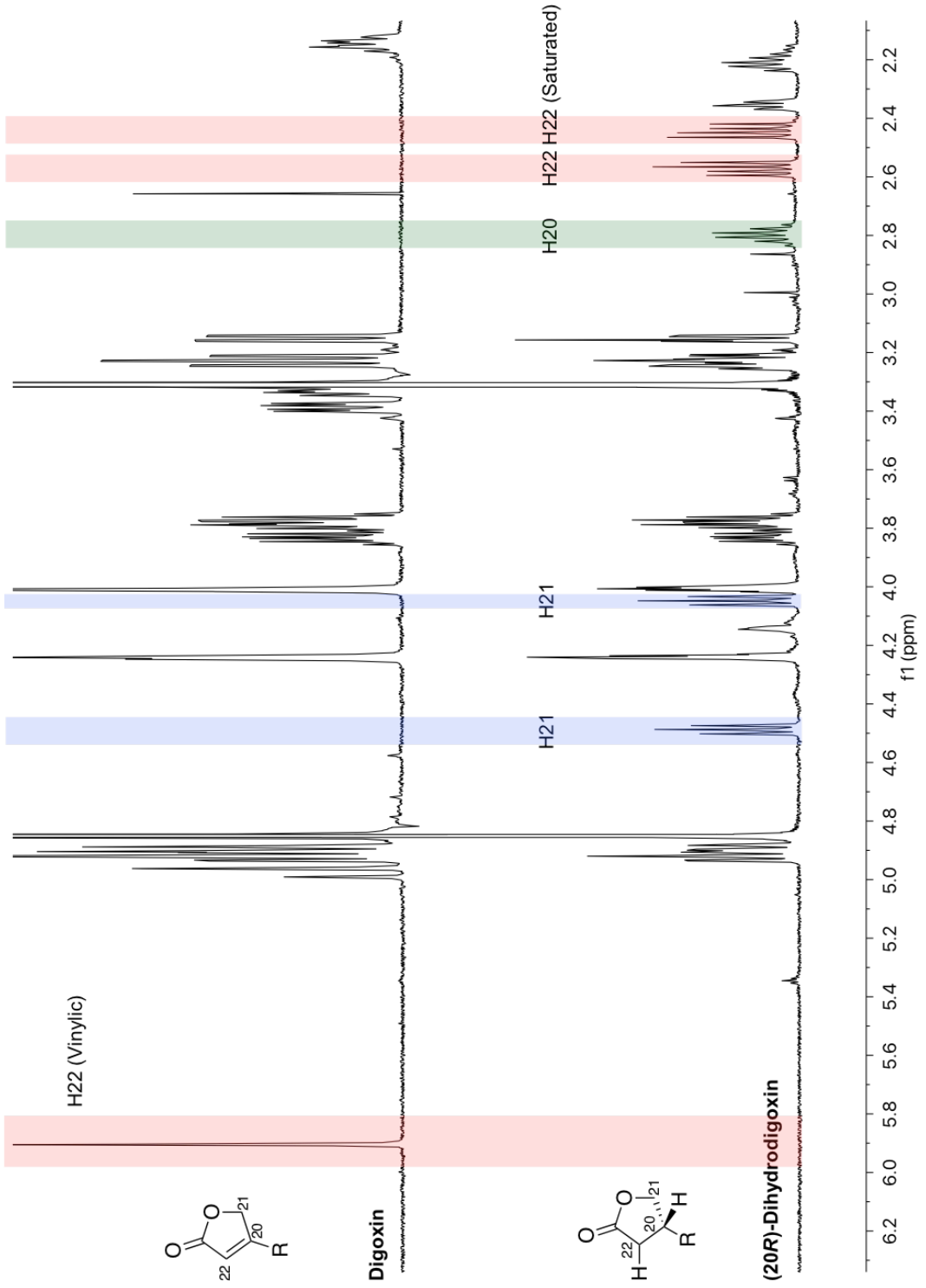


Figure 44: Overlay of ^{1}H NMR spectra of digoxin and dihydrodigoxin generated by Cgr2 (recorded in CD_3OD at 600 MHz) highlighting key protons in the lactone moiety.

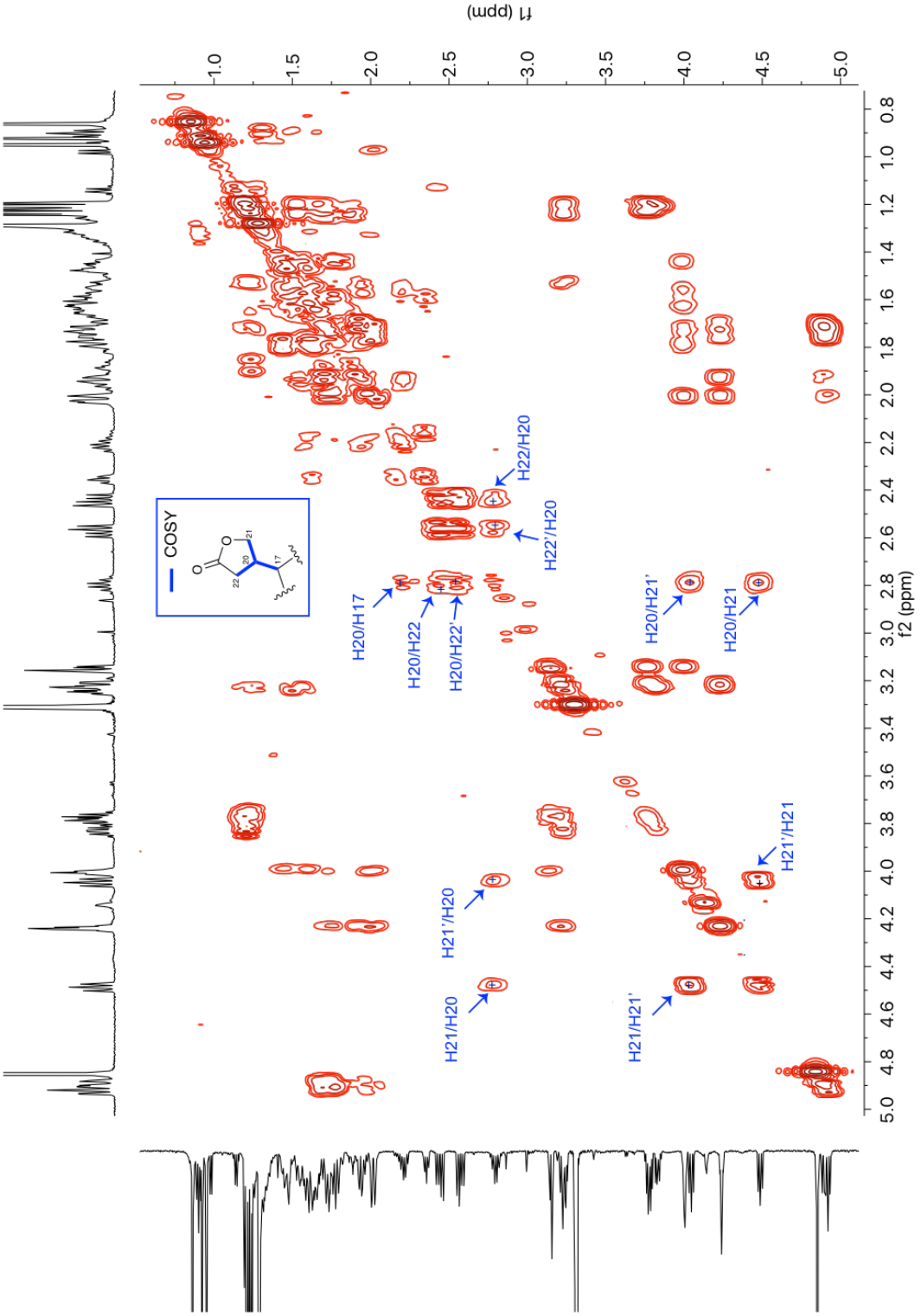


Figure 45: COSY spectra of dihydrodigoxin (recorded in CD_3OD at 600 MHz).

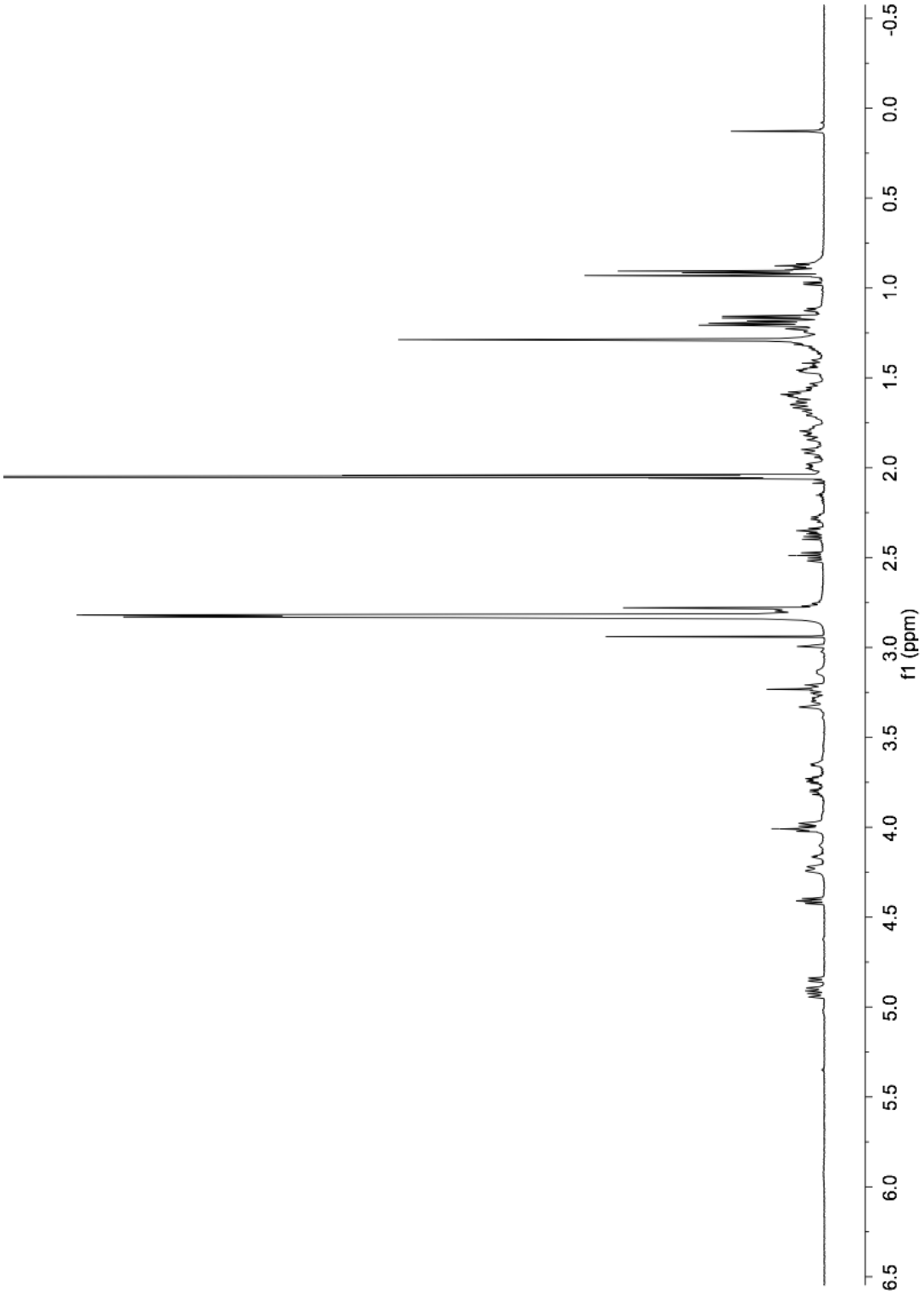


Figure 46: ^{1}H NMR spectrum of dihydrodigoxin generated by Cgr2 (recorded in CD_3COCD_3 at 600 MHz).

3.7. Discussion

In this Chapter, we unambiguously show that the *E. lenta* protein Cgr2 metabolizes the cardiac drug digoxin into the clinically inactive metabolite (20*R*)-dihydrodigoxin. This unique reductase requires an FAD cofactor, contains oxygen-sensitive, redox-active [4Fe-4S] clusters, and is stimulated by the addition of divalent metal cations *in vitro*. Based on their resemblance to other characterized bacterial reductase systems including anaerobic fumarate, sulfite, and nitrate reductases (*42-44*), our working model is that Cgr1 and Cgr2 form a membrane-anchored, extracellular complex that mediates electron transfer from an electron donor (e.g. the membrane quinone pool) through multiple cytochromes *c* in Cgr1 to Cgr2, which ultimately reduces the α,β-unsaturated γ-butyrolactone of digoxin (Figure 47A). We propose that the [4Fe-4S] cluster(s) of Cgr2 sequentially transfer electrons to the FAD cofactor. A resulting hydride equivalent is then transferred from the reduced flavin cofactor to the cardenolide, and proton transfer generates a fully reduced γ-butyrolactone (Figure 47B), yielding the therapeutically inactive metabolite dihydrodigoxin.

Bioinformatic analyses were unable to predict the presence of [4Fe-4S] cluster(s) of Cgr2, although these metalloclusters were essential both for protein stability and activity. Intriguingly, Cgr2 contains 16 cysteine residues, some of which are likely involved in [Fe-S] cluster ligation. Although we identified six residues that were important for both *in vivo* and *in vitro* activity of Cgr2, mutation of these residues to alanine did not decrease [4Fe-4S] cluster content as measured by EPR. These results suggest the presence of non-canonical [4Fe-4S] cluster binding motif(s) within Cgr2 that perhaps involve alternative amino acids (histidine, aspartate, serine, or backbone amides) (*29*). Through our attempts to optimize [4Fe-4S] cluster reconstitution, we also observed that divalent metal cations (Mg^{2+}, Mn^{2+}, and Fe^{2+}) stimulated the activity of wild-type Cgr2 *in vitro*, although the mechanistic basis of this observation is still unclear. Addition of Fe^{2+} to 3 of the 6 impaired cysteine to alanine Cgr2 mutants led to a significant increase in activity, while the remaining three mutants (C92A, C265A, and C535A) were unaffected by metal addition. This observation prompted the hypothesis that these latter three residues may influence divalent metal cation binding. In other protein, cysteine residues have been shown to bind

Mn^{2+}, but not Mg^{2+} centers (*45*), and direct coordination of Fe^{2+} centers would be highly unusual. A more likely scenario is that these cysteine residues could influence the substrate binding pocket or overall protein structure, rather than directly coordinating Fe^{2+} (*46*). Several key questions remain regarding the metal requirements of Cgr2, including: What are ligands and stoichiometry of the [4Fe-4S] cluster and divalent metal centers? Which metal (Mg^{2+}, Mn^{2+}, and Fe^{2+}) is endogenously found in Cgr2? And, what is the role of the divalent metal in digoxin metabolism. Ultimately, further structural characterization and spectroscopy (e.g. Mössbauer), or native purification of Cgr2 from *E. lenta* will be required to understand the composition and function of these metal centers in more detail.

Currently, further mechanistic studies of Cgr2 are hampered by a lack of homology to characterized enzymes and a divergent set of predicted active site residues (Section **2.2**). In addition, the remarkable conservation of all known *cgr2* sequences (only two known mutations) makes it challenging to narrow down which of the 560 amino acid residues are essential for activity. Structural characterization will thus be essential for determining which amino acids in Cgr2 are involved in metallocofactor binding, substrate binding, and catalysis. However, this conservation may also prove useful, as *cgr2* has now been validated as a biomarker that can be used to further study digoxin metabolism in environmental microbes, human microbiotas, and possibly in clinical settings.

While many mechanistic details of Cgr2's activity remain to be elucidated, we have shown that this unique and complex reductase is responsible for digoxin metabolism. *In vivo*, an active Cgr complex would presumably require coordination of many energetically costly processes including protein synthesis, flavinylation and [Fe-S] cluster assembly (Cgr2), cytochrome *c* maturation (Cgr1), and insertion or translocation across the bacterial membrane (Cgr1 and Cgr2). While the clinical significance of digoxin metabolism is obvious for patients, the physiological benefit associated with this energetically costly metabolic pathway in *E. lenta* remains unknown.

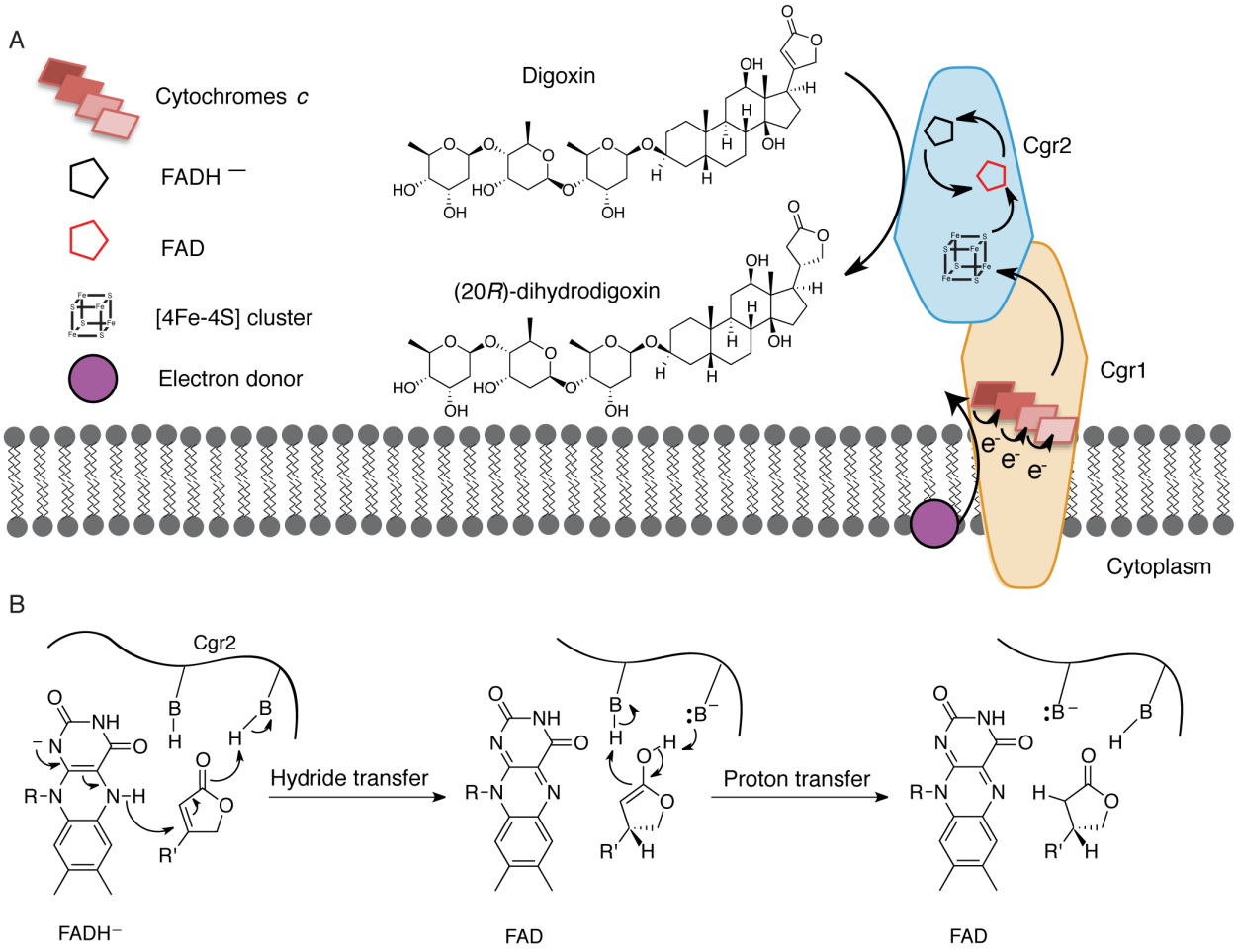


Figure 47: Preliminary model for digoxin metabolism by Cgr1 and Cgr2.
(A) Proposed biochemical model and (B) mechanism of digoxin reduction by Cgr proteins. Cgr1 is predicted to transfer electrons from a membrane-associated electron donor to at least one $[4Fe\text{-}4S]^{2+}$ cluster of Cgr2 via covalently bound heme groups. The reduced $[4Fe\text{-}4S]^{1+}$ cluster of Cgr2 could sequentially transfer two electrons to FAD, generating $FADH^-$, which could mediate hydride transfer to the β-position of the digoxin lactone ring. Protonation of the resulting intermediate would yield (20*R*)-dihydrodigoxin.

3.8. Experimental

Construction of *cgr1* and *cgr2* vectors for heterologous expression and purification

Cgr constructs and plasmids were amplified, digested, and ligated as described in Section **2.6**. The new *cgr* constructs and pMAL-c2x vector (Addgene) were processed as previously described, using restriction enzymes (New England Biolabs) listed in Table 4.

Table 4: Primers for heterologous expression and purification of Cgr1 and Cgr2.
Restriction sites are bolded.

Construct	For/ Rev	Sequence	Restriction sites	Vector, host	Annealing Temp (°C)	Extension time (sec)
MBP-Cgr1 (-39aa)	For	CATAA**GAATTC**CACGAGC AGCCGAGCT	EcoRI, PstI	pMAL-c2x, *E. coli*	65	30
	Rev	GTTTTA**CTGCAG**TTACGC CGCCGTCGAA				
MBP-Cgr2 (-48aa)	For	ACAATA**GAATTC**CAGACC GCGCCTGC	EcoRI, PstI	pMAL-c2x, *E. coli*	61	60
	Rev	TGTAT**CTGCAG**TTACTCCC ACGGCTCGA				
Cgr2-CHis$_6$	For	ACTGAC**CCATGG**GCATGG AATACGGAAAGTGCC	NcoI, HindIII	pTipQC1, *R. erythropolis*	65	60
	Rev	GTTAG**AAGCTT**CTCCCAC GGCTCGAG				
Cgr2 (-48aa)-NHis$_6$	For	TATTA**CCATGG**ATCAGAC CGCGCCTGC	NcoI, HindIII	pTipQC2, *R. erythropolis*	65	60
	Rev	ATACT**AAGCTT**CTCCCAC GGCTCGA				
Cgr2 (-48aa)-CHis$_6$	For	TATTA**CCATGG**ATCAGAC CGCGCCTGC	NcoI, HindIII	pTipQC1, *R. erythropolis*	65	60
	Rev	ATACT**AAGCTT**TTACTCCC ACGGCTCGA				

Heterologous expression of Cgr proteins in *E. coli*

The Cgr-MBP pMAL-c2x vectors were transformed into either Tuner (DE3) (Novagen) or BL21 (New England Biolabs) chemically competent *E. coli* cells. Cgr1 constructs were co-transformed with pEC86 (provided by Dr. Linda Thöny-Meyer), a vector containing the *ccmABCDEFGH* genes required for maturation of cytochrome *c* proteins. Transformed *E. coli* cells were plated on LB agar plates containing ampicillin (pMAL-c2X) or ampicillin and chloramphenicol (pMAL-c2x + pEC86). For Cgr2 constructs, single colonies were inoculated into 50 mL of LB (Lennox) broth (Alfa Aesar) and grown

overnight at 37 °C. Cgr1 constructs were similarly grown in 2xYT media (*1*) which contains higher levels of iron and heme precursors. Approximately 20 mL of starter culture was added to 2 L of the respective growth media + antibiotics and grown at 37 °C, 175 rpm. Upon reaching an OD_{600} of ~0.6, cultures were induced with 30 μM IPTG (Sigma-Aldrich). Protein expression was carried out overnight at 15 °C, 175 rpm and cells were harvested by centrifugation (10,800 rpm x 20 min). Cell pellets were immediately resuspended in 30 mL TES buffer (100 mM Tris, 20% sucrose, 0.5 mM EDTA, pH 8) containing Sigma Fast protease inhibitor cocktail (Sigma-Aldrich) and 0.5 mg/mL lysozyme (Sigma-Aldrich), and incubated at room temperature for 15 minutes. 30 mL of ice-cold water was added to cell suspensions and shaken on ice for 15 minutes at 100 rpm. The suspension was centrifuged at 13,000 rpm for 15 minutes.

Purification of MBP-Cgr fusion proteins and cleavage of MBP tag with Factor Xa protease

The supernatant containing periplasmic protein was diluted 3x with ice-cold column buffer (20 mM Tris, 200 mM NaCl, pH 8) to reduce viscosity and loaded onto a column containing 2 mL of amylose resin (New England Biolabs). The column was washed with 20 mL of column buffer and MBP-fusion proteins were eluted with 20 mL of elution buffer (column buffer + 10 mM maltose). Protein was concentrated using a 20 mL Spin-X UF 30k MWCO PES spin filter (Corning) and dialyzed into Factor Xa cleavage buffer (20 mM Tris, 100 mM NaCl, 2 mM $CaCl_2$, pH 8). Cgr-MBP fusion proteins were incubated with Factor Xa protease (New England Biolabs) at a 50:1 ratio for 2.5 - 21 hours at 4 °C. Reactions were loaded on an amylose resin column as previously described, releasing untagged Cgr proteins in the flow through. Samples were analyzed by SDS-PAGE analysis using a 4–15% Mini-PROTEAN® TGX™ Precast Gel (Bio-Rad) and Precision Plus Protein™ standards (Bio-Rad).

Heme-peroxidase gel stain assay

Cgr1-overexpressing strains were lysed and purified as previously described (Section **2.6**). Samples were analyzed by SDS-PAGE on a 4–15% Mini-PROTEAN® TGX™ Precast Gel (Bio-Rad), using a 13 kDa cytochrome *c* from bovine heart (Sigma-Aldrich) as a positive control. The gel was

transferred to a PVDF membrane (Thermo Fisher Scientific) (200 mA, 60 minutes). The membrane was washed in PBS buffer (Sigma-Aldrich), and incubated with 1 mL each of Supersignal West Femto kit Buffer 1 (Luminol) and Buffer 2 (hydrogen peroxide) for at least 5 minutes (Thermo Fisher Scientific). The membrane was imaged on a Chemi Imager (Bauer Core, Harvard University) using a 30 second exposure and a high sensitivity/low resolution setting.

Purification and [Fe-S] cluster reconstitution of Cgr2 from *R. erythropolis*

Culturing, heterologous expression, and LC-MS/MS analyses were described in Section **2.6**. All protein purification steps were carried out at 4 °C. Harvested cells were resuspended in 5 mL per g of cell pellet in lysis buffer (50 mM Tris, pH 8, 1 mM $MgCl_2$, 25 mM imidazole) containing Pierce EDTA-free protease inhibitor cocktail (Thermo Fisher Scientific). Cells were passaged through a cell disruptor (Avestin EmulsiFlex-C3) five times at 15,000-25,000 psi and centrifuged for 20 minutes at 13,000 rpm. The clarified lysate was incubated on a nutating mixer with 5-10 mL of HisPur Ni-NTA resin (Thermo Fisher Scientific) for 1 hour and then applied to a gravity flow column. The resin was washed with 50 mL of wash buffer (25 mM HEPES, 0.5 M NaCl, pH 8, 25 mM imidazole) and eluted with 25 mL of elution buffer (25 mM HEPES, 0.5 M NaCl, pH 8, 200 mM imidazole). Eluted protein was concentrated using a 20 mL Spin-X UF 30k MWCO PES spin filter (Corning) to a volume of 1-2.5 mL, and then desalted on a Sephadex G-25 PD-10 desalting column (GE Healthcare) that had been equilibrated with desalting buffer (50 mM HEPES, 100 mM NaCl, pH 8). Desalted protein was sparged with argon on ice for 30 - 45 minutes. Chemical reconstitution of [Fe-S] cluster(s) in Cgr2 was carried out at 4 °C in an anaerobic chamber (Coy Laboratory Products) under an atmosphere of 2% hydrogen and 98% nitrogen. A 30 µM solution of Cgr2 was prepared in reconstitution buffer (50 mM HEPES, 100 mM NaCl, pH 8, and 2 mM dithiothreitol (DTT)). $Fe(NH_4)_2(SO_4)_2 \cdot 6H_2O$ (Sigma-Aldrich) was added in four aliquots over 60 minutes, followed by addition of $Na_2S \cdot 9H_2O$ (Sigma-Aldrich) in four aliquots over 60 minutes to final concentrations of 0.24 or 0.375 mM (8 or 12.5 equivalents relative to Cgr2), and stirred for 16-24 hours. The reaction was filtered through a 0.25 mm, 0.2 µM pore-size PES syringe filter (VWR) to remove

precipitant and concentrated in a 6-mL Spin-X UF 30k MWCO PES spin filter inside a 50 mL conical-bottom centrifuge tube with plug seal cap (Corning). The concentrated protein (1-2.5 mL) was desalted on a PD-10 column into desalting buffer. Protein was aliquoted into 0.5 mL PP conical tubes with skirt (Bio Plas), sealed inside 18 x 150 mm Hungate tubes (Chemglass Life Sciences) and stored at –80 °C. Protein concentration was determined by Bradford using Protein Assay Dye Reagent (Bio-Rad) and bovine serum albumin (BSA) (Sigma-Aldrich) as a reference standard. Typical protein yields were ~20 mg/L of culture for both wild-type and point mutants of Cgr2(–48aa)-NHis6x, ~20 mg/L of Cgr2(–48aa)-TEV-NHis6x, ~8 mg/L for Cgr2(–48aa)-Chis6x, and ~1 mg/L for Cgr2-Chis6x. The iron and sulfur content of Cgr2 samples (protein concentrations between 20-50 μM) was determined using previously reported colorimetric assays (*32*).

TEV cleavage to remove His6x tag from Cgr2

Cgr2(–48aa)-TEV-NHis6x was incubated with TEV protease (New England Biolabs) at a 50:1 ratio for 15 hours at 4 °C. Reactions were loaded on a column containing HisPur Ni-NTA resin (Thermo Fisher Scientific) as previously described, to release untagged Cgr2 in the flow through. Samples were analyzed by SDS-PAGE analysis and size exclusion chromatography.

Site-directed mutagenesis of Cgr2

Site-directed mutagenesis was performed in 25 μL reactions using 200 ng of template DNA (Cgr2(–48aa)-$NHis_6$ in pTipQC2), 0.5 μM of each primer pair (Table 5), 0.5 mM dNTP, and 1 μL of Pfu Turbo polymerase AD (VWR). The following thermocycling parameters were used: denaturation at 95 °C for 1 minute; 18 cycles of 95 °C for 30 seconds, 65 °C for 50 seconds, and 68 °C for 22 minutes (2 min/kb); and a final extension at 68 °C for 7 minutes. The template plasmid was digested with 1 μL of DpnI (New England Biolabs) for 1 hour at 37 °C, and 2 μL of the reaction were transformed into chemically competent One Shot Top10 *E. coli* cells (Thermo Fisher Scientific).

Table 5: Primers for site-directed mutagenesis of Cgr2.
Residue numbering based on native Cgr2. Introduced mutations are bolded.

Mutant	F/R	Sequence
C82A	For	CAGCGGCGGCACG**GCC**GCGGCCATCG
	Rev	CCTCGATGGCCGC**GGC**CGTGCCGCCG
C111A	For	GCGGCAACTCGGCACTA**GCC**GGTGGATACAT
	Rev	CCAGCATGTATCCACC**GGC**TAGTGCCGAGTTG
C158A	For	ATATGATCCGCGAGGCG**GCC**TTGCGCTCCGGC
	Rev	GCCTCGCCGGAGCGCAA**GGC**CGCCTCGCGGAT
C187A	For	GCCCCCGGTCTGGTCA**GCC**GGCGACACGG
	Rev	GGCCTCCGTGTCGCC**GGC**TGACCAGACCGG
C231A	For	CGAAATCGAGATGGGC**GCC**GAGGTGGCGCAC
	Rev	GATGTGCGCCACCTC**GGC**GCCCATCTCGAT
C265A	For	GGCGTGGTCATGGCG**GCC**GCTTCGGTGGA
	Rev	GTTGTCCACCGAAGC**GGC**CGCCATGACCA
C321A	For	GATCGGTGCTGAGCTT**GCC**ATGCAGCAGGC
	Rev	CACGGCCTGCTGCAT**GGC**AAGCTCAGCACC
C327A	For	CATGCAGCAGGCCGTG**GCC**ATGAACGATTCT
	Rev	GATAGAATCGTTCAT**GGC**CACGGCCTGCTG
C371A	For	GACCGGCAGACGGTTT**GCC**CAGGACGATGCCG
	Rev	CTCGGCATCGTCCTG**GGC**AAACCGTCTGCC
C384A	For	CTATGTCATGCACGAG**GCC**GCGCAAGCTGCA
	Rev	CCATGCAGCTTGCGC**GGC**CTCGTGCATGAC
C425A	For	CATACGCCCGACACG**GCC**GATACTACGTTC
	Rev	CGAGAACGTAGTATC**GGC**CGTGTCGGGCGT
C443A	For	GCCGAGTTTATCGGC**GCC**GATCCGACCGC
	Rev	GAGGGCGGTCGGATC**GGC**GCCGATAAACTC
C459A	For	GAGGTGGAACTCTTTC**GCC**GAGGCCGGTTTG
	Rev	CATCCAAACCGGCCTC**GGC**GAAAGAGTTCCA
C483A	For	GACGCCGCCGTTCTAC**GCC**GATGTCGTGCGC
	Rev	GGGGCGCACGACATC**GGC**GTAGAACGGCGG
C521A	For	CTGTACGGCGCCGGG**GCC**ATCATCGGGGGT
	Rev	GTTACCCCCGATGAT**GGC**CCCGGCGCCGTA
C535A	For	GCCTTCTACTTCGGC**GCC**GGCTGGTCCATC
	Rev	CGTGATGGACCAGCC**GGC**GCCGAAGTAGAA
Y333N	For	GCATGAACGATTCTATC**AAC**GTAGGCGGCATCA
	Rev	TCGCTGATGCCGCCTAC**GTT**GATAGAATCGTTCA
Y532F	For	GATGCCGAGTGGGGC**TTT**GTCATGCACG
	Rev	GCACTCGTGCATGAC**AAA**GCCCCACTCG
G536A	For	TTCTACTTCGGCTGC**GCC**TGGTCCATCA
	Rev	GTTCGTGATGGACCA**GGC**GCAGCCGAAG

Thermal denaturation assay

Thermal denaturation assays of purified and reconstituted Cgr2 were prepared on ice in 0.2 mL skirted 96-well PCR plates (VWR) sealed with optical adhesive covers (Life Technologies). Each reaction contained 10 μg of purified or reconstituted Cgr2, Sypro Orange protein gel stain (Thermo Fisher Scientific) diluted 5000-fold, and buffer containing 100 mM buffering agent and 100 mM NaCl in a total volume of 30 μL. The following buffering agents were used: acetate/acetic acid for pH 4-6, HEPES for pH 7, Tris-HCl for pH 8-9, and glycine-NaOH for pH 10. For metal binding assays, metal salts (Sigma-Aldrich) were dissolved in pH 8 buffer to generate 100 mM stock solutions and added to a final concentration of 48 μM to (8 equivalents relative to Cgr2). Data was collected on a CFX96 Touch Real-Time PCR machine (Bio-Rad) using the "FRET" filter setting with FAM excitation and HEX emission channels (485 nm and 556 nm respectively). The following temperature-scan protocol was used: 25 °C for 30 seconds, then ramp from 25 °C to 100 °C at a rate of 0.1 °C/ sec. Reconstituted samples were prepared anaerobically and immediately analyzed upon removal from anaerobic chambers.

Size exclusion chromatography

Gel filtration experiments were carried out on a Superdex 200 10/300 GL column (GE Healthcare) attached to a BioLogic DuoFlow chromatography system (Bio-Rad). Experiments were carried out either aerobically or anaerobically inside an anaerobic chamber (Coy Laboratory Products). 100 μL protein samples (50-100 μM) were loaded onto the column at a rate of 0.2 mL/min for 1 mL followed by an isocratic flow of 0.33 mL/min for 30 mL with 50 mM HEPES, 100 mM NaCl, pH 8. The molecular weight for Cgr2(–48aa)-NHis$_6$ is 55.7 kDa. A gel filtration standard (Bio-Rad) containing thyroglobulin (670 kDa), γ-globulin (158 kDa), ovalbumin (44 kDa), myoglobin (17 kDa), and vitamin B12 (1.35 kDa) was used to determine the molecular weight of Cgr2-containing peaks.

UV-vis spectroscopy

Cgr2 was diluted to a final concentration of 50-100 μM in UV-Star UV-transparent 96-well microplates (Greiner Bio-One). The absorbance was measured between 250 - 750 nm using a PowerWave HT Microplate Spectrophotometer (BioTek) inside of an anaerobic glovebox (Mbraun). Curves were baseline subtracted using respective absorbance values at 700 nm. To determine whether the [Fe-S] cluster(s) were redox active, Cgr2 was incubated with 10 equivalents of sodium dithionite (Sigma-Aldrich) for 15 minutes at room temperature prior to taking additional absorption spectra. To assess the oxygen sensitivity of [Fe-S] cluster(s), Cgr2 was taken out of the anaerobic chamber and exposed to oxygen, and the absorption spectra was measured aerobically on a PowerWave HT Microplate Spectrophotometer (BioTek). Oxygen-exposed Cgr2 was then sparged for 30 minutes with argon (on ice) and brought back into the Mbraun glovebox for activity assays.

EPR spectroscopy

All samples were prepared in 50 mM HEPES, 100 mM NaCl, pH 8 under oxygen-free conditions in an anaerobic glovebox (Mbraun). For all EPR experiments the final concentration of Cgr2 was either 150 or 200 μM. When required, the samples were reacted with an excess of sodium dithionite (10-20 equivalents) for 20-30 minutes at 22 °C prior to freezing in liquid N_2. Spin quantification was carried out against a Cu^{2+}-EDTA standard containing an equimolar concentration of $CuSO_4$ in 10 mM EDTA (150 or 200 μM), under non-saturating conditions. Samples (250 μL) were loaded into 250 mm length, 4 mm medium wall diameter Suprasil EPR tubes (Wilmad LabGlass) and frozen in liquid N_2 under oxygen-free conditions. EPR spectra were acquired on a Bruker E500 Elexsys continuous wave (CW) X-Band spectrometer (operating at approx. 9.38 GHz) equipped with a rectangular resonator (TE102) and a continuous-flow cryostat (Oxford 910) with a temperature controller (Oxford ITC 503). The spectra were recorded at variable temperatures between 10-40 K at a microwave power of 0.2 mW, using a modulation amplitude of 0.6 mT, a microwave frequency of 9.38 GHz, a conversion time of 82.07 ms, and a time constant of 81.92 ms.

[Fe-S] cluster bioinformatics

A literature search was performed for known [2Fe-2S], [3Fe-3S], and [4Fe-4S] cluster binding motifs, and is summarized in Table 6 below (*11-27*).

Table 6: [Fe-S] cluster-binding motifs that are not found in Cgr2.

Cluster type	Protein	Motif
4Fe-4S	Dph2	$CX_{103}CX_{123}C$
	Radical SAM superfamily	CX_3CX_2C
	ThiC	CX_2CX_4C
	HmdA	CX_5CX_2C
	Nqo3 subunit of NDH-1	$CX_2CX_3CX_{27}C$
	Formate dehydrogenase; nitrate reductase	$CX_2CX_3CX_{26}C$
	RumA	$CX_5CX_2CX_{71}C$
	[NiFe] hydrogenase	$CX\mathbf{C}CX_{94}CX_4\mathbf{C}X_{28}C$
	Dihydropyrimidine dehydrogenase	$CX_{38}CX_5CXQ$
	CbiX	$MXCX_2C$
	MutY and endonuclease III	$CX_6CX_2CX_5C$
	IspG (aka GcpE)	$CX_2CX_{31}C_5GE$
	p58C	$CX_{79}CX_{16}CX_{39}C$
	FldI	$CX_{38}C$
3Fe-4S	Ferredoxin	$CX_2DX_2C\ldots CP$
2Fe-2S	Rieske proteins (Box I)	CXHCGC
	Rieske proteins (Box II)	CXCHX(S/A/G)X(Y/F)
	Xanthine oxidoreductase	$CX_2C//CXC$
	Human anamorsin	$CX_2CX_7CX_2C$; CX_8CX_2CXC
	Protoporphyrin ferrocheletase	CX_6CXC; $CX_{170}CX_6CXC$
	FdIV	$CX_5CX_2CX_{35}C$ (bacteria); $CX_4CX_2CX_{29\text{-}30}C$ (plant)

In vitro substrate reduction assays

Methyl viologen (paraquat) dichloride hydrate (Sigma-Aldrich) that had been reduced with sodium dithionite was used as an artificial electron donor (*28*) to initiate anaerobic Cgr2-mediated reduction of digoxin in vitro. Assays were carried out at 25 °C in an anaerobic glovebox (Mbraun) under an atmosphere of nitrogen and < 5 ppm oxygen. Reagents were brought into the glovebox as solids or

sparged liquids and resuspended in anoxic buffer inside the chamber: flavin (FAD or FMN) and methyl viologen (MV) were resuspended in 50 mM HEPES, 100 mM NaCl, pH 7 to generate stock solutions of 1 mM and 50 mM respectively; sodium dithionite was resuspended in 50 mM HEPES, 100 mM NaCl, pH 8 to generate a stock solution of 25 mM, and digoxin (Sigma-Aldrich) was dissolved in DMF to generate a stock solutions of 25 mM. The final assay mixture (100 µL) contained 5 µM Cgr2, 50 µM flavin, 0.375 mM MV, 0.25 mM sodium dithionite, and was initiated by addition of 0.5 mM substrate. As only 0.25 mM of reduced MV was obtained under these assay conditions, only half of the substrate could be consumed to produce a maximum of 250 µM product. For metal activation studies, metal salts were dissolved in pH 7 buffer (1 mM) and added to a final concentration of 40 µM. For flavin dependence studies, assays contained 5 µM Cgr2, 1.5 mM MV, 1 mM sodium dithionite, 0.5 mM digoxin, and 25-750 µM FAD (5-150x equivalents). Assays were prepared in a 96-well polysterene microplate (Corning) and activity was continuously monitored by measuring the absorbance at 600 nm on a PowerWave HT Microplate Spectrophotometer (BioTek); a decrease in the absorbance at 600 nm corresponded to MV oxidation coupled to substrate reduction. For endpoint assay, reactions were quenched in methanol, diluted to a final concentration of 1 µM in 50% methanol, and analyzed by LC-MS/MS as described in Section **2.6**.

Large-scale *in vitro* assay and NMR spectroscopy

A 5 mL reaction containing 0.5 mM digoxin, 5 µM Cgr2, 1.5 mM MV, 1 mM sodium dithionite, and 0.5 mM FAD was prepared in 50 mM HEPES, 100 mM NaCl, pH 7. The reaction was carried out for 1.5 hours at 20 °C, and then extracted three times with 5 mL of dichloromethane. The pooled organic fractions were concentrated using a rotary evaporator, and then frozen in liquid nitrogen and lyophilized overnight to remove residual water. The solid product (~2 mg) and digoxin (5 mg) were separately dissolved in 200 µL of deuterated methanol (CD_3OD) (Cambridge Isotope Laboratories) and transferred to Norell® Select Series™ high-throughput NMR tubes (Sigma-Aldrich). The product spectra was also collected in deuterated acetone (CD_3OCD_3). Nuclear magnetic resonance (NMR) spectra were collected

on a 600 MHz spectrophotomer (Agilent DD2 600) at 25 °C. Chemical shifts are reported in parts per million (ppm), using either the CD_3OD (3.31 ppm) or CD_3OCD_3 (2.05 ppm) resonance as an internal standard. Chemical resonances of the product generated by Cgr2 matched previously published data for the (20*R*) diastereomer (*41*). Spectra were visualized with MestReNova v. 11.0.2-18153.

3.9. References

1. Y. Y. Londer, Expression of recombinant cytochromes *c* in *E. coli*. *Methods Mol Biol* **705**, 123-150 (2011).

2. S. Nallamsetty, B. P. Austin, K. J. Penrose, D. S. Waugh, Gateway vectors for the production of combinatorially-tagged His6-MBP fusion proteins in the cytoplasm and periplasm of *Escherichia coli*. *Protein Sci* **14**, 2964-2971 (2005).

3. L. Thöny-Meyer, Cytochrome *c* maturation: a complex pathway for a simple task? *Biochemical Society Transactions* **30**, 633-638 (2002).

4. N. Nakashima, T. Tamura, A novel system for expressing recombinant proteins over a wide temperature range from 4 to 35 degrees C. *Biotechnol Bioeng* **86**, 136-148 (2004).

5. Y. Mitani, X. Meng, Y. Kamagata, T. Tamura, Characterization of LtsA from *Rhodococcus erythropolis*, an enzyme with glutamine amidotransferase activity. *J Bacteriol* **187**, 2582-25891 (2005).

6. N. Nakashima, T. Tamura, Isolation and characterization of a rolling-circle-type plasmid from *Rhodococcus erythropolis* and application of the plasmid to multiple-recombinant-protein expression. *Appl Environ Microbiol* **70**, 5557-55568 (2004).

7. T. M. Iverson, C. Luna-Chavez, L. R. Croal, G. Cecchini, D. C. Rees, Crystallographic studies of the *Escherichia coli* quinol-fumarate reductase with inhibitors bound to the quinol-binding site. *J Biol Chem* **277**, 16124-16130 (2002).

8. D. Leys *et al.*, Structure and mechanism of the flavocytochrome *c* fumarate reductase of *Shewanella putrefaciens* MR-1. *Nature Structural Biology* **6**, 1113-1117 (1999).

9. M. Kern, O. Einsle, J. Simon, Variants of the tetrahaem cytochrome c quinol dehydrogenase NrfH characterize the menaquinol-binding site, the haem c-binding motifs and the transmembrane segment. *Biochem J* **414**, 73-79 (2008).

10. A. Miseta, P. Csutora, Relationship between the occurrence of cysteine in proteins and the complexity of organisms *Molecular Biology and Evolution* **17**, 1232-1239 (2000).

11. Y. Zhang *et al.*, Diphthamide biosynthesis requires an organic radical generated by an iron-sulphur enzyme. *Nature* **465**, 891-896 (2010).

12. E. Nakamaru-Ogiso, T. Yano, T. Ohnishi, T. Yagi, Characterization of the iron-sulfur cluster coordinated by a cysteine cluster motif (CXXCXXXCX27C) in the Nqo3 subunit in the proton-translocating NADH-quinone oxidoreductase (NDH-1) of Thermus thermophilus HB-8. *J Biol Chem* **277**, 1680-1688 (2002).

13. T. T. Lee, S. Agarwalla, R. M. Stroud, Crystal structure of RumA, an iron-sulfur cluster containing E. coli ribosomal RNA 5-methyluridine methyltransferase. *Structure* **12**, 397-407 (2004).

14. M. E. Pandelia *et al.*, Characterization of a unique [FeS] cluster in the electron transfer chain of the oxygen tolerant [NiFe] hydrogenase from Aquifex aeolicus. *Proc Natl Acad Sci U S A* **108**, 6097-6102 (2011).

15. K. D. Schnackerz, D. Dobritzsch, Y. Lindqvist, P. F. Cook, Dihydropyrimidine dehydrogenase: a flavoprotein with four iron-sulfur clusters. *Biochim Biophys Acta* **1701**, 61-74 (2004).

16. H. K. Leech *et al.*, Characterization of the cobaltochelatase CbiXL: evidence for a 4Fe-4S center housed within an MXCXXC motif. *J Biol Chem* **278**, 41900-41907 (2003).

17. A. A. Gorodetsky, M. C. Buzzeo, J. K. Barton, DNA-mediated electrochemistry. *Bioconjug Chem* **19**, 2285-2296 (2008).

18. M. Lee *et al.*, Biosynthesis of isoprenoids: crystal structure of the [4Fe-4S] cluster protein IspG. *J Mol Biol* **404**, 600-610 (2010).

19. B. E. Weiner *et al.*, An iron-sulfur cluster in the C-terminal domain of the p58 subunit of human DNA primase. *J Biol Chem* **282**, 33444-33451 (2007).

20. S. Klinge, J. Hirst, J. D. Maman, T. Krude, L. Pellegrini, An iron-sulfur domain of the eukaryotic primase is essential for RNA primer synthesis. *Nat Struct Mol Biol* **14**, 875-877 (2007).

21. S. Dickert, Pierik, A.J., Buckel, W., Molecular characterization of phenyllactate dehydratase and its initiator from Clostridium sporogenes. *Molecular microbiology* **44**, 49-60 (2002).

22. R. C. Conover, Kowel, A.T., Fu, W., Park, J., Aono, S., Adams, M.W.W., Johnson, M.K., Spectroscopic characterization of the novel iron-sulfur cluster in Pyrococcus furiosus ferredoxin. *Journal of Biological Chemistry* **265**, 8533-8541 (1990).

23. D. Schneider, C. L. Schmidt, Multiple Rieske proteins in prokaryotes: where and why? *Biochim Biophys Acta* **1710**, 1-12 (2005).

24. T. Iwasaki, Okamoto, K., Nishino, T., Mizushima, J., Hori, H., Nishino, T., Sequence motif-specific assignment of two [2Fe-2S] clusters in rat xanthine oxidoreductase studied by site-directed mutagenesis. *J Biochem-Tokyo* **127**, 771-778 (2000).

25. L. Banci, Ciofi-Baffoni, S., Mikolajczyk, M., Winkelmann, J., Bill, E., Pandelia, M.E., Human anamorsin binds [2Fe-2S] clusters with unique electronic properties. *J Biol Inorg Chem* **18**, 883-893 (2013).

26. T. A. Dailey, H. A. Dailey, Identification of [2Fe-2S] Clusters in Microbial Ferrochelatases. *Journal of Bacteriology* **184**, 2460-2464 (2002).

27. Y. S. Jung *et al.*, Structure of C42D Azotobacter vinelandii FdI. A Cys-X-X-Asp-X-X-Cys motif ligates an air-stable [4Fe-4S]2+/+ cluster. *J Biol Chem* **275**, 36974-36983 (2000).

28. T. Watanabe, K. Honda, Measurement of the extinction coefficient of the methyl viologen cation radical and the efficiency of its formation by semiconductor photocatalysis. *Journal of Physical Cemistry* **86**, 2617-2619 (1982).

29. C. Ayala-Castro, A. Saini, F. W. Outten, Fe-S cluster assembly pathways in bacteria. *Microbiol Mol Biol Rev* **72**, 110-125 (2008).

30. G. Hanson, L. Berliner, *Metals in biology: Applications of high-resolution EPR to metalloenzymes*. (Springer Science, New York, 2010), vol. 29.

31. T. Yano, S. Magnitsky, V. D. Sled, T. Ohnishi, T. Yagi, Characterization of the putative 2x[4Fe-4S]-binding NQO9 subunit of the proton-translocating NADH-quinone oxidoreductase (NDH-1) of Paracoccus denitrificans. Expression, reconstitution, and EPR characterization. *J Biol Chem* **274**, 28598-28605 (1999).

32. S. Craciun, J. A. Marks, E. P. Balskus, Characterization of choline trimethylamine-lyase expands the chemistry of glycyl radical enzymes. *ACS Chem Biol* **9**, 1408-1413 (2014).

33. J. A. Imlay, Iron-sulphur clusters and the problem with oxygen. *Mol Microbiol* **59**, 1073-1082 (2006).

34. M. Laursen, J. L. Gregersen, L. Yatime, P. Nissen, N. U. Fedosova, Structures and characterization of digoxin- and bufalin-bound Na+,K+-ATPase compared with the ouabain-bound complex. *Proc Natl Acad Sci U S A* **112**, 1755-1760 (2015).

35. K. M. Weigand *et al.*, Na(+),K(+)-ATPase isoform selectivity for digitalis-like compounds is determined by two amino acids in the first extracellular loop. *Chem Res Toxicol* **27**, 2082-2092 (2014).

36. K. S. Hewitson *et al.*, The iron-sulfur center of biotin synthase: site-directed mutants. *J Biol Inorg Chem* **7**, 83-93 (2002).

37. S. E. Iismaa, Vazquez, A.E., Jensen, G.M., Stephens, P.J., Butt, J.N., Armstrong, F.A., Burgess, B.K., Site-directed mutagenesis of Azotobacter vinelandii ferredoxin I. Changes in [4Fe-4S] cluster reduction potential and reactivity. *Journal of Biological Chemistry* **266**, 21563-21571 (1991).

38. A. E. Martin, Burgess, B.K., Stout. C.D., Cash, V.L., Deans, D.R., Jensen, G.M., Stephens, P.J., Site-directed mutagenesis of Azotobacter vinelandii ferredoxin I: [Fe-S] cluster-driven protein rearrangement. *Proc Natl Acad Sci U S A* **87**, 598-602 (1990).

39. L. W. Robertson, A. Chandrasekaran, R. H. Reuning, J. Hul, B. D. Rawal, Reduction of digoxin to 20*R*-dihydrodigoxin by cultures of *Eubacterium lentum*. *Applied and Environmental Microbiology* **51**, 1300-1303 (1986).

40. R. H. Reuning, T. A. Shephard, B. E. Morrison, H. N. Bockbrader, Formation of [20*R*]-dihydrodigoxin from digoxin in humans. *Drug Metab Dispos* **13**, 51-57 (1985).

41. H. N. Bockbrader, R. H. Reuning, Spectral analysis of the configuration and solution conformation of dihydrodigoxigenin epimers. *Journal of Pharmaceutical Sciences* **72**, 271-274 (1983).

42. G. A. Reid, C. S. Miles, R. K. Moysey, K. L. Pankhurst, S. K. Chapman, Catalysis in fumarate reductase. *Biochim Biophys Acta* **1459**, 310-315 (2000).

43. C. J. Morris, Black, A.C., Pealing, S.L., Manson, F.D.C., Chapman, S.K., Reid, G.A., Gibson, D.M., Ward, F.B., Purification and properties of a novel cytochrome: flavocytochrome c from Shewanella putrefaciens. *Biochem J* **302**, 587-593 (1994).

44. G. L. Kemp *et al.*, Kinetic and thermodynamic resolution of the interactions between sulfite and the pentahaem cytochrome NrfA from Escherichia coli. *Biochem J* **431**, 73-80 (2010).

45. S. Schottler, W. Wende, V. Pingoud, A. Pingoud, Identification of Asp218 and Asp326 as the principal Mg2+ binding ligands of the homing endonuclease PI-SceI. *Biochemistry* **39**, 15895-15900 (2000).

46. A. M. Orville *et al.*, Thiolate ligation of the active site Fe2+ of isopenicillin N synthase derives from substrate rather than endogenous cysteine: spectroscopic studies of site-specific Cys to Ser mutated enzymes. *Biochemistry* **31**, 4602-4612 (1992).

Chapter 4: Investigations of the substrate specificity and biological role of Cgr2

4.1. Probing the physiological implications of digoxin metabolism

After demonstrating that Cgr2 was uniquely responsible for digoxin inactivation, we wanted to understand whether this metabolism conferred a physiological benefit to *cgr*$^+$ *E. lenta* strains. To that end, we examined the interaction between Cgr2 and dietary components (L-arginine) that had previously been shown to influence digoxin metabolism by *E. lenta* in culture and in gnotobiotic mouse models. In addition, we cultured *E. lenta* under a range of conditions to determine whether digoxin metabolism was associated with a growth advantage. Finally, we characterized the kinetic parameters and substrate scope of Cgr2 to understand whether digoxin is likely the endogenous substrate of this enzyme, or whether Cgr2 may perform a broader range of reductive chemistry within the human gut.

4.1.1 Dietary protein influences *E. lenta* growth and digoxin metabolism

Following the discovery that a subset of *E. lenta* strains was responsible for digoxin inactivation (*1*), scientists aimed to identify additional factors that contribute to variable metabolism of this drug in the broader population. Mathan, Alam, and co-workers found that the incidence of digoxin reduction differed for individuals living under different geographic, socioeconomic and environmental conditions (*2, 3*). For example, dihydrodigoxin was detected at higher levels in healthy volunteers in New York (35%) as compared to individuals in South India (14%) or Bangladesh (14%) (*2, 3*). In addition, South Indians living in urban environments were significantly more likely to metabolize digoxin than those living in rural settings (Figure 48A) (*2, 3*). Higher levels of digoxin reduction were further correlated with socioeconomic factors (e.g. higher income and education), as well as higher consumption of animal protein (Figure 48B) (*2, 3*).

Although diet is undoubtedly an important modulator of gut microbial composition (*4*), digoxin-reducing *E. lenta* strains were also isolated from the stools of individuals that did not produce dihydrodigoxin *in vivo* (*2, 3*). These results indicated that variable diets did not necessarily preclude colonization by *cgr*$^+$ *E. lenta* strains and suggested that additional metabolic interactions influenced digoxin metabolism in humans. Interestingly, higher consumption of animal protein was correlated with

higher digoxin metabolism in humans (*2, 3*), whereas dietary protein led to lower metabolism in GF mice that were mono-colonized with a digoxin-metabolizing *E. lenta* strain (*1, 5*). Furthermore, the amino acid L-arginine led to lower digoxin metabolism in pure culture of *E. lenta* (*1, 5*). An additional important consideration is that *E. lenta* species metabolize a wide range of plant-derived bioactive compounds (*6-8*), and further work is required to understand how distinct metabolic pathways are regulated when multiple substrates (e.g. mixture of natural products from plants) and nutrients (e.g. amino acids, proteins) are present in the diet. Although many questions remain, it is clear that host diet significantly influences the extent of digoxin metabolism by *E. lenta*.

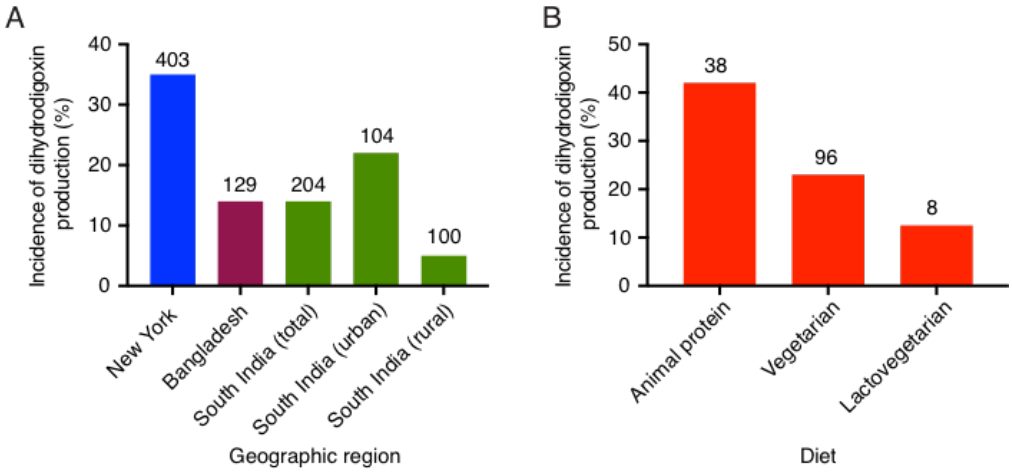


Figure 48: Geographic and dietary factors correlate with digoxin metabolism in humans. Rate of dihydrodigoxin incidence across different (A) geographic regions and (B) diets. The total number of individuals in each category is shown above the corresponding bar.
**Summary of data from Mathan, Alam, and co-workers (2, 3).*

The influence of dietary protein on digoxin metabolism is perhaps not surprising, given that various amino acids substantially influence *E. lenta* growth and physiology. In particular, this non-saccharolytic bacteria can use L-arginine as its sole energy source through the arginine dihydrolase pathway (*1, 9*). *E. lenta* DSM2243 harbors an operon (Elen_1946-1949) that encodes for all necessary

enzymes in this pathway, which generates ATP through the metabolism of arginine (Figure 49) (*10*). This operon also encodes an antiporter that imports L-arginine into the cell. Arginine deiminase is the first enzyme in this pathway that deaminates L-arginine to produce L-citrulline. Ornithine transcarbamylase then catalyzes the conversion of inorganic phosphate and L-citrulline into carbamyl phosphate and L-ornithine, the latter of which is exported from the cell via the antiporter. Finally, carbamate kinase performs substrate-level phosphorylation by transferring the phosphoryl group from carbamyl phosphate onto ADP, generating ATP as well as ammonia and carbon dioxide as byproducts. Several labs have demonstrated that adding L-arginine and L-citrulline to growth media augments *E. lenta* growth, while adding ornithine does not (*5, 9*).

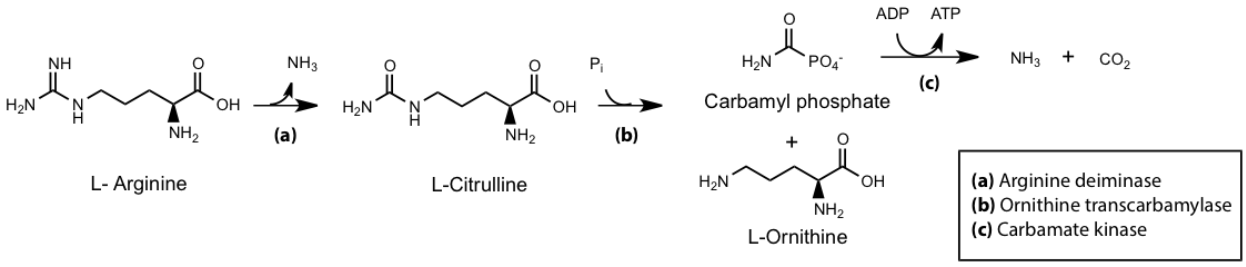


Figure 49: *E. lenta* generates ATP through the arginine dihydrolase pathway.

In addition, L-arginine was shown to significantly influence the digoxin-inactivating activity of *E. lenta*. Haiser *et al.* demonstrated that L-arginine substantially reduced *cgr*2 transcription levels and digoxin metabolism in culture. Furthermore, supplementation of dietary protein inhibited digoxin metabolism in GF mice that were mono-associated with *E. lenta* DSM2243 (*5*).

Using our *in vitro* system, we aimed to test whether L-arginine affected drug metabolism through direct inhibition of the digoxin reductase Cgr2. However, we found that the *in vitro* activity of Cgr2 was unaffected by the presence of up to 1000-fold excess L-arginine (5 mM) (Figure 50). As L-arginine does not inhibit Cgr2, it is likely influencing drug metabolism indirectly by affecting other global cellular

processes (e.g. energy utilization, transcriptional regulation) in *E. lenta*. Regardless of the mechanism by which L-arginine and other amino acids influence digoxin metabolism, this example highlights the complex nature of gut microbiota-xenobiotic interactions, which can vary not only due to the presence of specific microbial metabolic pathways, but also in response to dietary and environmental factors.

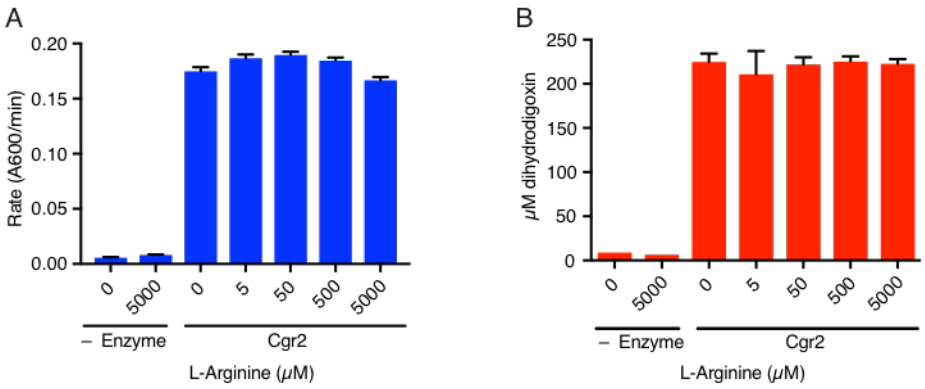


Figure 50: L-arginine does not directly inhibit Cgr2.
(A) Initial *in vitro* rates of digoxin-dependent MV oxidation by Cgr2 in the presence of varying concentration of L-arginine. (B) LC-MS/MS analysis of *in vitro* reactions quenched at 20 minutes. Data represents mean ± SEM (n = 3).

4.1.2 Digoxin reduction does not influence *E. lenta* growth

The high sequence conservation and levels of *cgr* operon transcription in response to digoxin incubation suggest that digoxin metabolism may provide a physiological benefit to *E. lenta*. In related anaerobic reeducate systems, substrate (e.g. fumarate, urocanate) reduction supports microbial growth (*11, 12*). We thus postulated that digoxin could serve as an alternative terminal electron acceptor, which would enable *E. lenta* to generate ATP in the absence of its primary energy source, L-arginine (*9*). To determine whether digoxin reduction conferred a growth advantage to cgr^+ strains, *E.* lenta DSM2243 was cultured in rich media under conditions that led to transcription of the *cgr* operon (*5*). However, we monitored the growth of *E. lenta* and found that digoxin supplementation did not alter bacterial growth

under these conditions (Figure 51) (*5*). In addition, *E. lenta* growth was unaffected by addition of either the common terminal electron acceptor fumarate or the related cardenolides digitoxin, digoxigenin, ouabain, and ouabagenin (Figure 51). Like digoxin, these cardenolides stimulate expression of the *cgr* operon (*5*) and are also reduced by *E. lenta* cells and by Cgr2 *in vitro* (Section **4.3**).

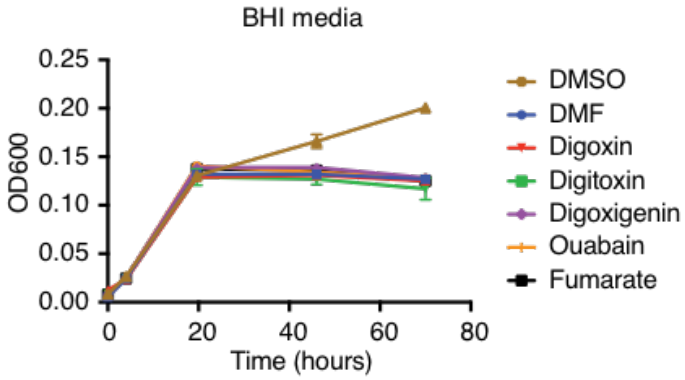


Figure 51: Digoxin and related cardenolides do not affect *E. lenta* growth in rich media. *E. lenta* DSM2243 was grown in 10 mL of BHI media supplemented with 10 µM of each substrate or an equivalent volume (0.1% v/v) of solvent (DMSO and DMF). Cultures were incubated at 37 °C. Data represents mean ± SEM (n = 3).

We next reasoned that digoxin metabolism might only be advantageous under nutrient-limiting conditions and tested *E. lenta* growth in a basal media lacking terminal electron acceptors (*13*). However, no growth advantage was observed in various media containing cardenolides as the sole electron acceptors and a variety of electron donors (e.g. H_2 or sodium acetate) (Figure 52). Fortuitously, we found that *E. lenta* growth was enhanced in the presence of dimethylsulfoxide (DMSO), which we had originally used to dissolve the cardenolide substrates. Homologs of all necessary anaerobic DMSO reductase and ATP synthase subunits were present in the *E. lenta* DSM2243 genome (KEGG database). Therefore, anaerobic respiration of DMSO is a viable pathway for generating ATP in *E. lenta*, despite previous conclusions that this bacterium derives all of its energy from substrate level phosphorylation using the arginine dihydrolase pathway (*9*). In conclusion, these results suggest that cardiac glycosides are

not likely to serve as terminal electron acceptors, as their reduction does not affect *E. lenta* growth in either rich or minimal growth medias. This observation suggested that there may be another, unidentified benefit associated with metabolism, or alternatively that digoxin is not the endogenous substrate of Cgr2.

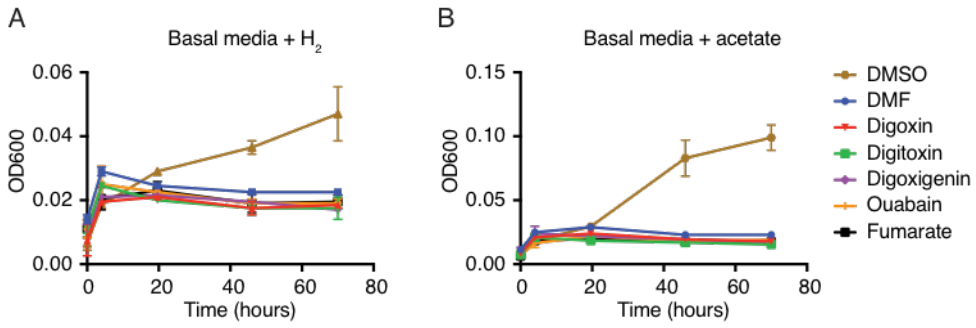


Figure 52: Cardenolides do not affect *E. lenta* growth in basal growth media.
E. lenta DSM2243 growth in basal media lacking terminal electron acceptors and supplemented with (A) 5% H_2 or (B) 10 mM sodium acetate as electron donors. Cultures were grown in 10 mL of media supplemented with 10 µM of each substrate or an equivalent volume (0.1% v/v) of solvent (DMSO and DMF). Cultures were incubated at 37 °C. Data represents mean ± SEM (n = 3).

4.2. Kinetics of Cgr2 towards digoxin

We next performed experiments to characterize the kinetic parameters and substrate specificity of Cgr2 in order to determine if digoxin is likely to be its endogenous substrate. In preliminary experiments, reconstituted Cgr2 (Y333) was incubated with varying concentrations of digoxin, and *in vitro* time-course reactions were quenched with methanol and analyzed by LC-MS/MS (Figure 53). Due to the low solubility of digoxin in aqueous solution, saturating V_{max} conditions could not be obtained under these *in vitro* conditions. Nevertheless, Cgr2 displayed Michaelis–Menten like behavior within the range of accessible concentrations. However, the measured K_m = 325 ± 55 µM and a catalytic efficiency of 5.0 ± 0.1 x10^2 M^{-1} s^{-1}, indicated the activity of Cgr2 towards digoxin was very low.

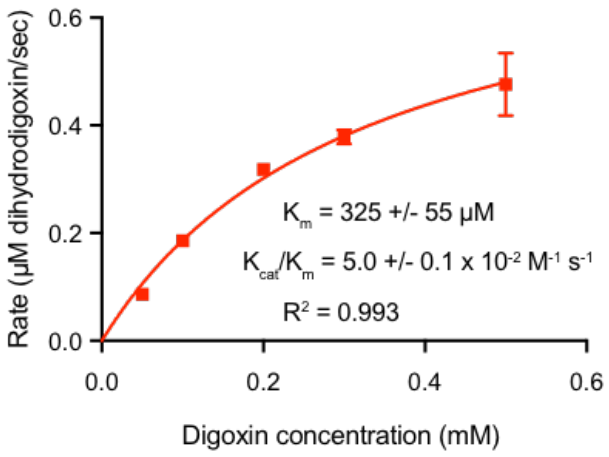


Figure 53: Initial kinetic characterization of Cgr2 towards digoxin.
In vitro reactions (200 µL) contained 5 µM Cgr2, 50 µM FAD, 0.375 mM MV, 0.25 mM sodium dithionite and 0.05-0.5 mM digoxin. Reactions were quenched in methanol and analyzed by LC-MS/MS. Data represents mean ± SEM (n = 3).

While these values could suggest that digoxin is not the endogenous substrate of Cgr2, our *in vitro* system was lacking several important components that could contribute to artificially low activity. We first addressed issues with cofactor binding and incorporation by optimizing [Fe-S] reconstitution conditions (Section **3.3**) and determining the optimal FAD concentrations (Section **3.5**) for maximizing the activity of Cgr2 *in vitro*. In addition, as we could not purify Cgr1, we relied on artificial electron donors to provide reducing equivalents for digoxin reduction. However, it is likely that the rate of electron transfer from an artificial electron donor is lower than electron transfer from the presumed biological reductase Cgr1. We originally used methyl viologen (MV) as an electron donor because it is a strong reductant (–440 mV) and enables facile monitoring of kinetic activity due its strong absorbance at 600 nm in the reduced state (*14*). To identify a more optimal electron donor we screened a variety of electron mediators (Figure 54A) with a range of reduction potentials (–660 mV to –255 mV) (*15, 16*). This redox potential range should be able to mediate a reduction cascade involving a typical $[4Fe\text{-}4S]^{2+}$ cluster (–715 to –280 mV), FAD cofactor (–220 mV), and ultimately digoxin (250 - 350 mV) (*17, 18*). Using colorimetric and LC-MS/MS assay, we observed that MV and diquat were both viable electron donors, although MV provided the highest rates of electron transfer and product formation (Figure 54B-D). Interestingly, product formation was not observed for BV, which has a similar reduction potential to diquat, suggesting that other features are important for electron transfer to Cgr2. For example, BV may be too large to bind near or engage with the [4Fe-4S] clusters of Cgr2. This could provide further support to the hypothesis that artificial electron donors are not as efficient as a protein (Cgr1) that likely evolved to transfer electrons within a protein complex.

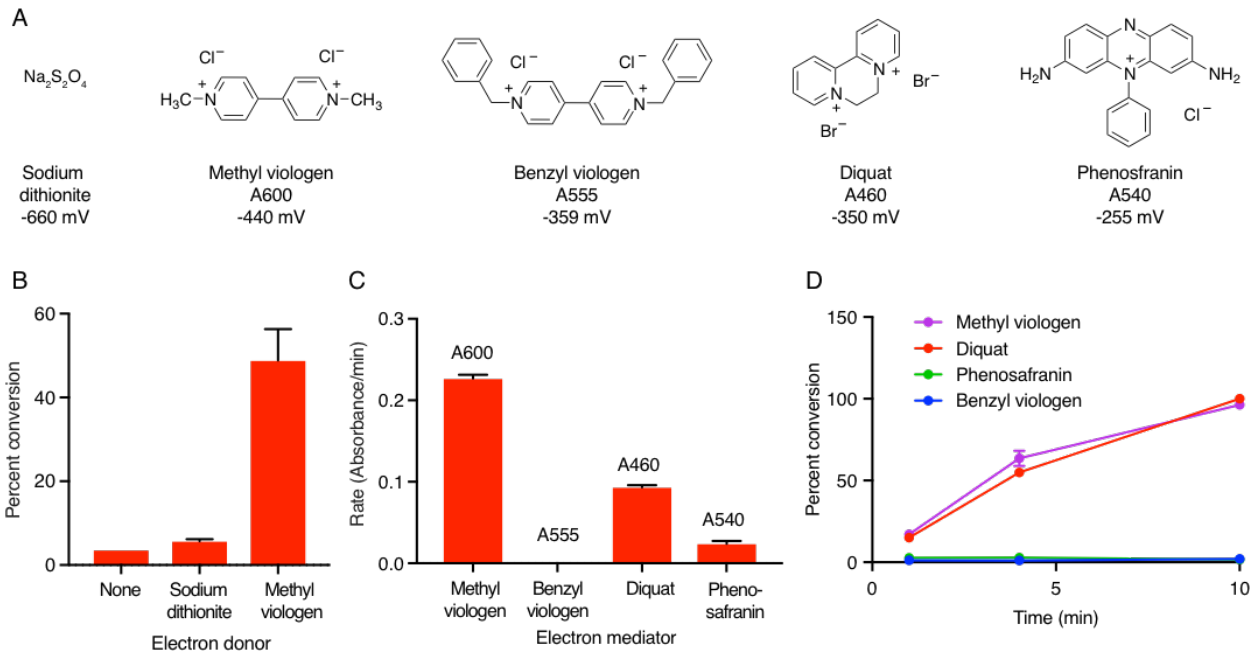


Figure 54: Methyl viologen is the optimal electron donor for Cgr2 *in vitro*.
(A) Chemicals used to provide reducing equivalents to Cgr2 *in vitro*. Maximum absorption peaks and reduction potentials shown below each compound. (B) LC-MS/MS analysis of *in vitro* reactions (quenched at 5 minutes) using sodium dithionite or MV as electron donors. (C) Initial *in vitro* rates showing decrease in maximum absorbance for each electron mediator over time. (D) Percent conversion of digoxin to dihydrodigoxin over time. *In vitro* reactions were quenched at various time points and analyzed by LC-MS/MS. Data represents mean ± SEM (n = 3).

Finally, we wanted to test the impact of detergents on Cgr2 activity, as we predict that this enzyme is localized in or near the *E. lenta* cell membrane (Figure 47). In addition, we had previously observed that the [Fe-S] clusters of Cgr2 were labile during additional purification steps and that Cgr2 displayed low affinity for FAD (Section **3.5**). These findings led us to hypothesize that the cofactor binding sites may be in a hydrophobic or exposed region that would normally be capped by Cgr1 or the cell membrane, and we wondered whether detergents could serve this role *in vitro*. However, when the detergent DDM was added to *in vitro* assays, Cgr2 activity was diminished (Figure 55). These data suggest that digoxin may have a high affinity for membranes, and that the substrate may be inaccessible to Cgr2 in this context when the membrane associated Cgr1 is not present. Alternatively, detergents may

disrupt Cgr2 folding, and further experiments will be required to tease apart the interaction between Cgr2 and detergents and membrane components. Ultimately, while we were able to optimize Cgr2 activity to some degree, these results further suggest that the activity of Cgr2 may be impaired *in vitro* in the absence of all of the Cgr machinery.

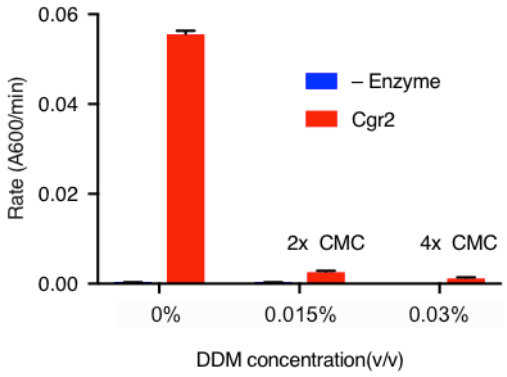


Figure 55: Cgr2 activity towards digoxin is reduced in the presence of detergent.
Initial *in vitro* rates using the detergent DDM above the critical micelle concentration (CMC). Reactions contained 5 µM Cgr2, 50 µM FAD, 0.375 mM MV, 0.25 mM sodium dithionite and were initiated upon addition of 0.5 mM digoxin. Reactions were quenched in methanol and analyzed by LC-MS/MS. Data represents mean ± SEM (n = 3).

After optimizing our *in vitro* assay, we determined the kinetic parameters for purified, reconstituted Cgr2 Y333 as well as the second naturally occurring variant N333. We had previously noticed that *E. lenta* strains encoding the N333 variant generally displayed a decreased ability to metabolize digoxin as compared to Y333-encoding strains (Figure 29C). This difference in activity was more pronounced in *in vitro* assays using the two Cgr2 variants (Figure 56A). Using optimal assay conditions, wild-type Cgr2 (Y333) exhibited a K_m of 94.6 ± 7.1 µM and a catalytic efficiency of 2.4 ± 0.8 x 10^3 M^{-1} s^{-1}. However, saturating V_{max} conditions could not be reached for the N333 variant within the

range of concentrations where digoxin is soluble (≤ 0.5 mM). Although it is possible that the Y333 residue is involved in substrate binding or catalysis (*19*), we were unable to resolve whether Cgr2 N333 was less active because of lower substrate binding or lower turnover relative to Cgr2 Y333. Despite its reduced activity, Cgr2 N333 converted digoxin to dihydrodigoxin at a comparable efficiency to Y333 at longer times (Figure 56B). It is possible that the marked differences in activity of the Cgr2 variants was not as readily apparent in *E. lenta* culture (Figure 29C) due to the considerably longer incubation times (2 days in culture vs. 4.5 hours *in vitro*) and lower digoxin concentrations (10 µM in culture vs. 500 µM *in vitro*) used in whole cell experiments. Although it is tempting to consider that these *in vitro* differences in activity would translate to differential clinical outcomes, further experiments are required to determine how Cgr2 variants impact digoxin pharmacokinetics in humans (Section **5.3**).

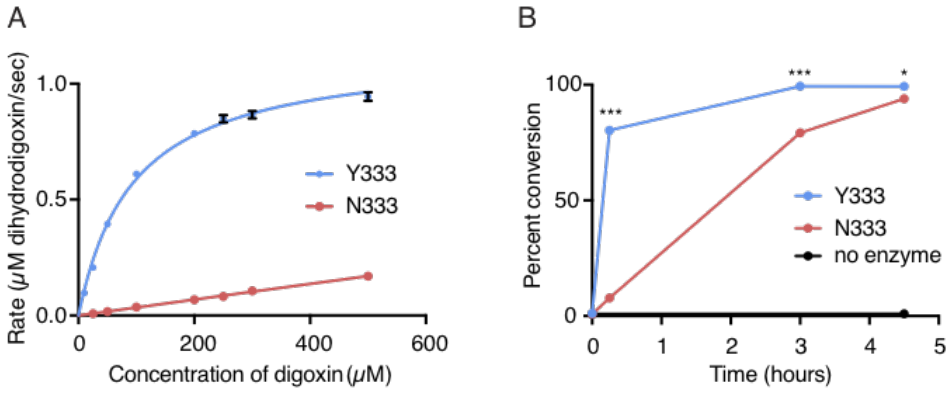


Figure 56: Optimized kinetics of Cgr2 variants for digoxin reduction.
(A) Michaelis–Menten kinetics of Cgr2 for digoxin reduction revealed that the Y333 variant is significantly more active than the N333 variant. Data represents mean ± SEM (n = 3). (B) *In vitro* time course of the conversion of digoxin to dihydrodigoxin by Cgr2 Y333 and N333 variants. Reaction aliquots were quenched in methanol and analyzed by LC-MS/MS. Values represent mean ± SEM (n = 3). Asterisks indicate statistical significance at each time-point of Y333 vs. N333 conversion, by Student's t test (*$p < 0.05$, ***$p < 0.001$).

Despite considerable improvement of its *in vitro* activity compared to our initial attempts (Figure 53), under optimized conditions Cgr2 Y333 was still less efficient for digoxin reduction as compared to the activity of other FAD-dependent reductases towards their native substrates (*11, 19-21*) (Table 7). This decreased activity could arise from impaired cofactor binding or reconstitution, or inefficient electron transfer *in vitro* in the absence of Cgr1. Alternatively, these results could indicate that digoxin, which is administered at very low concentrations (0.125-0.25 mg daily; ~1 nM in serum) due to its toxicity (*22, 23*), is not the endogenous substrate of Cgr2.

Table 7: Comparison of Cgr2 kinetic parameters with those of related reductases.

Enzyme	Substrate	K_{cat} (s^{-1})	K_m (μM)	K_{cat}/K_m (M^{-1} s^{-1})
E. lenta Cgr2 wild-type	Digoxin	0.23 ± 0.01	94.6 ± 7.1	2.4 ± 0.8 x 10^3
R. erythropolis SQ1 3-ketosteroid-Δ1-dehydrogenase	4-Androstene-3,17-dione	68.1 ± 1.4	37.0 ± 0.8	1.8 ± 0.1 x 10^6
S. putrefaciens flavocytochrome *c*	Fumarate	250 ± 50	21 ± 10	1.2 ± 0.5 x 10^7
S. oneidensis MR-1 UrdA	Urocanate	360	<<10	n/a
E. coli NrfA	Nitrite	22 ± 7	769 ± 20	3.5 x 10^7
E. coli NrfA	Sulfite	0.03 ± 0.01	70 ± 15	4.3 x 10^3

4.3. Identifying potential substrates of Cgr2

Together, the poor kinetics of Cgr2 towards digoxin (Figure 56), the lack of an obvious physiological benefit for *cgr*$^+$ *E. lenta* strains (Figure 51-Figure 52), and the low concentrations of digoxin administered to patients (*22, 23*) led us to question whether digoxin was the endogenous substrate for this enzyme. To identify potential endogenous substrate for Cgr2, we selected a panel of 28 small molecules based on their chemical similarity to digoxin, presence at higher concentrations in the GI tract, or physiological relevance, as discussed in the following sections. We then tested the activity of Cgr2 towards these alternative substrates in the hopes of understanding its physiological role in the human gut.

4.3.1 Cardiac glycosides

Cardiac glycosides are highly potent natural products produced by diverse plants (e.g. European foxglove plant, East African Ouabaio wood, *Hycinthaceae* bulbs, Indian *Leguminoseae* seeds) and some animals (e.g. toads, snakes, sea onions) (*24, 25*). This class of molecules has been used for centuries to treat a variety of ailments including cardiac dysfunction (*26*). Structural features of these natural products include a steroid core, a hydroxyl group at the C–3 position that is modified with up to four sugars (typically L-rhamnose, D-glucose, D-digitoxose, or D-digitalose), and a variable structural motif at C–17 (*27, 28*). The functional group at C–17 further stratifies cardiac glycosides into cardenolides, which contain an α,β-unsaturated butyrolactone (Figure 57A), or bufadienolides, which contain a 2H-pyran-2-one (Figure 58A) (*27, 28*). Both subclasses of cardiac glycosides inhibit Na^+/K^+ ATPases, although cardenolides bind in complex with a single Mg^{2+} ion (Figure 57B), whereas bufadienolides bind in complex with two K^+ ions (Figure 58B) (*29, 30*). Na^+/K^+ ATPase inhibition leads to increased intracellular concentrations of Ca^{2+}, which influences a variety of downstream pathways including cardiac muscle contraction and apoptosis (Section **1.4.2**) (*27*). Although the high toxicity of cardiac glycosides has led to a marked decline in their use for cardiac indications (*31*), the latter finding has prompted considerable interest in their use as anti-cancer agents (*27, 32, 33*). Cardenolides are more commonly

used in Western medicine as pure substances, whereas bufadienolides are the active constituents in Asian and Chinese traditional medicines (*24*).

Cardiac glycosides are often produced as a complex mixture of compounds (*26*), with different analogs possessing a range of selectivities, potencies, and pharmacokinetic properties (*27*). Cardenolides are exclusively biosynthesized by plants and include the pharmaceutical agents digoxin and digitoxin, which are produced by *Digitalis* plants, and ouabain, which is produced by the Ouabaio wood and was used as an arrowhead poison (Figure 57A) (*24*). Digoxin is the predominantly prescribed cardenolide worldwide, while digitoxin is more frequently used in Scandinavia (*34*). Although both drugs share the same target (Figure 57B), their pharmacokinetic properties differ considerably. Digoxin has lower bioavailability and plasma protein binding than digitoxin (*34*). In addition, while digoxin is not metabolized by the host, digitoxin undergoes extensive metabolism by liver enzymes (up to 70% of dose), (*35*) including β-oxidation at C12, and a cytochrome P450-dependent oxidation of the 15' hydroxyl in the terminal digitoxose group, which leads to successive cleavage of the digitoxose sugars (*36-38*). Finally, digitoxin is also reduced to dihydrodigitoxin in humans, suggesting that it may also be a substrate of Cgr2 (*39*).

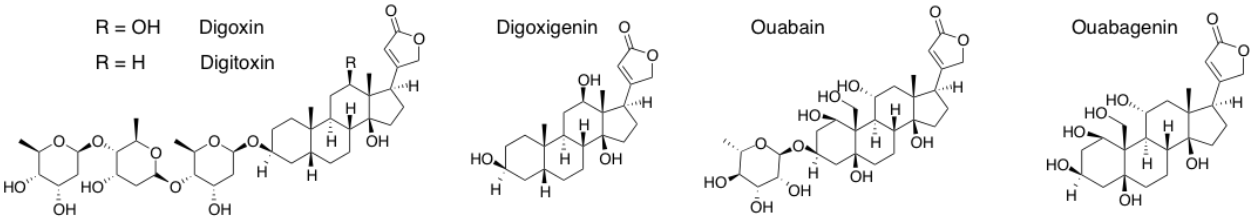


Figure 57: Cardenolide natural products.
Cardenolides contain a hydroxylated steroid core, an α,β-unsaturated butyrolactone moiety at C–17, and a range of sugar moieties at C–3. Cardenolides bind and inhibit Na^+/K^+ ATPases in cardiac cells, leading to downstream muscular contraction, and higher concentrations lead to toxicity and in extreme cases, death.

Bufadienolides are biosynthesized by plants and animals, including toads (*Bufo*), and are present in various Eastern medicines. Bufalin is the main bioactive substance in Ch'an Su, a traditional Chinese

medicine prepared from toad venom with anti-tumor, anesthetic, and caridotonic properties (*24, 27*). Hellebrin and proscillaridin are additional bufadienolides with anti-cancer and immunoregulatory properties (*27, 40*).

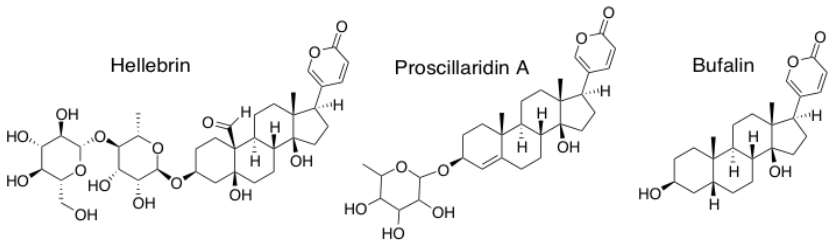


Figure 58:Bufadienolide natural products.
Bufadienolides contain a hydroxylated steroid core, a 2H-pyran-2-one moiety at C–17, and a range of sugars at C–3. Bufadienolides bind and inhibit Na^+/K^+ ATPases. Bufadienolides are active components of traditional Chinese medicines (*41*).

4.3.2 Furanones

Furanones are naturally occurring, redox-active molecules that are implicated in a broad range of inter-organismal signaling pathways and biological activities (*42*). These compounds are typically small, fragrant, and fairly unstable (prone to oxidation), which likely enables their use as temporal signaling molecules (*42*). Furanones are biosynthesized by various organisms (bacteria, yeast, plants) and also form spontaneously upon heating of amino acids and pentose sugars in foods in a process known as the Maillard reaction (*42, 43*). They are also present in a variety of fermented foods, roasted meats, seeds, and wine, and are commonly used as food additives due to their strong aromatic properties (*42, 44*). While there is considerable inter-individual variability in their consumption, dietary furanones can be ingested in multigram daily quantities, which is considerably higher than digoxin doses (< 0.25 mg daily) (*42*). We selected various 2(5H)- (Figure 59A) and 3(2H)- furanones for testing (Figure 59B), including the anti-oxidant vitamin C, the anti-carcinogen EMHF (5-ethyl-4-hydroxy-2-methyl-3(*2H*)-furanone), a broad spectrum anti-bacterial and anti-fungal agent DMHF (4-Hydroxy-2,5-dimethyl-3(*2H*)-furanone), the dietary flavorants emoxyfuranone (5-ethyl-3-hydroxy-4-methyl-2(5H)-furanone), sotolon (4,5-dimethyl-

3-hydroxy-2,5-dihydrofuran-2-one), DMMF (2,5-dimethyl-4-methoxy-3(*2H*)-furanone), and MHF (4-hydroxy-5-methyl-3-furanone) (*42, 45*). We also tested 2(5H)-furanone, the functional group in digoxin that is reduced by Cgr2, to probe the minimal requirements for substrate binding and metabolism by this enzyme. This compound is also encountered in the diet as a component of liquid smoke food flavorants (*44*).

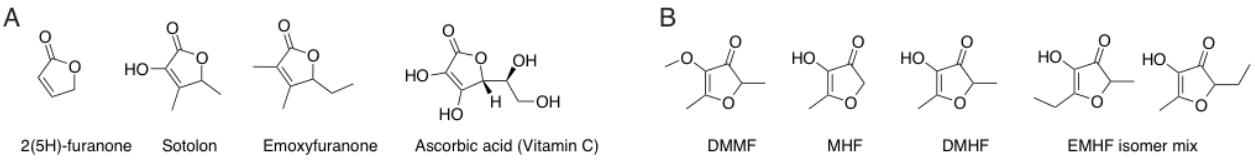


Figure 59: Furanones encountered in the diet.
(A) 2(*5H*)-furanone and (B) 3(*2H*)-furanone substrates.

4.3.3 α,β-unsaturated carboxylic acids

We also selected and tested several α,β-unsaturated carboxylic acids, including fumarate and urocanate, which are the endogenous substrates of the most similar homologs of Cgr2 (*11, 46*), as well as the antibiotic fusidic acid (*47*). Fumarate is one of the most common terminal electron acceptors used in bacterial anaerobic respiration (*48*). This substrate is likely to be encountered by bacteria in the GI tract as fumarate is used as a natural food preservative (*49*) and is ingested in high quantities (1-300 mg daily) in the form of iron supplements (ferrous fumarate) and psoriasis treatments (*50, 51*). Urocanate is a UV-protectant that arises from L-histidine metabolism in humans (*52*). Although urocanate is predominantly found in the epidermis, it has also been detected in human fecal samples, suggesting that gut microbes may encounter this molecule (http://www.hmdb.ca/; accession number HMDB0000301). Finally, we tested the activity of Cgr2 towards the Gram-positive antibiotic fusidic acid. While this molecule may not be present in the diet or the GI tract outside of the context of acute antibiotic treatment, it shares several features with digoxin, including a hydroxylated steroid core and an α,β-unsaturated carbonyl-containing functional group at C–17 that is essential for antibiotic activity (*53*).

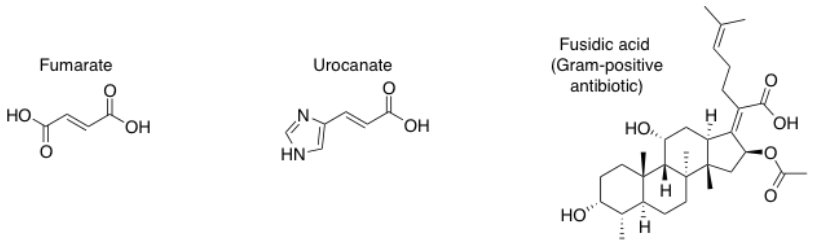


Figure 60: α,β-unsaturated carboxylic acids compounds.

4.3.4 Ketosteroids

In addition to digoxin, *E. lenta* species metabolize a variety of bile acid and steroid substrates in the human body using reductive transformations (*54, 55*). Bioinformatic analyses also showed that Cgr2 resembles ketosteroid dehydrogenases (Section **2.1**). We thus compiled a panel of additional human and microbial ketosteroids (Figure 61). This includes several human sex and stress hormones that have receptors in the GI tract or are known to impact GI function and motility, and thus may be present in the gut lumen (*56-58*). In addition, we tested two synthetic, orally administered contraceptive drugs, 19-norethindrone and levonorgestrel, that are reduced at the α,β-unsaturated ketone within the human body (*59*). Although it is not clear where this metabolism occurs, co-administration of antibiotics with 19-norethrindrone led to higher plasma levels of the drug in women, which could implicate gut bacteria in contraceptive drug inactivation (*60*). Finally, we tested a putative gut microbial metabolite of cholesterol, which is proposed to undergo reduction to the non-absorbable metabolite coprostanol in the GI tract (*61*).

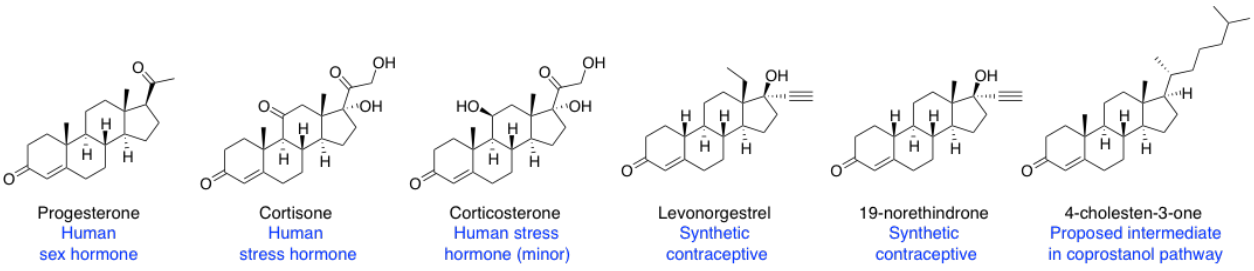


Figure 61: Ketosteroids encountered in the human gut.
Progesterone, cortisone, and corticosterone are endogenously produced hormones in humans. Levonorgestrel and 19-norethindrone are synthetic, oral contraceptives which are reduced in the human body (*59*). 4-cholesten-3-one is a putative intermediate in the gut bacterial pathway that metabolizes cholesterol into coprostanol.

4.3.5 Prostaglandins

Prostaglandins (PGs) are an important class of host lipids with hormone-like effects throughout the human body (*62, 63*). Unlike steroidal hormones, PGs are synthesized throughout many host tissues and are believed to act locally (*62*). All PGs derive from arachidonic acid and contain a 20-carbon skeleton, which includes a five-membered ring and two lipophilic tails. Different PG classes (PGA-K) mediate diverse functions in humans, including modulation of vascular constriction and blood pressure (PGA1) (*64*), alteration of GI tract motility and function (*62*), and regulation of inflammatory processes (PGA1, PGJ2) (*63, 65*). As *E. lenta* is often associated with inflammation (*66, 67*), we postulated that metabolism of anti-inflammatory PGs (PGA1, PGJ2) could provide a physiological advantage to this bacterium by promoting an inflammatory host environment. Additionally as PGs and their metabolites are found in the feces (*62*), it is feasible that *E. lenta* could interact with these compounds in the GI tract.

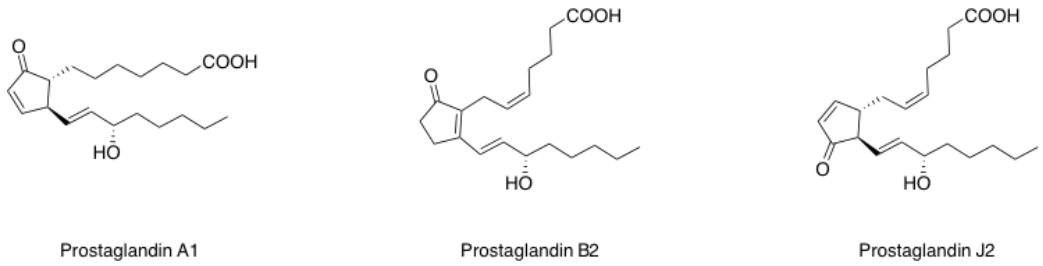


Figure 62: Host-derived prostaglandins involved in inflammatory and GI processes.

4.3.6 *In vitro* activity of Cgr2 towards the substrate panel

Cgr2 activity was tested *in vitro* against the full panel of cardiac glycosides, naturally occurring furanones, α,β-unsaturated carboxylic acids, ketosteroids, and prostaglandins, all of which may be relevant in the context of the human gut. Cgr2 displayed robust activity towards all cardenolides tested, and the cardenolide aglycones digoxigenin and ouabagenin were metabolized at a significantly faster rate

than their glycosylated forms digoxin and ouabain, respectively (** $p < 0.01$, Student's t test) (Figure 63). Dihydrocardenolide formation was confirmed for all metabolized substrates by MS analysis (Figure 64).

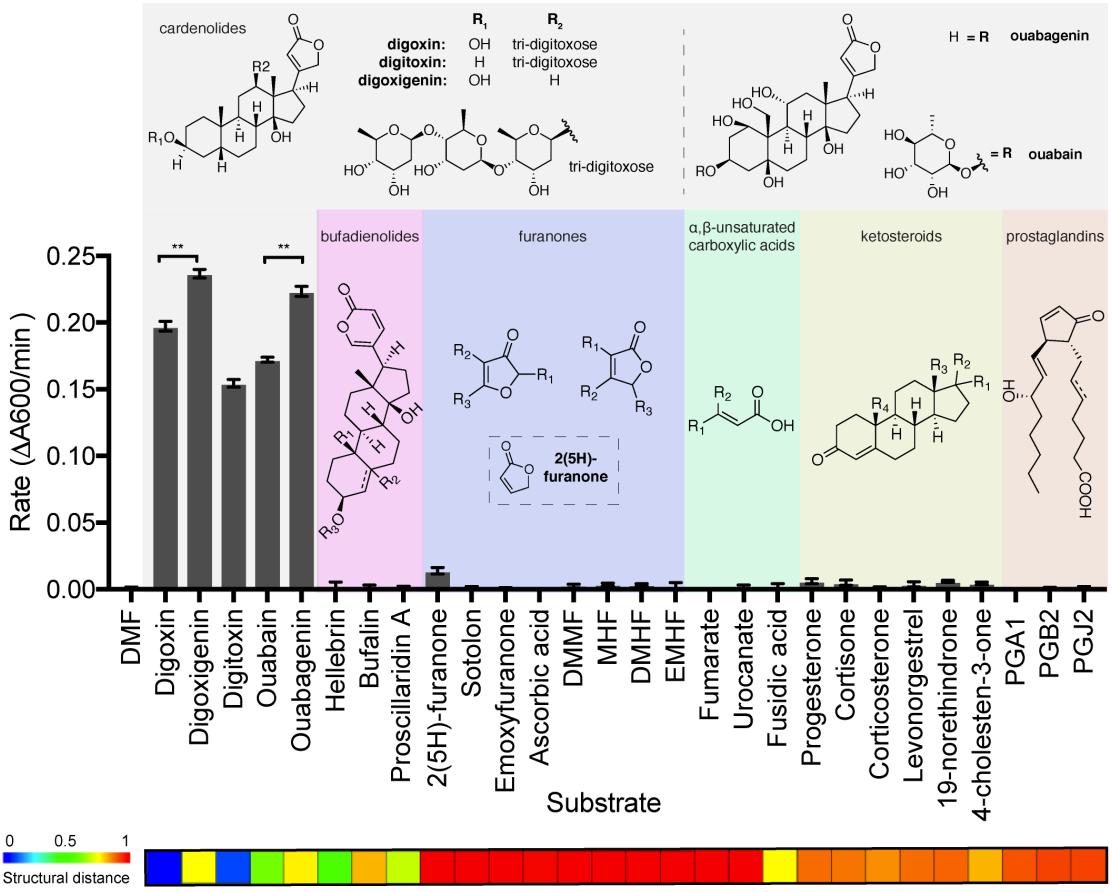


Figure 63: The substrate scope of Cgr2 is restricted to cardenolides.
Rate of methyl viologen oxidation coupled to substrate reduction by Cgr2. Colors denote different substrate classes. With the exception of the cardenolides, a representative substrate structure is shown. Values represent mean ± SEM (n = 3). ** $p < 0.01$, Student's t test. The heatmap generated in ChemMine (*68*) represents the structural similarity of each compound relative to digoxin. Structural distance matrix is calculated as (1- Tanimoto coefficient), where lower values represent more structurally similar compounds.

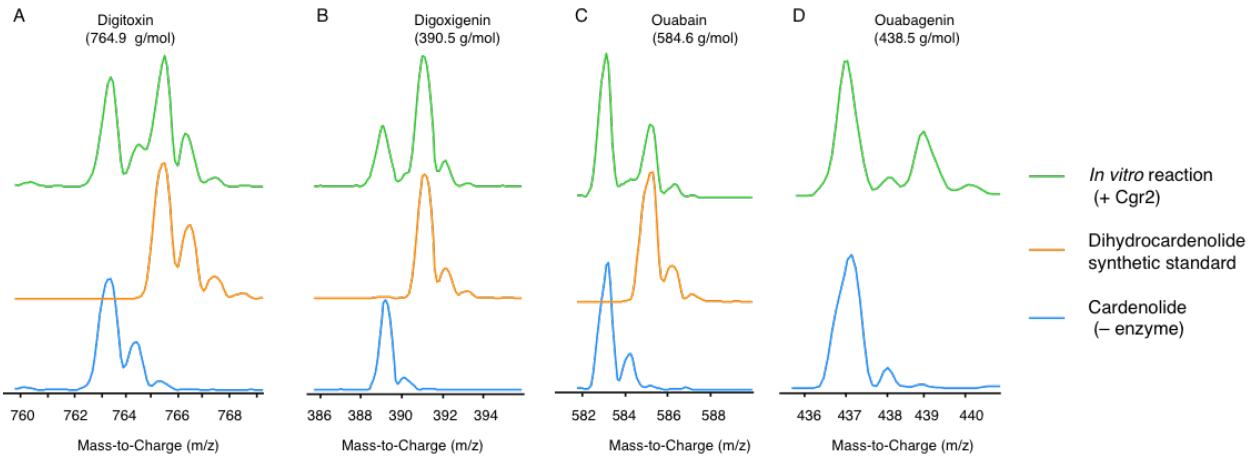


Figure 64: Confirmation of dihydrocardenolide product formation from *in vitro* reactions with Cgr2.
In vitro reactions were quenched in 10 equivalents of methanol, further diluted in 50% methanol to a 1 μM final concentration and analyzed by MS_1 scan in negative ion mode. Traces of cardenolide and dihydrocardenolide standards and *in vitro* reactions are shown for (A) digitoxin, (B) digoxigenin, (C) ouabain, and (D) ouabagenin (a dihydroouabain product standard was not available).

Compared to digoxin, the isolated lactone 2(5H)-furanone was minimally processed as assessed by colorimetric assays (Figure 63). We repeated these experiments to determine whether electron transfer was accompanied by γ-butyrolactone product formation. Again, we observed modest levels of MV oxidation in the presence of 2(5H)-furanone (Figure 65A) that increased with higher Cgr2 concentrations (Figure 65B). However, even after 24 hours, minimal γ-butyrolactone was produced as detected by high-resolution GC-MS (Figure 65C). Furthermore, no significant differences were observed between a no enzyme control and samples containing Cgr2 (Figure 65C). These results demonstrate that 2(5H)-furanone is a poor substrate for Cgr2, and that a *cis-trans-cis* steroid core (or a steroid fragment) greatly enhances binding and processing by the enzyme. However, the qualitatively similar rates observed for reduction of the various cardenolides demonstrates that the enzyme tolerates differences in the number and position of hydroxyl groups on the steroid scaffold. No activity was observed toward any other substrate tested, including fumarate and urocanate, the substrates of the most similar protein homologs of Cgr2, as well as ascorbic acid, host steroids and prostaglandins, and various abundant dietary furanones

(Figure 63). Overall, these results suggest that Cgr2 activity is restricted to cardenolides and does not extend to other structurally related endogenous or exogenous compounds.

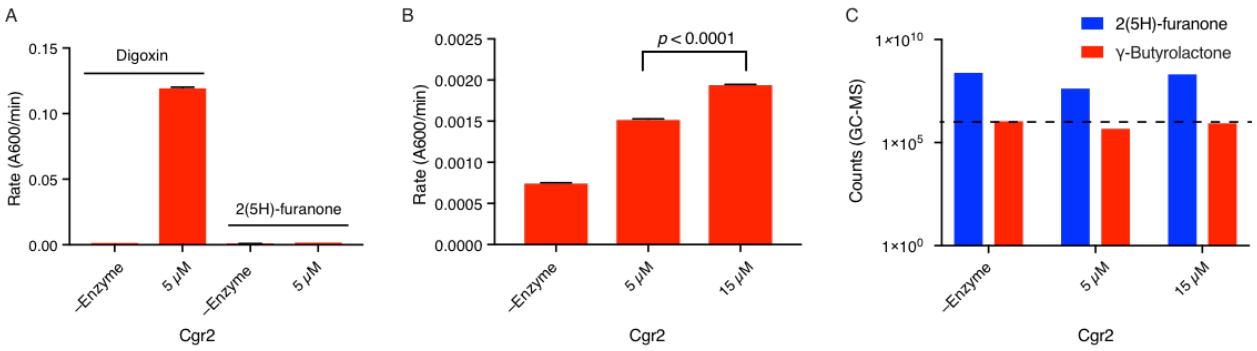


Figure 65: Modest MV oxidation is observed in *in vitro* reactions of Cgr2 with 2(5H)-furanone, although minimal product is formed.
(A) Initial *in vitro* rates of methyl viologen oxidation coupled to digoxin or 2(5H)-furanone reduction by Cgr2. (B) Higher rates of methyl viologen oxidation were observed for 2(5H)-furanone with increasing enzyme concentrations. Reaction conditions for panels (A) and (B) were 5 µM or 15 µM Cgr2, 500 µM FAD, 1.5 mM MV, 1 mM sodium dithionite, and 500 µM of substrate. Values represent mean ± SEM (n = 3) and statistical significance by Student's *t* test is shown. (C) *In vitro* reactions (24 hours) were extracted into ethyl acetate and analyzed by GC-MS. Minimal product formation was observed, and no significant differences were observed for reactions containing 0, 5, or 15 µM Cgr2.

4.4. Discussion

In this work, we aimed to understand the conditions that lead to inter-individual differences in digoxin metabolism and decipher the physiological impact of this metabolism for *E. lenta* and the human host. Our *in vitro* system enabled us to measure the kinetics of digoxin reduction by Cgr2, study the effects of L-arginine on activity, and assess the substrate scope and specificity of this enzyme. Using a combination of culturing and *in vitro* experiments, we hoped to clarify whether digoxin is the endogenous substrate for this enzyme and what benefits this metabolic pathway might confer on *cgr*$^+$ *E. lenta* strains.

Over several decades, scientists have uncovered dietary and geographic factors that are associated with differential digoxin metabolism in humans. Higher consumption of animal protein was correlated with increased production of dihydrodigoxin in human volunteers (*3*), whereas L-arginine inhibited *cgr* operon transcription and digoxin metabolism by pure cultures of *E. lenta.* Dietary protein also inhibited digoxin metabolism in GF mice that were mono-colonized with the bacteria (*5*). This variability suggested that dietary components could affect digoxin metabolism through direct interactions with the *cgr* pathway. However, using our *in vitro* assay, we showed that L-arginine does not directly inhibit Cgr2 (Figure 50), and may instead impart its effects indirectly by altering global cellular processes. Although L-arginine has been more thoroughly studied in the context of digoxin metabolism, other aspects of the arginine dihydrolase pathway (Figure 49) may also contribute to variable digoxin metabolism *in vivo,* and merit further investigation. For example, while L-arginine leads to decreased transcription of the *cgr* operon, is downstream metabolite L-ornithine actually results in higher transcription (>1000x and ~250-fold induction with and without L-ornithine, respectively) (*5*). Elevated levels of L-ornithine may prevent import of L-arginine into the cell by the L-arginine/L-ornithine antiporter or inhibit the ornithine transcarbamylase enzyme. Together, these observations demonstrate the complexities of dietary-xenobiotic interactions in human subjects. To fully understand digoxin metabolism in patients, it will be essential not only to determine whether *cgr*$^+$ *E. lenta* strains are present, but also to address the variability and extent of metabolism of dietary components (protein, amino acids, etc.) and their interplay with the *cgr* operon.

While the implications of digoxin metabolism are clinically significant for patients, the physiological effects of this activity for *E. lenta* remain unclear. The incredibly high amino acid sequence conservation (Figure 27) and high levels of *cgr* operon transcription in response to digoxin incubation (*5*) suggested that digoxin metabolism might provide a physiological benefit to *E. lenta*. We initially postulated that digoxin could serve as an alternative terminal electron acceptor, as its metabolism is down-regulated by L-arginine, a preferred energy source for *E. lenta* (*9*). However, unlike in the related fumarate and urocanate reductase systems (*11, 12*), substrate reduction did not affect *E. lenta* growth in any of the minimal or rich medias tested (Figure 51 and Figure 52), and preliminary data in GF mice suggests that *cgr2*$^{+}$ strains do not have a clear competitive advantage over their *cgr2*$^{-}$ counterparts (preliminary data, Turnbaugh lab). Additionally, Cgr2 displayed low *in vitro* activity towards digoxin in comparison to that of related reductases towards their native substrates (Table 7). This observation prompted us to investigate whether digoxin is the endogenous substrate of the Cgr enzymes by assessing the activity of Cgr2 toward a panel of alternative candidate substrates. However, Cgr2 appears to be restricted to the cardenolides (Figure 63), and could not reduce a variety of chemically similar, abundant, and physiologically relevant compounds (Section **4.3**). These results suggest that cardenolides are the endogenous substrates of Cgr2. However, several explanations are provided below which attempt to disentangle the discrepancies of this system, including the high amino acid conservation of the *cgr* operon, the unknown physiological implication of this pathway for *E. lenta* strains, the low *in vitro* activity of Cgr2, and the narrow substrate scope of this enzyme.

This specificity of Cgr2 towards cardenolides was surprising as digoxin and digitoxin are ingested at low concentrations (*22, 23*) and these compounds are rarely encountered outside of the context of cardiac therapy, except for acute poisoning. There have been numerous reports that humans synthesize endogenous cardiac glycosides that regulate heart rate (*24, 69, 70*). This would suggest that cardiac glycosides are not only used as a defense mechanism against other organisms (e.g. plant toxins, toad venoms), but also regulate (healthy) physiological processes within an organism by engaging with specific cellular receptors (*24, 71*). If these findings are true, these "endogenous digitalis-like factors"

could exert a constant, selective pressure on *E. lenta* strains to maintain Cgr2. However, these results are contentious as many studies relied on antibody-based detection methods, and later investigations with NMR and MS methods failed to recapitulate these findings (*24*). Thus, it is possible that digoxin metabolism by *E. lenta* is simply adventitious and we cannot rule out the possibility that Cgr2 has an as-yet unidentified endogenous substrate that is present at higher concentrations in the human gut, or whose reduction imparts a distinct physiological advantage to these bacteria.

If digoxin or another cardenolide are indeed the substrates of Cgr2, several plausible explanations exist for the low observed *in vitro* activity of this enzyme. The most straightforward explanation is that limitations in our *in vitro* system (e.g. heterologous expression, chemical reconstitution, inability to access Cgr1) could contribute to lower activity than would be expected for an enzyme towards it endogenous substrate. After optimizing heterologous expression (Section **3.1**), metallocofactor reconstitution methods (Section **3.3**), and *in vitro* biochemical assays (Section **4.2**), we identified conditions that reproducibly lead to digoxin reduction by Cgr2. However, the *in vitro* activity of Cgr2 may be artificially low due to incomplete metallocofactor reconstitution (Section **3.3**), poor FAD binding (Section **3.5**), poor substrate binding, and lower rates of electron transfer from an artificial electron donor as compared to the predicted biological reductase. Cgr1 is likely essential for digoxin reduction in *E. lenta* and would presumably improve both the stability and activity of Cgr2 *in vitro*. However, we were unable to successfully obtain any Cgr1 using a variety of protein constructs, heterologous hosts, and expression conditions (Section **3.1**). Further attempts to reconstitute the full Cgr complex may include heterologous expression in additional hosts that may be more suitable for producing cytochromes *c* (Cgr1) or natively purifying both proteins from *E. lenta* DSM2243 (Section **5.2.3**). In addition, an improved [Fe-S] reconstitution method may enhance the stability and activity of Cgr2. This may be achieved *in vitro* using different reconstitution methods (e.g. using a cysteine desulfurase enzyme to provide sulfide equivalents), or through co-expression of Cgr2 with [Fe-S] cluster assembly machinery (*72*). Until we obtain a full, active Cgr complex, it will not be possible to resolve whether digoxin is the endogenous substrate of Cgr2 purely on the basis of kinetic parameters.

Although digoxin reduction has not been linked to a growth or colonization advantage for *E. lenta* strains, it is possible that this pathway may impart an advantage under particular conditions that we have not yet recapitulated in culture or in a model host. Although we tested *E. lenta* growth under several conditions (Section **4.1**), we do not know where *E. lenta* colonizes the human GI tract. This gut biogeography would dictate which environmental cues (e.g. chemicals, pH levels, nutrients) and other microorganisms *E. lenta* would encounter. These factors could alter bacterial metabolic activities, including digoxin reduction, and their biological implications. While preliminary studies in germ-free mice did not show higher colonization by *cgr*$^+$ *E. lenta* strains as compared to *cgr*$^-$ strains, these strains were not otherwise isogenic. Until genetic manipulation is possible in this organism, it will be hard to definitively conclude whether the *cgr* operon is beneficial for *E. lenta*.

Alternatively, digoxin metabolism may not actually be beneficial in the context of the human gut. Although *E. lenta* strains have only been found in humans, at some point, this bacterium or its close relatives at may have lived in close proximity to cardenolide-producing plants or resided in the GI tracts of animals that ingest these plants. Outside of cardenolide-rich environments, *E. lenta* strains may no longer have an evolutionary impetus to maintain this metabolic capability and may be losing rather than acquiring the *cgr* operon over time. However, as the *cgr* operon is highly conserved, restricted to human-associated *E. lenta* strains (Section **2.4**) and distinct from other enzymes (Section **2.2**), phylogenetic analyses cannot be used to test this hypothesis.

A final, provocative possibility is that *cgr*$^+$ *E. lenta* strains may have evolved to protect humans from cytotoxic cardenolides. Although this bacterium is often present in low abundance, it is commonly found in the human gut microbiota (Figure 30). Furthermore, *E. lenta* metabolizes a wide range of human-derived and plant-based substrates into metabolites with beneficial properties (*1, 6-8, 55*). Thus *E. lenta* likely has a strong co-evolutionary history with the human host. Although toxin degradation has been observed in the gut microbiotas of insects (*73-75*), our studies would provide the first example of a specific human gut bacterial enzyme dedicated to inactivation of plant toxins. This gut bacterial pathway

may augment the detoxifying action of host intestinal and hepatic enzymes (Section **1.2.1**), helping to maintain GI health and ensure a habitat for *E. lenta* colonization.

4.5. Experimental

Growth curves

To understand whether metabolism of digoxin or other cardiac glycoside substrates conferred a growth advantage, *E. lenta* DSM2243 was grown in either rich (BHI) or defined media. Basal media lacking terminal electron acceptors was prepared as previously described (*13*) with the following modifications: yeast extract and tryptone were each added to 0.1% (w/v), L-cysteine concentration was 0.4 mM, sodium sulfide was not added, and either 5% H_2 or 10 mM sodium acetate were used as electron donors. Starter cultures were prepared as previously described (Section **2.6**) in BHI media supplemented with 1% arginine, and diluted 1:100 into media that had been supplemented with substrates (dissolved in DMF) to a final concentration of 10 μM. Cultures were grown anaerobically at 37 °C in biological triplicate. OD_{600} measurements were recorded on a Genesys20 spectrophotometer (Thermo Scientific).

Anaerobic DMSO reductase bioinformatics

The genes required for anaerobic DMSO reduction were identified using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (http://www.genome.jp/kegg/) (*76*). All necessary anaerobic DMSO reductase genes (Elen_0422-0426, Elen_0466, Elen_0609, Elen_1190-1, Elen_2787, Elen_2790) and ATP synthase subunits (Elen_1039-1046) were present in *E. lenta* DSM2243 (KEGG database T00984).

In vitro digoxin reduction assay in the presence of L-arginine

Methyl viologen (paraquat) dichloride hydrate (MV) that had been reduced with sodium dithionite was used as an artificial electron donor (*14*) to initiate anaerobic Cgr2-mediated reduction of digoxin *in vitro*. Assays were carried out at 25 °C in an anaerobic glovebox (Mbraun) under an atmosphere of nitrogen and < 5 ppm oxygen. Reagents (all purchased from Sigma-Aldrich) were brought into the glovebox as solids or sparged liquids and resuspended in anoxic buffer inside the chamber: FAD, MV, and L-arginine were resuspended in 50 mM HEPES, 100 mM NaCl, pH 7 to generate stock

solutions of 1 mM, 50 mM, and 100 mM respectively; sodium dithionite was resuspended in 50 mM HEPES, 100 mM NaCl, pH 8 to generate a stock solution of 25 mM. The final assay mixture (100 µL) contained 5 µM Cgr2 Y333, 50 µM FAD, 0.375 mM MV, 0.25 mM dithionite, 0 - 5 mM L-arginine, and the reaction was initiated by addition of 0.5 mM digoxin. Assays were prepared in triplicate in a 96-well polysterene microplate (Corning) and activity was continuously monitored by measuring the absorbance at 600 nm on a PowerWave HT Microplate Spectrophotometer (BioTek); a decrease in the peak absorbance corresponded to MV oxidation coupled to substrate reduction. For endpoint assays, 20 µL reaction aliquots were quenched in 180 µL of methanol, diluted to a final concentration of 1 µM in 50% methanol, and analyzed by mass spectrometry as described in Section **2.6**.

In vitro digoxin reduction assay using different electron donors

The following electron donors were purchased from Sigma-Aldrich: sodium dithionite, methyl viologen (MV), benzyl viologen (BV), diquat (DQ), and phenosafranin (PS). The final assay mixture (100 µL in 50 mM HEPES, 100 mM NaCl, pH 7) contained 5 µM Cgr2 Y333, 50 µM FAD, 2 mM electron donor, 1 mM dithionite, and was initiated by addition of 0.5 mM digoxin. Assays were prepared in triplicate in a 96-well polysterene microplate (Corning) and activity was continuously monitored by measuring the absorbance at 600 nm (MV), 555 nm (BV), 540 nm (PS) or 460 nm (DQ) on a PowerWave HT Microplate Spectrophotometer (BioTek); a decrease in the peak absorbance corresponded to electron mediator oxidation coupled to substrate reduction. Reactions using sodium dithionite as the electron donor could not be monitored colorimetrically. For endpoint assays, 20 µL reaction aliquots were quenched in 180 µL methanol, diluted to a final concentration of 1 µM in 50% methanol, and analyzed by mass spectrometry as described in Section **2.6**.

In vitro digoxin reduction assays in the presence of detergent

Assays were run in triplicate (100 µL) in 50 mM HEPES, 100 mM NaCl, pH 7 buffer containing 5 µM Cgr2 Y333, 50 µM FAD, 0.375 mM MV, and 0.25 mM sodium dithionite, and a variable

concentration of *n*-Dodecyl β-D-maltoside (DDM) detergent. Reactions were initiated by addition of digoxin as a solution in DMF to a final concentration 0.5 mM. Final concentrations of 0.015% and 0.03% (w/w) DDM corresponded to 2x and 4x the critical micelle concentration (CMC) for this detergent. Assays were prepared in a 96-well polysterene microplate (Corning) and activity was continuously monitored by measuring the absorbance at 600 nm on a PowerWave HT Microplate Spectrophotometer (BioTek).

Kinetics assays

Kinetic assays were performed in an anaerobic glovebox (Mbraun) at 25 °C. Initial assays (Figure 53) were run in triplicate (200 µL) in assay buffer containing 5 µM Cgr2 Y333, 50 µM FAD, 0.375 mM MV, and 0.25 mM sodium dithionite, and reactions were initiated by addition of digoxin as a solution in DMF to a final concentration of 0.05, 0.1, 0.2, 0.3, or and 0.5 mM. Optimized assays (Figure 56) were run in triplicate (200 µL) in assay buffer containing 5 µM Cgr2 (Y333 or N333), 500 µM FAD, 1.5 mM MV, and 1 mM sodium dithionite, and reactions were initiated by addition of digoxin as a solution in DMF to a final concentration of 0.01, 0.025, 0.05, 0.1, 0.2, 0.25, 0.3, or 0.5 mM. 20 µL reaction aliquots were quenched in 180 µL of ice-cold methanol in Costar® flat bottom polysteryene 96-well plates (Corning). The plates were sealed with adhesive aluminum foil for 96-well plates (VWR), brought out of the anaerobic chamber, and further diluted (50-fold) into 50% methanol. Digoxin and dihydrodigoxin standard curves were prepared in the full reaction matrix and processed identically such that final concentrations (after 500x total dilution) generated a standard curve between 0.01- 1 µM. Plates were centrifuged (4000 rpm x 10 min, 4 °C) and 200 µL of each reaction were transferred to a 0.5 mL PP 96-well plate (Agilent Technologies) sealed with EPS easy piercing seals (BioChromato). The reactions were monitored by LC-MS/MS as previously described (Section **2.6**), except that samples were directly injected (no column), and isocratic flow was used with 75% methanol with 1 mM ammonium hydroxide.

In vitro substrate reduction assays

All substrates (Figure 63) were dissolved in DMF to generate stock solutions of 25 mM, with the exception of sodium fumarate dibasic and urocanic acid, which were dissolved in water. All substrates and reagents were purchased from Sigma-Aldrich except for the bufadienolides (Enzo Life Sciences) and prostaglandins (Cayman Chemicals). The final assay mixture (100 µL in 50 mM HEPES, 100 mM NaCl, pH 7) contained 5 µM Cgr2 Y333, 50 µM FAD, 0.375 mM MV, 0.25 mM dithionite, and was initiated by addition of 0.5 mM substrate. Assays were prepared in a 96-well polysterene microplate (Corning) and activity was continuously monitored by measuring the absorbance at 600 nm on a PowerWave HT Microplate Spectrophotometer (BioTek). For endpoint assays, 20 µL reaction aliquots were quenched in methanol, diluted to a final concentration of 1 µM in 50% methanol, and analyzed by mass spectrometry as described below.

MS analysis of substrate reduction assays

In vitro reactions with cardenolide substrates were analyzed by MS_1 scan on an Agilent 6410 Triple Quad LC/MS using electrospray ionization in negative ion mode. The instrument settings were as follows: gas temperature (300°C), gas flow (10 L/min), nebulizer pressure (25 psi), capillary voltage (4000 V), fragmentor voltage (250 V) and chamber current (0.1 µA), buffer (0.125 mL/min flow with 75% methanol with 1 mM ammonium hydroxide), and direct injection (5 µL sample). Cardenolides were purchased from Sigma-Aldrich, and dihydrocardenolide standards were obtained through hydrogenation of digoxin as previously described (*5*). For 2(5H)-furanone samples, triplicate *in vitro* reactions were combined after 24 hours and extracted five times with ethyl acetate. The pooled organic fractions were concentrated using a rotary evaporator, and samples were resuspended in 300 µL of ethyl acetate. Samples were analyzed by high resolution GC-MS by Jennifer Wang at the Small Molecule Mass Spectrometry Facility at Harvard University. 2(5H)-furanone and γ-butyrolactone standards were purchased from Sigma-Aldrich.

Chemical similarity analysis

The chemical similarity of all substrates was assessed using the ChemMine software (http://chemminetools.ucr.edu) (*68*). Substrates were imported into ChemMine in SMILES format. The hierarchical clustering tool was used to generate a heatmap visualizing the structural distance matrix between each substrate and digoxin.

4.6. References

1. J. F. Dobkin, J. R. Saha, V. P. Butler, H. C. Neu, J. Lindenbaum, Digoxin-inactivating bacteria: Identification in human gut flora. *Science* **220**, 325-327 (1983).

2. A. N. Alam, J. R. Saha, J. F. Dobkin, J. Lindenbaum, Interethnic variation in the metabolic inactivation of digoxin by the gut flora. *Gastroenterology* **95**, 117-123 (1988).

3. V. I. Mathan, J. Wiederman, J. F. Dobkin, J. Lindenbaum, Geographic differences in digoxin inactivation, a metabolic activity of the human anaerobic gut flora. *Gut* **30**, 971-977 (1989).

4. L. A. David *et al.*, Diet rapidly and reproducibly alters the human gut microbiome. *Nature* **505**, 559-563 (2014).

5. H. J. Haiser *et al.*, Predicting and manipulating cardiac drug inactivation by the human gut bacterium *Eggerthella lenta*. *Science* **341**, 295-298 (2013).

6. M. Kutschera, W. Engst, M. Blaut, A. Braune, Isolation of a catechin-converting human intestinal bacteria. *Journal of Applied Microbiology* **111**, 165-175 (2011).

7. T. Clavel *et al.*, Intestinal bacterial communities that produce active estrogen-like compounds enterodiol and enterolactone in humans. *Applied and Environmental Microbiology* **71**, 6077-6085 (2005).

8. F. Rafii, The role of colonic bacteria in the metabolism of the natural isoflavone daidzin to equol. *metabolites* **5**, 56-73 (2015).

9. J. F. Sperry, T. D. Wilkins, Arginine, a growth-limiting factor for *Eubacterium lentum*. *J Bacteriol* **127**, 780-784 (1976).

10. L. Novak *et al.*, Arginine deiminase pathway enzymes: evolutionary history in metamonads and other eukaryotes. *BMC Evol Biol* **16**, 197 (2016).

11. A. V. Bogachev, Y. V. Bertsova, D. A. Bloch, M. I. Verkhovsky, Urocanate reductase: identification of a novel anaerobic respiratory pathway in Shewanella oneidensis MR-1. *Mol Microbiol* **86**, 1452-1463 (2012).

12. I. Schroder, R. P. Gunsalus, B. A. Ackrell, B. Cochran, G. Cecchini, Identification of active site residues of *Escherichia coli* fumarate reductase by site-directed mutagenesis. *J Biol Chem* **266**, 13572-13579 (1991).

13. F. E. Löffler, R. A. Sanford, K. M. Ritalahti, Enrichment, Cultivation, and Detection of Reductively Dechlorinating Bacteria. **397**, 77-111 (2005).

14. T. Watanabe, K. Honda, Measurement of the extinction coefficient of the methyl viologen cation radical and the efficiency of its formation by semiconductor photocatalysis. *Journal of Physical Cemistry* **86**, 2617-2619 (1982).

15. J. Simon *et al.*, The tetraheme cytochrome *c* NrfH is required to anchor the cytochrome *c* nitrite reductase (NrfA) in the membrane of *Wolinella succinogenes*. *Eur J Biochem* **268**, 5776-5782 (2001).

16. L. Michaelis, E. S. Hill, The viologen indicators. *J Gen Physiol* **16**, 859-873 (1933).

17. K. Chen *et al.*, Alteration of the reduction potential of the [4Fe-4S](2+/+) cluster of *Azotobacter vinelandii* ferredoxin I. *J Biol Chem* **274**, 36479-36487 (1999).

18. A. Sharma, B. Purkait, Identification of medicinally active ingredient in ultradiluted *Digitalis purpurea*: Fluorescence spectroscopic and cyclic-voltammetric study. *J Anal Methods Chem* **2012**, 109058 (2012).

19. A. Rohman, N. van Oosterwijk, A. M. Thunnissen, B. W. Dijkstra, Crystal structure and site-directed mutagenesis of 3-ketosteroid Delta1-dehydrogenase from *Rhodococcus erythropolis* SQ1 explain its catalytic mechanism. *J Biol Chem* **288**, 35559-35568 (2013).

20. C. J. Morris, Black, A.C., Pealing, S.L., Manson, F.D.C., Chapman, S.K., Reid, G.A., Gibson, D.M., Ward, F.B., Purification and properties of a novel cytochrome: flavocytochrome c from Shewanella putrefaciens. *Biochem J* **302**, 587-593 (1994).

21. G. L. Kemp *et al.*, Kinetic and thermodynamic resolution of the interactions between sulfite and the pentahaem cytochrome NrfA from Escherichia coli. *Biochem J* **431**, 73-80 (2010).

22. M. Gheorghiade, K. F. Adams, W. S. Colucci, Digoxin in the management of cardiovascular disorders. *Circulation* **109**, 2959-2964 (2004).

23. T. W. Smith, V. P. Butler, E. Haber, Determination of therapeutic and toxic serum digoxin concentrations by radioimmunoassay. *N Engl J Med* **281**, 1212-1216 (1969).

24. M. Michalak, K. Michalak, J. Wicha, The synthesis of cardenolide and bufadienolide aglycones, and related steroids bearing a heterocyclic subunit. *Nat Prod Rep* **34**, 361-410 (2017).

25. H. M. Gao, R. Popescu, B. Kopp, Z. M. Wang, Bufadienolides and their antitumor activity. *Natural Product Reports* **28**, 953-969 (2011).

26. K. Greeff, T. Akera, *Cardiac glycosides*. Handbook of experimental pharmacology (Springer-Verlag, Berlin ; New York, 1981).

27. K. Winnicka, K. Bielawski, A. Bielawski, Cardiac glycosides in cancer research and cancer therapy. *Acta poloniae pharmaceutica - Drug research* **63**, 109-115 (2006).

28. P. S. Steyn, F. R. van Heerden, Bufadienolides of plant and animal origin. *Nat Prod Rep* **15**, 397-413 (1998).

29. K. M. Weigand *et al.*, Na(+),K(+)-ATPase isoform selectivity for digitalis-like compounds is determined by two amino acids in the first extracellular loop. *Chem Res Toxicol* **27**, 2082-2092 (2014).

30. M. Laursen, J. L. Gregersen, L. Yatime, P. Nissen, N. U. Fedosova, Structures and characterization of digoxin- and bufalin-bound Na+,K+-ATPase compared with the ouabain-bound complex. *Proc Natl Acad Sci U S A* **112**, 1755-1760 (2015).

31. Digitalis Investigation Group, The effect of digoxin on mortality and morbidity in patients with heart failure. *N Engl J Med* **336**, 525-533 (1997).

32. F. Kayali, M. A. Janjua, D. A. Laber, D. Miller, G. Kloecker, Phase II trial of second-line erlotinib and digoxin for nonsmall cell lung cancer (NSCLC). *Open Access Journal of Clinical Trials* **3**, 9-13 (2011).

33. J. Lin *et al.*, A pilot phase II Study of digoxin in patients with recurrent prostate cancer as evident by a rising PSA. *Am J Cancer Ther Pharmacol* **2**, 21-32 (2014).

34. E. L. Oiestad, U. Johansen, M. Stokke Opdal, S. Bergan, A. S. Christophersen, Determination of digoxin and digitoxin in whole blood. *J Anal Toxicol* **33**, 372-378 (2009).

35. D. Perrier, M. Mayersohn, F. I. Marcus, Clinical pharmacokinetics of digitoxin. *Clin Pharmacokinet* **2**, 292-311 (1977).

36. I. Bauer, P. Neubert, W. Schaumann, Pharmacodynamics, pharmacokinetics and metabolism of digitoxin and derivatives in cats. *Naunyn Schmiedebergs Arch Pharmacol* **335**, 469-475 (1987).

37. A. Schmoldt, Increased digitoxin cleavage by liver microsomes of spironolactone-pretreated rats. *Naunyn Schmiedebergs Arch Pharmacol* **305**, 261-263 (1978).

38. H. F. Vohringer, L. Weller, N. Rietbrock, Influence of spironolactone pretreatment on pharmacokinetics and metabolism of digitoxin in rats. *Naunyn Schmiedebergs Arch Pharmacol* **287**, 129-139 (1975).

39. G. Bodem, E. V. Unruh, Enhanced transformafion of digitoxin to dihydrodigitoxin in humans with renal failure. *The Journal of Clinical Pharmacology* **19**, 195-199 (1979).

40. P. Terness, D. Navolan, C. Dufter, B. Kopp, G. Opelz, The T-cell suppressive effect of bufadienolides: structural requirements for their immunoregulatory activity. *Int Immunopharmacol* **1**, 119-134 (2001).

41. X. C. Ma *et al.*, Simultaneous quantification of seven major bufadienolides in three traditional Chinese medicinal preparations of chansu by HPLC-DAD. *Nat Prod Commun* **4**, 179-184 (2009).

42. J. Colin Slaughter, The naturally occurring furanones: formation and function from pheromone to food. *Biol Rev Camb Philos Soc* **74**, 259-276 (1999).

43. F. Ledl, E. Schleicher, New aspects of the Maillard reaction in foods and in the human body. *Angewandte Chemie International Edition in English* **29**, 565-594 (1990).

44. N. Montazeri, A. C. Oliveira, B. H. Himelbloom, M. B. Leigh, C. A. Crapo, Chemical characterization of commercial liquid smoke products. *Food Sci Nutr* **1**, 102-115 (2013).

45. W. S. Sung *et al.*, 2,5-dimethyl-4-hydroxy-3(2H)-furanone (DMHF); antimicrobial compound with cell cycle arrest in nosocomial pathogens. *Life Sci* **80**, 586-591 (2007).

46. G. A. Reid, C. S. Miles, R. K. Moysey, K. L. Pankhurst, S. K. Chapman, Catalysis in fumarate reductase. *Biochim Biophys Acta* **1459**, 310-315 (2000).

47. L. Verbist, The antimicrobial activity of fusidic acid. *J Antimicrob Chemother* **25 Suppl B**, 1-5 (1990).

48. A. Kroger, V. Geisler, E. Lemma, F. Theis, R. Lenger, Bacterial fumarate respiration. *Arch Microbiol* **158**, 311-314 (1992).

49. J. E. Comes, R. B. Beelman, Addition of fumaric acid and sodium benzoate as an alternative method to achieve a 5-log reduction of *Escherichia coli* O157:H7 populations in apple cider. *J Food Prot* **65**, 476-483 (2002).

50. N. H. Litjens *et al.*, Pharmacokinetics of oral fumarates in healthy subjects. *Br J Clin Pharmacol* **58**, 429-432 (2004).

51. P. Geisser, S. Burckhardt, The pharmacokinetics and pharmacodynamics of iron preparations. *Pharmaceutics* **3**, 12-33 (2011).

52. N. K. Gibbs, J. Tye, M. Norval, Recent advances in urocanic acid photochemistry, photobiology and photoimmunology. *Photochem Photobiol Sci* **7**, 655-667 (2008).

53. W. von Daehne, W. O. Godtfredsen, P. R. Rasmussen, Structure-activity relationships in fusidic acid-type antibiotics. *Adv Appl Microbiol* **25**, 95-146 (1979).

54. J. M. Ridlon, D. J. Kang, P. B. Hylemon, Bile salt biotransformation by human intestinal bacteria. *J. Lipid Res.* **47**, 241-259 (2006).

55. A. S. Devlin, M. A. Fischbach, A biosynthetic pathway for a prominent class of microbiota-derived bile acids. *Nat Chem Biol* **11**, 685-690 (2015).

56. S. Thevarajah, M. Polanceczky, E. J. Scherl, C. L. Frissora, Hormonal influences on the gastrointestinal tract and irritable bowel syndrome. *Practical Gastroenterology*, 62-74 (2005).

57. G. Zheng *et al.*, Corticosterone mediates stress-related increased intestinal permeability in a region-specific manner. *Neurogastroenterol Motil* **25**, e127-139 (2013).

58. M. W. Pfaffl, I. G. Lange, H. H. D. Meyer, The gastrointestinal tract as target of steroid hormone action: Quantification of steroid receptor mRNA expression (AR, ER alpha, ER beta and PR) in 10 bovine gastrointestinal tract compartments by kinetic RT-PCR. *J Steroid Biochem* **84**, 159-166 (2003).

59. F. Z. Stanczyk, S. Roy, Metabolism of levonorgestrel, norethindrone, and structurally related contraceptive steroids. *Contraception* **42**, 67-96 (1990).

60. H. Adlercreutz, M. O. Pulkkinen, E. K. Hamalainen, J. T. Korpela, Studies on the role of intestinal bacteria in metabolism of synthetic and natural steroid hormones. *J Steroid Biochem* **20**, 217-229 (1984).

61. D. Ren, L. Li, A. W. Schwabacher, J. W. Young, D. C. Beitz, Mechanism of cholesterol reduction to coprostanol by *Eubacterium coprostanoligenes* ATCC 51222. *Steroids* **61**, 33-40 (1996).

62. S. L. Waller, Prostaglandins and the gastrointestinal tract. *Gut* **14**, 402-417 (1973).

63. T. Lawrence, D. A. Willoughby, D. W. Gilroy, Anti-inflammatory lipid mediators and insights into the resolution of inflammation. *Nat Rev Immunol* **2**, 787-795 (2002).

64. E. W. Horton, R. L. Jones, Prostaglandins A1,A2 and 19-hydroxy A1; their actions on smooth muscle and their inactivation on passage through the pulmonary and hepatic portal vascular beds. *Br J Pharmacol* **37**, 705-722 (1969).

65. E. Ricciotti, G. A. FitzGerald, Prostaglandins and inflammation. *Arterioscler Thromb Vasc Biol* **31**, 986-1000 (2011).

66. I. Brook, E. H. Frazier, Significant recovery of nonsporulating anaerobic rods from clinical specimens. *Clin Infect Dis* **16**, 476-480 (1993).

67. J. M. Sahuquillo-Arce *et al.*, *Eggerthella lenta* bacteraemia in a patient with Caroli disease. *JMM Case Reports* **2**, 1-3 (2015).

68. T. W. Backman, Y. Cao, T. Girke, ChemMine tools: an online service for analyzing and clustering small molecules. *Nucleic Acids Res* **39**, W486-491 (2011).

69. W. Schoner, G. Scheiner-Bobis, Endogenous and exogenous cardiac glycosides and their mechanisms of action. *American Journal of Cardiovascular Drugs* **7**, 173-189 (2007).

70. Y. Komiyami, T. Mori, T. Murakami, M. Masuda, H. Takahashi, Production of ouabain-like factors of hypothalamo-pituitary origin, determined by a sensitive ELISA for ouabain, is increased in DOCA-salt hypertensive rats. *Pathophysiology* **2**, 35-40 (1995).

71. A. Y. Bagrov, J. I. Shapiro, O. V. Fedorova, Endogenous cardiotonic steroids: physiology, pharmacology, and novel therapeutic targets. *Pharmacol Rev* **61**, 9-38 (2009).

72. C. Ayala-Castro, A. Saini, F. W. Outten, Fe-S cluster assembly pathways in bacteria. *Microbiol Mol Biol Rev* **72**, 110-125 (2008).

73. C. U. Welte, J. F. Rosengarten, R. M. de Graaf, M. S. Jetten, SaxA-mediated isothiocyanate metabolism in phytopathogenic *Pectobacteria*. *Appl Environ Microbiol* **82**, 2372-2379 (2016).

74. C. U. Welte *et al.*, Plasmids from the gut microbiome of cabbage root fly larvae encode SaxA that catalyses the conversion of the plant toxin 2-phenylethyl isothiocyanate. *Environ Microbiol* **18**, 1379-1390 (2016).

75. N. A. Broderick, K. F. Raffa, R. M. Goodman, J. Handelsman, Census of the bacterial community of the gypsy moth larval midgut by using culturing and culture-independent methods. *Appl Environ Microbiol* **70**, 293-300 (2004).

76. M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27-30 (2000).

Chapter 5: Future directions and concluding remarks

5.1. Conclusions

Since the early 1980s, the human gut bacterium *E. lenta* was known to be responsible for inactivating the cardiac medication digoxin (*1*). However, the identity, specificity, and distribution of the enzymes responsible for this activity were unknown. We have unambiguously demonstrated that the *E. lenta* protein Cgr2 inactivates cardenolides, including the pharmaceutical agents digoxin and digitoxin, which have been used for over two centuries in the treatment of cardiac diseases. Cgr2 is a unique flavoprotein reductase that contains oxygen-sensitive [4Fe-4S] cluster(s) and a divergent set of predicted active site residues. The failure of bioinformatic analyses to identify this essential [4Fe-4S] cluster highlights the need for additional structural and mechanistic studies of the *cgr* operon and other gut microbial enzymes involved in xenobiotic metabolism. Our preliminary mechanistic model for digoxin reduction is that Cgr1 and Cgr2 form a membrane-anchored, extracellular complex that mediates electron transfer from an unknown electron donor through multiple cytochromes *c* in Cgr1 to the [4Fe-4S] cluster(s) and FAD cofactor in Cgr2. A hydride equivalent is transferred from the reduced FAD cofactor to the α,β-unsaturated γ-butyrolactone group of digoxin (and other cardenolides), followed by proton transfer from amino acid residues within Cgr2 to the cardenolide intermediate to yield the therapeutically inactive products.

We have demonstrated that *E. lenta* strains harboring the *cgr* operon are widespread in the human gut microbiota, which supports the high incidence of dihydrodigoxin production observed clinically (10 - 35%) (*2, 3*). Strikingly, the sequences of *cgr2* and its associated genes are highly conserved, with only two naturally occurring variants. This conservation is perhaps surprising given the strict specificity of Cgr2 towards cardenolides, which are ingested at very low concentrations to minimize toxicity in the context of cardiac therapy (*4, 5*). Although we cannot rule out the existence of an unidentified endogenous substrate for Cgr2, our results suggest that in a similar manner to intestinal or liver host enzymes, gut bacteria may have evolved enzymes that detoxify ingested xenobiotics to maintain host health. While it is not clear why the *cgr2* gene is maintained in *E. lenta* strains in the absence of a strong

selective pressure, these low abundance bacteria undoubtedly influence drug metabolism and response in humans.

Our results also highlight important considerations that must be addressed in order to accurately predict and manipulate gut microbial xenobiotic metabolism, particularly in the context of therapeutics. Not only do *E. lenta* strains vary in the presence or absence of the *cgr* operon, but we have also identified a naturally occurring, single amino acid substitution that causes a dramatic loss of activity in the Cgr2 enzyme. In addition, dietary factors influence the extent of digoxin metabolism in humans. Moving forward, it will be important to more systematically probe the effects of dietary components on *cgr*$^{+}$ *E. lenta* colonization and digoxin metabolism, to understand whether polymorphisms in the *cgr2* gene lead to different pharmacokinetics in humans, and to determine whether digoxin and dihydrodigoxin display differential activity towards emerging disease targets of this drug (Section **5.3**) (*6, 7*), (*8*) (*9*).

Finally, our work, which progresses from clinical observation of gut microbial metabolism of a pharmaceutical to validation and characterization of the relevant enzymes (*1*), provides a template for identifying the responsible genes, enzymes, and biochemical mechanisms associated with new gut microbial metabolic activities. These studies also highlight the promise of such investigations to reveal new types of chemistry within the complex gut microbial community. Our study of just one unique, clinically relevant gut microbial enzyme has illuminated underappreciated functional diversity within the broader flavin-dependent reductases, which are widespread among human-associated microbes. As reductive transformations represent a major route by which gut microbes metabolize xenobiotics, these unique, putative reductase enzymes provide a promising starting point for identifying additional gut microbial-xenobiotic interactions. As human gut microbial metabolism of dietary components and pharmaceuticals is increasingly implicated in human health and disease states (*10-12*), gaining a molecular understanding of these metabolic pathways may shape therapeutic interventions and ultimately improve human health.

5.2. Further mechanistic studies of Cgr proteins

We have demonstrated that Cgr2 is necessary, sufficient, and specific for cardenolide inactivation in cells and *in vitro*. Using various biochemical and spectroscopic assays, we have learned that Cgr2 is dependent on a FAD cofactor and at least one [4Fe-4S] cluster(s), and its activity is stimulated by divalent metal cations. Below, we describe additional spectroscopic, structural, and biochemical experiments that would augment our mechanistic understanding of Cgr2, as well as the full Cgr complex.

5.2.1 Metallocofactor characterization

We determined that Cgr2 contains at least one [4Fe-4S] cluster that is required for both protein stability and activity, and that divalent metal cations further augment this activity *in vitro* (Section **3.3**). As cysteine thiolates are the most common ligands for [Fe-S] clusters (*13, 14*), we individually mutated all 16 cysteine residues in Cgr2 to alanine and identified six residues that were important for activity, but which did not decrease $[4Fe\text{-}4S]^{1+}$ content (Figure 39). Using biochemical assays and EPR spectroscopy, we could not distinguish among several possible explanations for this observation: (1) Cgr2 contains multiple [Fe-S] clusters, only one of which can be reduced to the EPR-detectable 1+ state; (2) mutation of a single [4Fe-4S] cluster ligand may not be sufficient for disrupting cluster formation (*15-17*); or (3) alternative ligands (histidine, aspartate, serine, or backbone amides) may be involved in cluster ligation (*14*). Additional biochemical, spectroscopic, and structural studies are thus needed to establish the ligands and stoichiometry of the [4Fe-4S] clusters, determine whether additional Fe-containing metallocofactors exist, and understand their roles in Cgr2 structure and/or catalysis.

EPR spectroscopy was crucial for demonstrating that Cgr2 contains redox-active [4Fe-4S] clusters (Figure 36), and this technique may be applied to shed additional light on the catalytic role and physiological significance of our *in vitro* findings. While our EPR experiments were performed on purified or reconstituted Cgr2 that was reduced with sodium dithionite, the redox properties of some [Fe-S] clusters can be altered in the presence of additional cofactors or substrates (*18*). In preliminary experiments, we found that while digoxin did not influence the EPR signal of reduced Cgr2 samples

(Figure 66A-B), addition of FAD to this sample led to a time-dependent decrease in the $[4Fe\text{-}4S]^{1+}$ signal and a concomitant increase in the semiquinone FAD signal (Figure 66C-D). These preliminary results implicate the [4Fe-4S] cluster in transferring electrons to the FAD cofactor in Cgr2. Repeating these EPR experiments under single turnover conditions and with appropriate controls would provide more definitive evidence that the [4Fe-4S] cluster(s) of Cgr2 play a catalytic role in digoxin reduction.

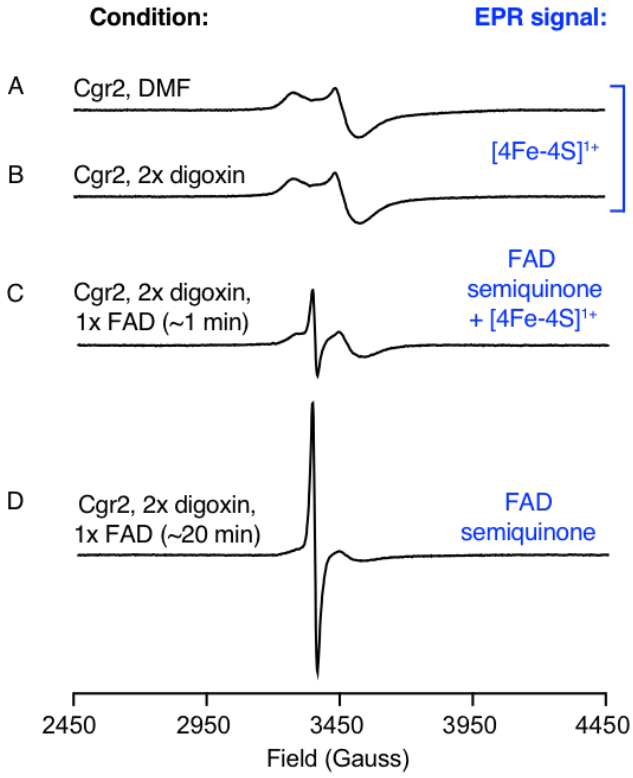


Figure 66: Preliminary EPR experiments suggest that the [4Fe-4S] cluster(s) of Cgr2 transfer electrons to the FAD cofactor.
EPR spectra of Cgr2 samples that had been reconstituted with 8 equivalents of iron and sulfide overnight and reduced with 15 equivalents of sodium dithionite. (A) Samples containing 200 µM of wild-type Cgr2 and (A) DMF or (B) 400 µM of digoxin had identical EPR signals corresponding to $[4Fe\text{-}4S]^{1+}$ clusters. (C) Addition of 1 equivalent of FAD to sample (B) led to a decrease in the EPR signal corresponding to $[4Fe\text{-}4S]^{1+}$ cluster(s) and a concomitant increase in the FAD semiquinone EPR signal that (D) increased with time. EPR measurements were conducted at 10 K, and experimental conditions were microwave frequency 9.38 GHz, microwave power 0.2 mW, modulation amplitude 0.6 mT, and receiver gain 40 dB.

During our attempts to optimize [4Fe-4S] cluster reconstitution, we discovered that divalent metal cations increased the activity of Cgr2 (Section **3.3**). In order to characterize these putative metal centers in Cgr2, we would need to employ new spectroscopic techniques. Unlike EPR spectroscopy, which detects unpaired electrons and is thus only sensitive to particular redox states of [Fe-S] clusters, Mössbauer spectroscopy can provide information about all distinct Fe species within a sample (*13*). This technique relies on the absorption and emission of γ rays from solid samples (e.g. frozen protein samples), and is commonly applied to detect Fe centers in proteins that have been expressed in ^{57}Fe-enriched media or chemically reconstituted with ^{57}Fe reagents (*13*). A Mössbauer spectrum is a composite of signals from all distinct ^{57}Fe species (e.g. $Fe^{2/3+}$, $[2Fe\text{-}2S]^{2/1/0+}$, or $[4Fe\text{-}4S]^{3/2/1/0+}$) in relative proportion to each other (*13*). Thus, reconstituting wild-type Cgr2 with ^{57}Fe and sulfide would enable us to determine whether additional metallocenters are present in the enzyme, or whether Cgr2 binds ^{57}Fe non-specifically. The inactive cysteine to alanine point mutants we have previously identified (Figure 39) could similarly be reconstituted and analyzed to determine whether they are important for direct ligation of metal centers that are not detectable by EPR, including additional [Fe-S] clusters or divalent metal centers. If these cysteine residues are not important for metallocofactor binding, additional mutagenesis experiments of non-cysteine ligands may be informative.

5.2.2 Structural characterization of Cgr2

As only two coding mutations have been identified in all *cgr2* sequences (Section **2.3**), and Cgr2 does not share conserved active site residues with characterized enzymes (Section **2.2**), it is difficult to deduce which active site amino acids are essential for Cgr2 activity. Although our mutagenesis studies of putative metallocenter binding residues (Figure 39) and the natural Y333/N333 variants (Figure 56) in Cgr2 identified several residues that are important for Cgr2 activity, there is a limit to how much we can learn about this enzyme through rational mutagenesis studies. We also generated various homology models of Cgr2 based on the X-ray crystal structures of the most similar homologs to Cgr2 (generated with HHPred modeler; https://toolkit.tuebingen.mpg.de/#/). However, these models contained many

unstructured regions and substrate binding pockets that were likely too small to accommodate digoxin. Coupled with the low sequence identity of Cgr2 to the structurally characterized enzymes (fumarate reductases and ketosteroid dehydrogenases), these models do not provide obvious insights into Cgr2 structure or activity. Structural studies of Cgr2 would therefore greatly aid our mechanistic understanding of digoxin reduction.

In collaboration with the Rosenberg lab (UCSF) and Dr. Manuel Ortega in the Drennan lab (MIT), we attempted to structurally characterize Cgr2 using X-ray crystallography. Crystallization was attempted under aerobic and anaerobic conditions using either purified or reconstituted Cgr2 protein, as well as a range of buffer conditions (pHs, salts, additives, buffering agents, etc.), FAD concentrations, and digoxin concentrations. However, we were unable to obtain reproducible, diffracting Cgr2 crystals. While we have not yet identified conditions that lead to Cgr2 crystallization, our screening has by no means been exhaustive. We found that [4Fe-4S] cluster reconstitution substantially improved Cgr2 stability (Figure 35), but our *in vitro* preps may not be optimally stable. Reconstitution could be further improved either *in vitro* or *in vivo* through co-expression of Cgr2 with [Fe-S] cluster assembly machinery (*14*). Thermal shift assays can also be used to screen and prioritize additional buffer conditions that stabilize Cgr2 (e.g. more acidic pH) (Figure 35B). Several recombinant and chemical modification techniques may also facilitate crystallization, including chemical modification of surface residues (e.g. methylation of surface lysines) (*19*), or generation of fusion proteins (e.g. Glutathione-*S*-transferase or MBP) that are more amenable to forming crystal contacts (*20, 21*). Finally, different protein truncations and tagged constructs may further enhance Cgr2 stability and rigidity. Our original crystallization experiments were attempted with the Cgr2(–48 aa)-Nhis6x construct, which was designed based on the predicted cleavage site from the Tat secretion signal. However, Cgr2 is predicted to contain floppy, disordered regions at both the N- and C- termini, corresponding to residues 49-66 and 551-560 (http://dis.embl.de/) (*22*). Removal of these putative disordered regions could thus enhance protein stability and enable crystallization. Although crystallography of Cgr2 has proven challenging, if successful, would provide invaluable information that cannot currently be achieved by other techniques.

5.2.3 Reconstitution of the full Cgr complex

While we have demonstrated that Cgr2 is necessary and sufficient for digoxin reduction in a heterologous host and *in vitro* using a chemical electron donor, Cgr1 is likely essential for this metabolic activity in *E. lenta.* The ability to reconstitute the entire *cgr* operon (Cgr2 and Cgr1) would presumably enhance the *in vitro* activity of Cgr2. Although expression of cytochromes *c* from (generally Gram-negative) bacteria and eukaryotes has been successful in *E. coli* (*23, 24*), we were unable to observe overexpression or heme *c* incorporation into Cgr1. This lack of expression could be due to an incompatibility between the cytochrome *c* maturation factors of *E. coli* ("system I") and cytochromes *c* enzymes from Gram-positive organisms, whose heme groups are normally installed with "System II" enzymes *(25).* While we also did not observe protein production in the Gram-positive host *R. erythropolis*, Cgr1 overexpression may be more fruitful in other heterologous hosts such as *Shewanella oneidensis* (*26, 27*) or *Wolinella succinogenes* (*28*) which utilize System II to endogenously and heterologously produce high titers of multi-heme cytochromes *c*.

Ideally, Cgr proteins could be overexpressed and isolated from *E. lenta*. As genetic manipulation is not currently possible, Cgr proteins would have to be natively purified from *E. lenta* cells grown in the presence of digoxin. However, this may prove challenging as *E. lenta* is slow growing and it is difficult to achieve high cell densities, especially under conditions in which the *cgr* operon is transcribed (low L-arginine). Alternative growth medias or conditions may be tested, including fermentation or growth in the presence of formate, which was essential for *E. lenta* metabolism of epigallocatechin (*29*), and which has been shown to accelerate *E. lenta* growth in the context of other xenobiotics (Vayu Rekdal). Assuming that the Cgr proteins form a high-affinity complex, recombinant Cgr2-Nhis6x may alternatively be used to pull down and purify Cgr1 from *E. lenta* cells. Ultimately, access to the full Cgr complex would enable more accurate mechanistic studies of digoxin metabolism by *E. lenta*.

5.3. Additional factors that may influence the clinical efficacy of digoxin

Prior to our work, only one *cgr2*$^+$ *E. lenta* strain had been sequenced and linked to digoxin metabolism (*30, 31*). In collaboration with the Turnbaugh lab, we tested an expanded *E. lenta* library and identified 7 new *cgr2*$^+$ strains with remarkable sequence conservation (only 2 amino acid variants) and the ability to produce dihydrodigoxin in culture (Figure 28). In addition, 14 *cgr2* sequences were reconstructed from human gut metagenomes, and only encoded one polymorphism (Y333/N333) (Figure 29). Although our analysis of 28 *cgr2* sequences demonstrates that this gene is highly conserved at the amino acid level, examination of larger *E. lenta* collections and metagenomes may reveal additional genetic diversity that exists within the *cgr* operon. Beyond expanding our analyses to a larger sample size, it will be interesting to assess the prevalence and conservation of *cgr2* across different cohorts. For example, if the *cgr* operon confers a fitness advantage to *E. lenta* strains in humans, one would expect to see higher levels of *cgr*$^+$ *E. lenta* in cardiac patients receiving digoxin treatment relative to patients taking other medications, or to healthy individuals. In addition, as distinct diets are associated with variable levels of dihydrodigoxin production *in vivo*, comparing *cgr* levels in the microbiotas of individuals consuming different foods will help clarify whether particular diets impact *E. lenta* colonization or affect regulation of *cgr* genes.

While we have shown that *cgr2* is a reliable biomarker for digoxin production by *E. lenta* strains, additional questions must be addressed before this gene can be used as a clinical biomarker. Several studies have shown that digoxin-metabolizing strains of *E. lenta* are present in the gut microbiotas of individuals that do not produce dihydrodigoxin *in vivo* (*32, 33*). While dietary factors likely explain some of these discrepancies, *cgr*$^+$ *E. lenta* colonization levels may also affect the extent of digoxin metabolism observed in humans. A direct comparison of colonization levels with dihydrodigoxin production in patients and in controlled animal studies will thus help to determine the threshold of *cgr*$^+$ *E. lenta* required to impart a clinically relevant level of metabolism. Animal models will also be useful in establishing whether the Y333 and N333 variants of Cgr2 lead to physiologically distinct outcomes *in vivo*. Although Cgr2 N333 was less active than Cgr2 Y333 towards digoxin in culture and *in vitro*, it

remains to be seen whether this observation will translate to different digoxin pharmacokinetics *in vivo*. These studies will help to establish the degrees of sensitivity (threshold) and accuracy (nucleotide-level precision) required to use *cgr2* as a clinical biomarker for digoxin inactivation.

Finally, it will be important to determine how microbial metabolism affects the utility of digoxin in non-cardiac indications (Section **1.4.2**), including cancer, HIV, and autoimmune diseases (*6-9, 34*). In some of these therapeutic contexts, digoxin is thought to mediate a beneficial response through inhibition of Na^+/K^+ ATPases, which activates downstream signaling pathways (*9*). However, isoforms of this enzymes are differentially expressed in particular tissues or disease states (*35, 36*) (*37*) and can activate different host responses. Although dihydrodigoxin is consistently less potent than digoxin, these compounds do not have the same binding affinities across all Na^+/K^+ ATPase isoforms (*35*). In addition, digoxin has been shown to interact with and inhibit new targets, including the transcription factor RORγt which mediates autoimmune and inflammatory responses (*8*), DNA double strand break (DSB) repair proteins (*34*), and the transcription factor HIF-1α which is overexpressed in many cancers (*6*). In light of the renewed interest in digoxin therapy, it will be critical to assess the efficacy of both digoxin and its microbial metabolite towards the relevant Na^+/K^+ ATPase isoforms (*37*) or new targets (*8, 34*) in these disease contexts.

5.4. Examining the broader relevance of gut microbes to pharmaceutical metabolism

The human gut microbiota is currently known to transform over 50 pharmaceuticals spanning many indications and host targets into metabolites with altered pharmacological properties (*10, 38*). However, the examples of xenobiotic metabolism surveyed in Section **1.3** likely represent a small fraction of the transformations taking place in this microbial community. The human gut microbiome encodes a broad diversity of enzymes, many of which are exclusively microbial, thus expanding the repertoire of metabolic reactions occurring within the human body. Although clinical studies have revealed marked interindividual variability in these microbial transformations (*39, 40*), most of these reactions have incompletely understood consequences. However, gut microbial xenobiotic metabolites are known to have altered bioactivity, bioavailability, and toxicity, and can interfere with the activities of human xenobiotic-metabolizing enzymes to affect the fates of other ingested molecules (Section **1.3**). Despite the diverse and physiologically significant consequences of these modifications, relatively few mechanistic details are known. The future challenges lie in identifying the relevant microorganisms, genes, and enzymes involved in both known and currently unappreciated metabolic pathways involving pharmaceuticals, elucidating the effects of this chemistry on both host and microbiota, and assessing the presence and variability of these pathways within the human population.

5.4.1 Strategies and considerations for discovering new gut microbial-xenobiotic interactions

As discussed in Section **1.3**, most examples of gut microbial drug metabolism have associated individual activities with the whole gut community. Such transformations were identified as early as the 1950s in humans, animal models, fecal samples, and individual microbes. In these studies, changes in metabolism in the absence of microbes (e.g. germ-free animals) or upon microbial perturbation (e.g. antibiotic treatment or dietary modulation) indicated the gut microbiota's involvement in xenobiotic processing (Figure 67) (*38, 41*). A central challenge moving forward is to connect these examples of gut microbial drug metabolism to specific organisms, genes, and enzymes. This knowledge will enable us to understand the extent to which metabolism varies among individuals, the mechanisms by which these

microbial activities influence human biology, and how we may rationally manipulate these reactions to improve human health. Surveys of human gut isolates, such as Human Microbiome Project (HMP) reference strains (*42*) or culturing from human fecal samples can identify individual organisms with metabolic capabilities of interest. Examining intact communities using approaches like stable isotope probing, fluorescence *in situ* hybridization, and imaging mass spectrometry, can also be used to detect individual cells that metabolize particular compounds (*43*). Coupled with single cell genomics, these methods may aid investigations of xenobiotic-processing activities associated with gut microbes that are challenging to cultivate (*44*).

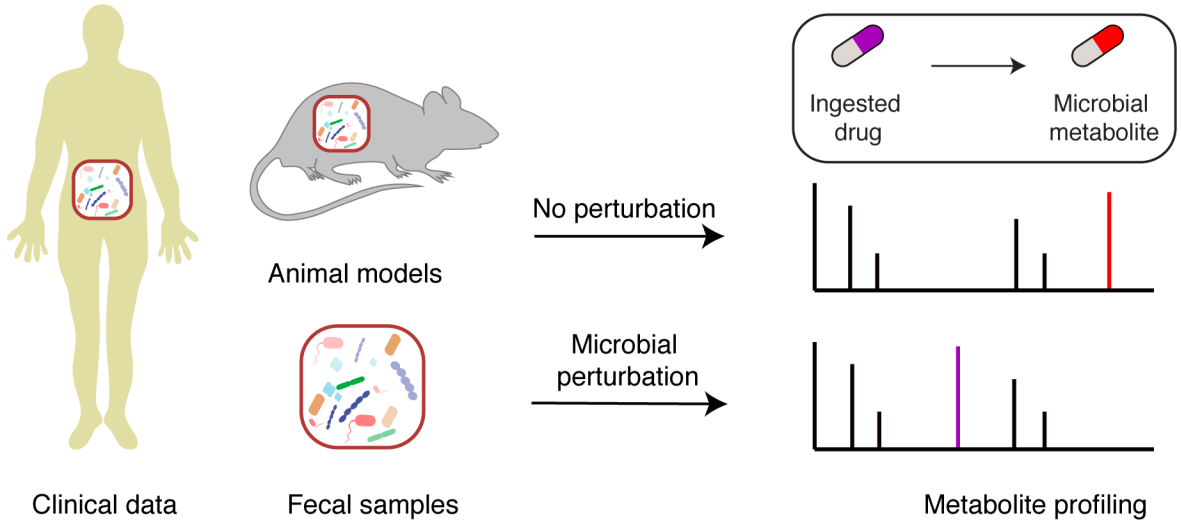


Figure 67: Approaches to linking the gut microbial community to drug metabolism.
The gut microbiota is implicated in drug metabolism when novel metabolites are detected in patients, animals, or microbial communities that were subject to microbial perturbations, such as dietary or antibiotic interventions.

Both traditional approaches (chemical or transposon mutagenesis and activity-guided protein purification) and more recent strategies (rational genome mining, comparative genomics, RNA-Seq, and functional metagenomics) may be used to link microbial metabolic activities with genes (*31, 45, 46*). However, a lack of genetic tools for many gut microbes and the narrow range of heterologous hosts available for functional characterization still limit progress toward this objective. An appreciation of the chemical reactivity required for xenobiotic degradation is thus vital for informing genome and metagenome mining, as well as interpreting and rationalizing results from other methods. Microbes in which enzymes and metabolic pathways are better studied, including pathogens and environmental strains involved in bioremediation (*47*), may also offer clues to support microbiota-focused discovery efforts.

Metabolomics can also play a key role in illuminating currently unappreciated activities, either through surveying individual gut isolates for their ability to consume various xenobiotics or through studying more complex systems, including human fecal samples, model organisms, and patients (*48-50*). These latter endeavors are critical, as metabolic processes in this community may depend on other microbial and host activities (*51, 52*). Metabolomics experiments may also enable better characterization of complex ingested substances (dietary components, herbal supplements, and traditional medicines), to reveal new substrates for gut microbial metabolism, and assess the influence of these components on the metabolism of pharmaceuticals.

Finally, deciphering how gut microbial transformation of xenobiotics affects host health will require integrating clinical studies with mechanistic experiments in model systems and organisms. Data from patients will be particularly important for identifying connections between microbial xenobiotic metabolism and host biology. Existing epidemiological studies linking diet or xenobiotic exposure to health outcomes may be reexamined with gut microbial participation in mind, and new or ongoing clinical studies should incorporate the gut microbiota as a key element of study design. Integrating -omics methods (metabolomics, proteomics, genomics, transcriptomics) (*50, 53*) into these efforts will help to connect xenobiotics, microbial functions, and host health. Such studies will require identification of microbial genes and/or metabolites that are reliable, diagnostic markers for the activities of interest.

5.4.2 Moving from clinical observation to biomarkers of drug metabolism

With recent advances in sequencing, culturing, genetic, computational, and analytical techniques (*54-57*), scientists are now poised to identify and characterize the gut microorganisms, genes, and enzymes that influence human health. The studies presented in this thesis provide a roadmap for discovering and mechanistically characterizing novel gut microbial-xenobiotic interactions, which may be selected and prioritized from the catalog of known clinically relevant examples (Section **1.3**), or from newly identified metabolites of drug candidates (Section **5.4.1**). Isolation of individual xenobiotic-metabolism strains is a crucial first step towards uncovering the genetic and biochemical bases of these transformations. Starting from complex microbial communities (e.g. human fecal samples), individual strains can be selectively enriched (*58, 59*) or isolated (*1*) to enable screening and identification of microbes that metabolize a xenobiotic of interest. The observation that many xenobiotic-processing or transporting genes are only upregulated in the presence of substrate can be leveraged to identify xenobiotic-metabolizing genes. For example, RNA-seq (*31*) or native protein purification (*58*) can be performed on metabolizing strains grown in the presence or absence of a drug to identify genes or enzymes that are likely to be involved in substrate processing. Genetic and/or heterologous expression techniques can then help to validate the role of the identified genes and enzymes in drug metabolism. Finally, the responsible genes can serve as candidate biomarkers to probe the distribution and potential for drug metabolism in human samples or in relevant clinical populations.

5.4.3 Incorporating gut microbial metabolism into drug design and development

Elucidating the molecular basis for gut microbial xenobiotic metabolism and the mechanisms by which these activities influence host biology holds enormous promise for improving human health. Such information will not only help to predict how individual subjects will metabolize ingested compounds, but will also be critical for developing approaches to control these activities. Such techniques could become an instrumental part of drug design and development, as well as personalized medicine.

It is possible that the idiosyncratic toxicity and lack of efficacy that derail approved drugs (*60, 61*) and many late-stage clinical candidates may arise from gut microbial metabolic activities present in a fraction of subjects. Such processes may be identified earlier in development by screening for microbial metabolism. Just as assays with panels of cytochrome P450s and hepatocytes can identify likely human metabolites of current drug candidates, screening fecal samples, gut microbial isolates, or collections of microbial enzymes could reveal additional modifications that alter efficacy or toxicity (Figure 68A) (*62, 63*). While this type of screening has not been widely adopted, several pharmaceutical companies have begun to incorporate microbiota screening (e.g. with human fecal suspensions) into the drug development process to identify novel microbial metabolites of drug candidates (*63-65*).

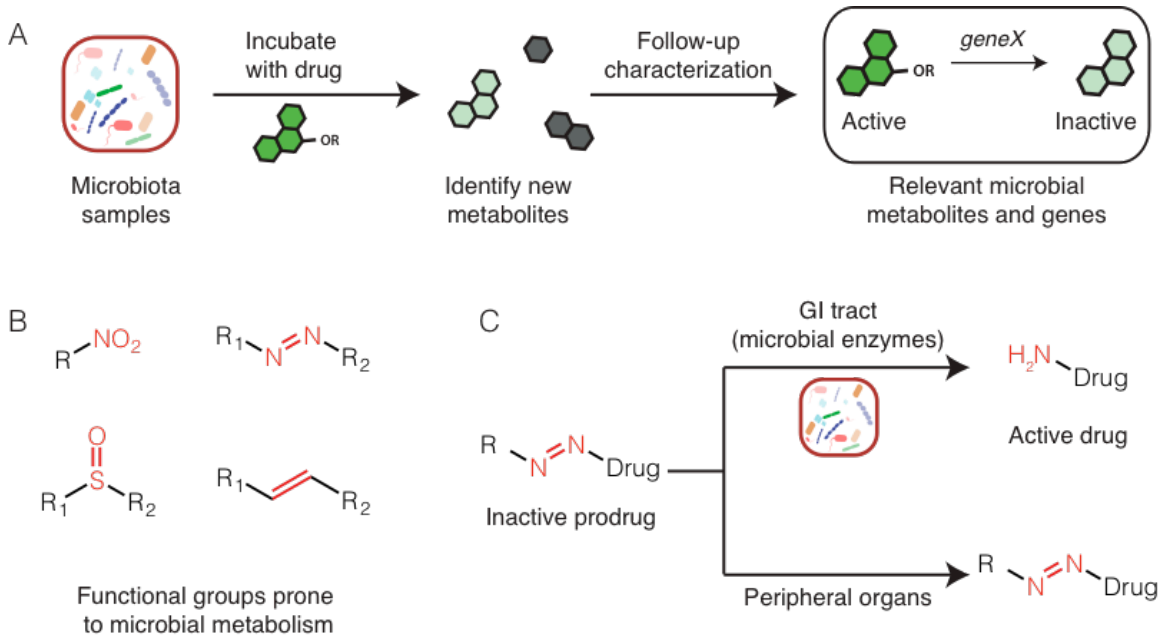


Figure 68: Incorporating microbial metabolism into drug discovery and development. (A) Microbial samples from patients may be incubated with drug candidates to identify and prioritize novel metabolites for further characterization. (B) Functional groups that are prone to gut microbial metabolism may be removed or replaced by bioisosteres to limit metabolism in humans, or (C) incorporated into prodrug scaffolds to afford selective activation in the GI tract and limit side effects in peripheral organs.

Understanding which functional groups are prone to microbial metabolism would also allow medicinal chemists to avoid these structural features or to replace reactive functionality with resistant bioisosteres, analogous to common drug modifications (e.g. fluorination) that limit metabolism by cytochrome P450s (Figure 68B) (*66*). Alternatively, these microbially labile functional groups can be leveraged to generate prodrugs that are selectively activated within the GI tract, potentially reducing systemic side effects (Figure 68C) (*67, 68*). Gut microbial enzymes that mediate harmful metabolic activities may also represent new classes of drug targets (*69, 70*).

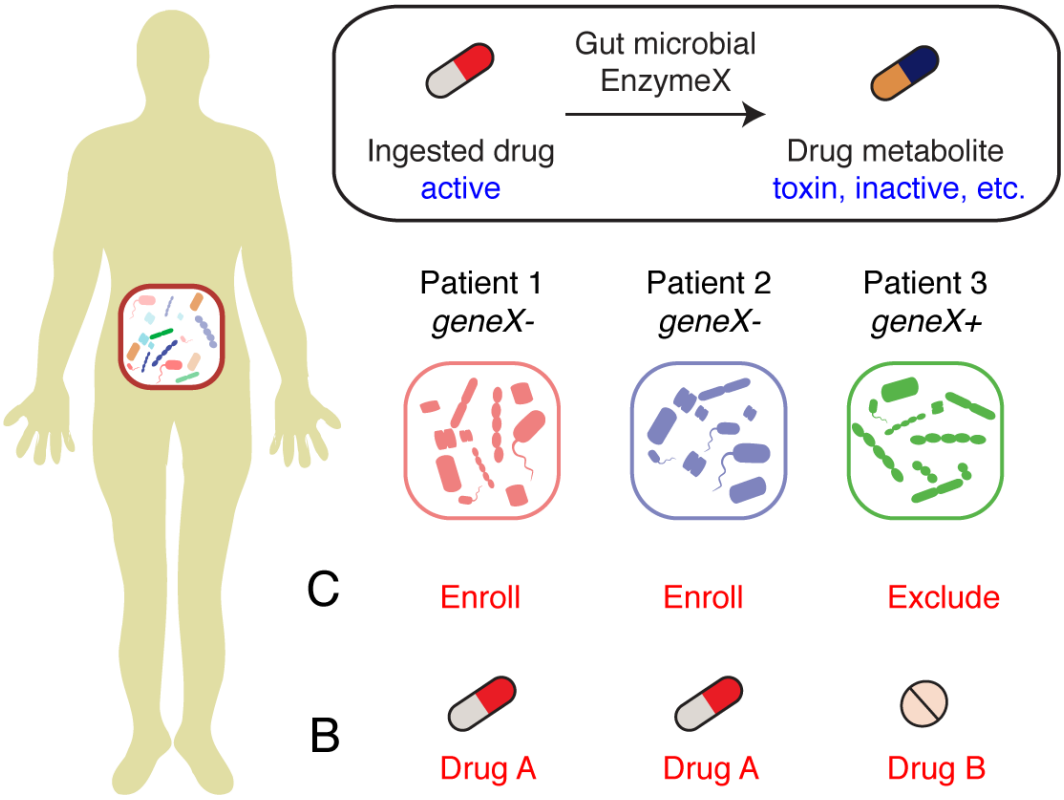


Figure 69: Biomarkers of microbial drug metabolism can inform clinical trial design and personalized medicine.
(A) The presence of gut microbial drug-metabolizing genes (e.g. in patient fecal samples) may be used as a criterion for patient enrollment or exclusion from clinical trials. (B) Patients harboring drug-metabolizing microbiotas may be prescribed alternative drugs.

Knowledge of gut microbial drug metabolism may also inform clinical trial design, allowing for selection of patients based on the presence or absence of particular activities and potentially improving outcomes (Figure 69A). Finally, an understanding of how the gut microbiota transforms drugs could shape clinical practice by allowing physicians to screen for detrimental or beneficial activities in order to prescribe the optimal medications (Figure 69B). The use of microbial genes as biomarkers would be analogous to "molecular typing" of tumors, which has caused a paradigm shift in the administration of cancer therapeutics (*71*). Although we are undoubtedly far from understanding the full metabolic potential of human associated microbes (*72*), continued research in this field can uncover novel chemistry, shed light on complex host-microbial evolutionary interactions, and influence drug development and design to ultimately improve human health.

5.5. Experimental

EPR spectroscopy of Cgr2 in the presence of digoxin and FAD

Samples were prepared in 50 mM HEPES, 100 mM NaCl, pH 8 under oxygen-free conditions in an anaerobic glovebox (Mbraun). Measurements were performed on wild-type Cgr2 that had been reconstituted with 8 equivalent of Fe and S overnight as described in Section **3.8**. Samples 1 and 2 corresponding to Figure 66A and Figure 66B contained 200 μM Cgr2 and either 1.5 % (w/w) DMF or 400 μM digoxin in DMF (2 equivalents), respectively. Each sample was reacted with an excess of sodium dithionite (15 equivalents) for 20 minutes at 22 °C. Samples (250 μL) were loaded into 250 mm length, 4 mm medium wall diameter Suprasil EPR tubes (Wilmad LabGlass) and frozen in liquid N_2 under oxygen-free conditions. EPR spectra were acquired on a Bruker E500 Elexsys continuous wave (CW) X-Band spectrometer (operating at approx. 9.38 GHz) equipped with a rectangular resonator (TE102) and a continuous-flow cryostat (Oxford 910) with a temperature controller (Oxford ITC 503). The spectra were recorded at 10 K, using a microwave power of 0.2 mW, a modulation amplitude of 0.6 mT, a microwave frequency of 9.38 GHz, a conversion time of 82.07 ms, and a time constant of 81.92 ms.

After recording the EPR spectrum of sample 2 (from above), the EPR tube was brought back into the anaerobic chamber and allowed to thaw at 22 °C. 50 μL of 1 mM FAD was added to the EPR tube and flicked to mix. The 300 μL reaction containing 2 equivalents of digoxin (333 μM) and 1 equivalent of FAD (167 μM) was capped and refrozen in liquid N_2 approximately 1 minute after addition of FAD (Figure 66C). The EPR spectrum was recorded as described in the previously described, and then brought back into the chamber. After thawing, the tube was inverted several times and incubated for approximately 20 minutes at 22 °C. The tube was frozen again, and a final EPR spectrum was recorded as previously described (Figure 66D).

5.6. References

1. J. F. Dobkin, J. R. Saha, V. P. Butler, H. C. Neu, J. Lindenbaum, Digoxin-inactivating bacteria: Identification in human gut flora. *Science* **220**, 325-327 (1983).

2. E. Watson, D. R. Clark, S. R. Kalman, Identification by gas chromatography-mass spectroscopy of dihydrodigoxin- a metabolite of digoxin in man. *Journal of Pharmacology and Experimental Therapeutics* **184**, 424-431 (1973).

3. D. R. Clark, S. R. Kalman, Dihydrodigoxin: a common metabolite of digoxin in man. *Drug Metab Dispos* **2**, 148-150 (1974).

4. T. W. Smith, V. P. Butler, E. Haber, Determination of therapeutic and toxic serum digoxin concentrations by radioimmunoassay. *N Engl J Med* **281**, 1212-1216 (1969).

5. M. Gheorghiade, K. F. Adams, W. S. Colucci, Digoxin in the management of cardiovascular disorders. *Circulation* **109**, 2959-2964 (2004).

6. J. Lin *et al.*, A pilot phase II Study of digoxin in patients with recurrent prostate cancer as evident by a rising PSA. *Am J Cancer Ther Pharmacol* **2**, 21-32 (2014).

7. F. Kayali, M. A. Janjua, D. A. Laber, D. Miller, G. Kloecker, Phase II trial of second-line erlotinib and digoxin for nonsmall cell lung cancer (NSCLC). *Open Access Journal of Clinical Trials* **3**, 9-13 (2011).

8. J. R. Huh *et al.*, Digoxin and its derivatives suppress TH17 cell differentiation by antagonizing RORγt activity. *Nature* **472**, 486-490 (2011).

9. R. W. Wong, A. Balachandran, M. A. Ostrowski, A. Cochrane, Digoxin suppresses HIV-1 replication by altering viral RNA processing. *PLoS Pathog* **9**, e1003241 (2013).

10. P. Spanogiannopoulos, E. N. Bess, R. N. Carmody, P. J. Turnbaugh, The microbial pharmacists within us: a metagenomic view of xenobiotic metabolism. *Nat Rev Microbiol* **14**, 273-287 (2016).

11. N. Koppel, V. M. Rekdal, E. P. Balskus, Chemical transformation of xenobiotics by the human gut microbiota. *Science* **356**, eaag2770 (2017).

12. N. Koppel, E. P. Balskus, Exploring and understanding the biochemical diversity of the human microbiota. *Cell Chemical Biology* **23**, 19-30 (2016).

13. M. E. Pandelia, N. D. Lanz, S. J. Booker, C. Krebs, Mossbauer spectroscopy of Fe/S proteins. *Biochim Biophys Acta* **1853**, 1395-1405 (2015).

14. C. Ayala-Castro, A. Saini, F. W. Outten, Fe-S cluster assembly pathways in bacteria. *Microbiol Mol Biol Rev* **72**, 110-125 (2008).

15. K. S. Hewitson *et al.*, The iron-sulfur center of biotin synthase: site-directed mutants. *J Biol Inorg Chem* **7**, 83-93 (2002).

16. S. E. Iismaa, Vazquez, A.E., Jensen, G.M., Stephens, P.J., Butt, J.N., Armstrong, F.A., Burgess, B.K., Site-directed mutagenesis of Azotobacter vinelandii ferredoxin I. Changes in [4Fe-4S] cluster reduction potential and reactivity. *Journal of Biological Chemistry* **266**, 21563-21571 (1991).

17. A. E. Martin, Burgess, B.K., Stout. C.D., Cash, V.L., Deans, D.R., Jensen, G.M., Stephens, P.J., Site-directed mutagenesis of Azotobacter vinelandii ferredoxin I: [Fe-S] cluster-driven protein rearrangement. *Proc Natl Acad Sci U S A* **87**, 598-602 (1990).

18. E. O'Brien *et al.*, The [4Fe4S] cluster of human DNA primase functions as a redox switch using DNA charge transport. *Science* **355**, eaag1789 (2017).

19. P. Sledz *et al.*, New surface contacts formed upon reductive lysine methylation: improving the probability of protein crystallization. *Protein Sci* **19**, 1395-1404 (2010).

20. Z. S. Derewenda, The use of recombinant methods and molecular engineering in protein crystallography. *Methods* **34**, 354-363 (2004).

21. A. F. Moon, G. A. Mueller, X. Zhong, L. C. Pedersen, A synergistic approach to protein crystallization: combination of a fixed-arm carrier with surface entropy reduction. *Protein Sci* **19**, 901-913 (2010).

22. R. Linding *et al.*, Protein disorder prediction: implications for structural proteomics. *Structure* **11**, 1453-1459 (2003).

23. C. Sanders, H. Lill, Expression of prokaryotic and eukaryotic cytochromes *c* in *Escherichia coli*. *Biochim Biophys Acta* **1459**, 131-138 (2000).

24. L. Thöny-Meyer, Cytochrome *c* maturation: a complex pathway for a simple task? *Biochemical Society Transactions* **30**, 633-638 (2002).

25. J. Simon, L. Hederstedt, Composition and function of cytochrome *c* biogenesis System II. *FEBS J* **278**, 4179-4188 (2011).

26. K. Ozawa, F. Yasukawa, Y. Fujiwara, H. Akutsu, A simple, rapid, and highly efficient gene expression system for multiheme cytochromes *c*. *Biosci Biotechnol Biochem* **65**, 185-189 (2001).

27. K. Ozawa, A. I. Tsapin, K. H. Nealson, M. A. Cusanovich, H. Akutsu, Expression of a tetraheme protein, *Desulfovibrio vulgaris* Miyazaki F cytochrome *c*(3), in *Shewanella oneidensis* MR-1. *Appl Environ Microbiol* **66**, 4168-4171 (2000).

28. M. Kern, J. Simon, Production of recombinant multiheme cytochromes *c* in *Wolinella succinogenes*. *Methods Enzymol* **486**, 429-446 (2011).

29. A. Takagaki, Y. Kato, F. Nanjo, Isolation and characterization of rat intestinal bacteria involved in biotransformation of (-)-epigallocatechin. *Arch Microbiol* **196**, 681-695 (2014).

30. E. Saunders *et al.*, Complete genome sequence of *Eggerthella lenta* type strain (IPP VPI 0255). *Stand Genomic Sci* **1**, 174-182 (2009).

31. H. J. Haiser *et al.*, Predicting and manipulating cardiac drug inactivation by the human gut bacterium *Eggerthella lenta*. *Science* **341**, 295-298 (2013).

32. V. I. Mathan, J. Wiederman, J. F. Dobkin, J. Lindenbaum, Geographic differences in digoxin inactivation, a metabolic activity of the human anaerobic gut flora. *Gut* **30**, 971-977 (1989).

33. A. N. Alam, J. R. Saha, J. F. Dobkin, J. Lindenbaum, Interethnic variation in the metabolic inactivation of digoxin by the gut flora. *Gastroenterology* **95**, 117-123 (1988).

34. Y. V. Surovtseva *et al.*, Characterization of cardiac glycoside natural products as potent inhibitors of DNA double-strand break repair by a whole-cell double immunofluorescence assay. *J Am Chem Soc* **138**, 3844-3855 (2016).

35. K. M. Weigand *et al.*, Na(+),K(+)-ATPase isoform selectivity for digitalis-like compounds is determined by two amino acids in the first extracellular loop. *Chem Res Toxicol* **27**, 2082-2092 (2014).

36. A. Katz *et al.*, Selectivity of digitalis glycosides for Iisoforms of human Na,K-ATPase. *Journal of Biological Chemistry* **285**, 19582-19592 (2010).

37. J. Muller-Ehmsen, J. Wang, R. H. Schwinger, A. A. McDonough, Region specific regulation of sodium pump isoform and Na,Ca-exchanger expression in the failing human heart-right atrium vs left ventricle. *Cell Mol Biol* **47**, 373-381 (2001).

38. T. Sousa *et al.*, The gastrointestinal microbiota as a site for the biotransformation of drugs. *Int J Pharm* **363**, 1-25 (2008).

39. R. A. van Hogezand *et al.*, Bacterial acetylation of 5-aminosalicylic acid in faecal suspensions cultured under aerobic and anaerobic conditions. *Eur J Clin Pharmacol* **43**, 189-192 (1992).

40. J. G. Nutt, N. H. Holford, The response to levodopa in Parkinson's disease: imposing pharmacological law and order. *Ann Neurol* **39**, 561-573 (1996).

41. H. Danielsson, B. Gustafsson, On serum-cholesterol levels and neutral fecal sterols in germ-free rats - bile acids and steroids. *Arch Biochem Biophys* **83**, 482-485 (1959).

42. Human Microbiome Jumpstart Reference Strains Consortium, A catalog of reference genomes from the human microbiome. *Science* **328**, 994-999 (2010).

43. D. Berry *et al.*, Host-compound foraging by intestinal microbiota revealed by single-cell stable isotope probing. *Proc Natl Acad Sci U S A* **110**, 4720-4725 (2013).

44. R. S. Lasken, Genomic sequencing of uncultured microorganisms from single cells. *Nat Rev Microbiol* **10**, 631-640 (2012).

45. S. Craciun, E. P. Balskus, Microbial conversion of choline to trimethylamine requires a glycyl radical enzyme. *Proc Natl Acad Sci U S A* **109**, 21307-21312 (2012).

46. L. Tasse *et al.*, Functional metagenomics to mine the human gut microbiome for dietary fiber catabolic enzymes. *Genome Res* **20**, 1605-1612 (2010).

47. C. S. Karigar, S. S. Rao, Role of microbial enzymes in the bioremediation of pollutants: a review. *Enzyme Res* **2011**, 805187 (2011).

48. W. R. Wikoff *et al.*, Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proc Natl Acad Sci U S A* **106**, 3698-3703 (2009).

49. A. M. Aura *et al.*, Drug metabolome of the simvastatin formed by human intestinal microbiota *in vitro*. *Mol Biosyst* **7**, 437-446 (2011).

50. Z. Wang *et al.*, Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **472**, 57-63 (2011).

51. F. P. Martin *et al.*, A top-down systems biology view of microbiome-mammalian metabolic interactions in a mouse model. *Mol Syst Biol* **3**, 112 (2007).

52. M. Li *et al.*, Symbiotic gut microbes modulate human metabolic phenotypes. *Proc Natl Acad Sci U S A* **105**, 2117-2122 (2008).

53. D. Zeevi *et al.*, Personalized nutrition by prediction of glycemic responses. *Cell* **163**, 1079-1094 (2015).

54. A. L. Goodman *et al.*, Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. *Proc Natl Acad Sci U S A* **108**, 6252-6257 (2011).

55. B. Lim, M. Zimmermann, N. A. Barry, A. L. Goodman, Engineered Regulatory Systems Modulate Gene Expression of Human Commensals in the Gut. *Cell* **169**, 547-558 e515 (2017).

56. L. Ma *et al.*, Gene-targeted microfluidic cultivation validated by isolation of a gut bacterium listed in Human Microbiome Project's Most Wanted taxa. *Proc Natl Acad Sci U S A* **111**, 9768-9773 (2014).

57. M. H. Sarafian *et al.*, Bile acid profiling and quantification in biofluids using ultra-performance liquid chromatography tandem mass spectrometry. *Anal Chem* **87**, 9662-9670 (2015).

58. T. Kumano, E. Fujiki, Y. Hashimoto, A. Kobayashi, Discovery of a sesamin-metabolizing icroorganism and a new enzyme. *Proc Natl Acad Sci U S A* **113**, 9087-9092 (2016).

59. A. Martinez-del Campo *et al.*, Characterization and detection of a widely distributed gene cluster that predicts anaerobic choline utilization by human gut bacteria. *MBio* **6**, (2015).

60. H. Okuda, K. Ogura, A. Kato, H. Takubo, T. Watabe, A possible mechanism of eighteen patient deaths caused by interactions of sorivudine, a new antiviral drug, with oral 5-fluorouracil prodrugs. *J Pharmacol Exp Ther* **287**, 791-799 (1998).

61. J. H. Carter, M. A. Mclafferty, P. Goldman, Role of the gastrointestinal microflora in amygdalin (Laetrile)-induced cyanide toxicity. *Biochem Pharmacol* **29**, 301-304 (1980).

62. H. Okuda *et al.*, Lethal drug interactions of sorivudine, a new antiviral drug, with oral 5-fluorouracil prodrugs. *Drug Metabolism and Disposition* **25**, 269-273 (1997).

63. M. McCabe *et al.*, Defining the role of gut bacteria in the metabolism of deleobuvir: *In vitro* and *in vivo* studies. *Drug Metab Dispos* **43**, 1612-1618 (2015).

64. R. S. Sane *et al.*, Contribution of major metabolites toward complex drug-drug interactions of deleobuvir: *In vitro* predictions and *in vivo* outcomes. *Drug Metab Dispos* **44**, 466-475 (2016).

65. Y. Li, J. Xu, W. G. Lai, A. Whitcher-Johnstone, D. J. Tweedie, Metabolic switching of BILR 355 in the presence of ritonavir. II. Uncovering novel contributions by gut bacteria and aldehyde oxidase. *Drug Metab Dispos* **40**, 1130-1137 (2012).

66. H. J. Bohm *et al.*, Fluorine in medicinal chemistry. *Chembiochem* **5**, 637-643 (2004).

67. K. Lavrijsen *et al.*, Reduction of the prodrug loperamide oxide to its active drug loperamide in the gut of rats, dogs, and humans. *Drug Metab Dispos* **23**, 354-362 (1995).

68. M. A. Peppercorn, P. Goldman, Role of intestinal bacteria in metabolism of salicylazosulfapyridine. *Journal of Pharmacology and Experimental Therapeutics* **181**, 555-562 (1972).

69. B. D. Wallace *et al.*, Alleviating cancer drug toxicity by inhibiting a bacterial enzyme. *Science* **330**, 831-835 (2010).

70. K. S. Saitta *et al.*, Bacterial beta-glucuronidase inhibition protects mice against enteropathy induced by indomethacin, ketoprofen or diclofenac: mode of action and pharmacokinetics. *Xenobiotica* **44**, 28-35 (2014).

71. S. E. Jackson, J. D. Chester, Personalised cancer medicine. *Int J Cancer* **137**, 262-266 (2015).

72. Human Microbiome Project Consortium, Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207-214 (2012).